



Universiteit
Leiden
The Netherlands

Word order, information structure and agreement in Teke-Kukuya Li, Z.

Citation

Li, Z. (2024, September 5). *Word order, information structure and agreement in Teke-Kukuya*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/4054947>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4054947>

Note: To cite this publication please use the final published version (if applicable).

Word order, information structure and
agreement in Teke-Kukuya

Published by

LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461
e-mail: lot@uva.nl
<http://www.lotschool.nl>

Cover illustration: The main road of Lékana District, by Zhen Li

ISBN: 978-94-6093-461-2
DOI: <https://dx.medra.org/10.48273/LOT0677>
NUR: 616

Copyright © 2024 Zhen Li. All rights reserved.

Word order, information structure and agreement in Teke-Kukuya

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 5 september 2024
klokke 11.30 uur

door

Zhen Li

geboren te Tianjin, People's Republic of China
in 1994

Promotores: dr. G. J. van der Wal
prof. dr. M. Mous

Promotiecommissie: prof. dr. L. L.-S. Cheng
prof. dr. K. Bostoen
(Ghent University)
prof. dr. L. Marten
(SOAS, University of London)
prof. dr. J.-W. Zwart
(University of Groningen)

To my grandfather
献给我的爷爷李联合

Contents

1	Introduction	1
1.1	Research questions	1
1.2	Teke-Kukuya: language and people	3
1.3	Methodology and data	6
1.4	Outline of the thesis	7
2	A grammar sketch of Teke-Kukuya	11
2.1	Segmental phonology	12
2.1.1	Consonants	13
2.1.2	Vowels	24
2.1.3	Syllable structure	30
2.2	Prosody	33
2.2.1	Tone	33

2.2.2	Stem-initial accent	45
2.3	Nominal morphology	49
2.3.1	Noun classes	49
2.3.2	Noun phrase	67
2.4	Verbal morphology and TAM	94
2.4.1	Verbal affixes	95
2.4.2	TAM conjugations	108
2.5	Clause structure	117
2.5.1	Complementisers	117
2.5.2	Relative constructions	122
2.5.3	Copula and non-verbal predication	134
3	Word order and topic/focus expressions	137
3.1	Canonical word order	138
3.1.1	SVO as canonical word order	138
3.1.2	Focus expressions in SVO	140
3.1.3	Thetic sentences	144
3.2	Dedicated IBV focus position	148
3.2.1	Argument and adjunct focus in IBV	149
3.2.2	Predicate(-centered) focus and IBV	161
3.2.3	IBV as a dedicated focus position	173

3.2.4	Interpretational properties of IBV focus	179
3.3	Topical elements in the preverbal domain	187
3.3.1	Multiple topics in the preverbal domain	187
3.3.2	Functional passives	198
3.4	Cleft constructions	209
4	On the origin of IBV focus	215
4.1	Hypothesis: IBV focus originates from a cleft	217
4.2	Subject marking alternation in IBV focus and relatives . . .	219
4.2.1	Subject marking alternation in IBV focus	219
4.2.2	Subject marking in relative (cleft) constructions .	228
4.3	More tonal evidence on the connection between IBV focus and cleft	237
4.4	Origin and grammaticalisation of IBV focus	251
4.4.1	IBV focus: mono-clausal or bi-clausal?	251
4.4.2	From cleft to IBV focus: the grammaticalisation path	256
4.5	Microvariation in preverbal focus in West-Coastal Bantu . .	264
5	The syntax of IBV focus and subject marking asymmetry	271
5.1	Structural representation of SVO	275
5.1.1	Position of the verb	275
5.1.2	Position of the subject in SVO	281

5.2	Where is the IBV focus position?	289
5.2.1	IBV focus as A-bar syntax	289
5.2.2	Structural position of the IBV element	294
5.3	Subject marking asymmetry and IBV focus	306
5.3.1	Agreement in subject and non-subject relatives	306
5.3.2	Featural account of the subject marking asymmetry	320
5.4	Subject marking alternation and IBV focus	337
5.4.1	The derivation of cleft constructions	338
5.4.2	The derivation of the IBV focus construction	347
5.5	Summary and remaining questions	351
6	Conclusion	357
6.1	Main findings in the thesis	357
6.2	Remaining research questions	360
	Appendix	365
	Bibliography	375
	English summary	401
	Samenvatting in het Nederlands	405

Acknowledgements

This dissertation would never have been accomplished without the invaluable help and support of so many people throughout my PhD journey. First and foremost, I would like to express my deepest thanks to my supervisor Jenneke van der Wal, without whom I would never have had the chance to conduct PhD research on Bantu linguistics. Her meticulous guidance and insightful feedback during these five years have been instrumental in shaping this research, and her methods of conducting linguistic research and organizing things will always be an admirable example in my academic life. I also feel incredibly grateful for having my promotor Maarten Mous, whose expertise and sensitivity in African languages are always taken for granted by me as an inexhaustible treasure. I thank his foresight in suggesting Teke-Kukuya for my PhD project, which turned out to be a quite intriguing language and a good starting point for my future research. He has always been at the helm in instructing me to write this dissertation, especially the grammar sketch, throughout the time. I express special thanks to my “only” local colleague in the BaSIS project, Elizabeth Kerr, for always sharing shining insights on my research and for her companionship during all the (online) meetings, research trips, and conferences/talks, which made me never feel lonely. I also thank the BaSIS project for its intellectual and financial support, and all the collaborators from Africa who I admire so much: Allen Asiimwe, Patrick Kanampiu, Ernest Nshemezimana, Amani Lusekelo and Nelsa Nhantumbo. I will miss the BaSIS tea each week, the

numerous Skype calls during the COVID-19 pandemic, and our gathering camp in Malawi.

The most unforgettable periods of my PhD journey were the two field-work trips in the Republic of Congo, which would not have been the same without the people who offered me kind help during those nine months. I thank all my consultants from Lékana and Brazzaville who dedicated their time to sharing their knowledge of the Kukuya language with me: Gilbert Mbou, Gabriel Ntsiebele, Zacharie Ngouloubi, Ngolo Martin and Alain Mbiambourou. I appreciate Gilbert for the drives and dinners he offered and Zacharie for all his words as a kind grandfather to me. I also thank Senhua Zhang for generously hosting me in his home-like hotel in Brazzaville and Yongkang Wang for introducing local students to me from the Confucius Institute. My special thanks go to Ruth Rahary in SIL Congo who introduced the Kukuya speakers to me and shared her outstanding expertise on Teke languages in our productive discussions.

I love starting my day with a venti cup of coffee from the common room in Reuvenplaats, but what really motivated me to frequently go to the workplace was being together with my fellow linguists at LUCL. I enjoyed having lunch with my colleagues on workdays while chatting about our research and everyday news. I would like to thank Jiang Wu for all the conversations we had online and in person on the practical and nonsense stuff in life, as well as his encouragement and comfort over the years. I also feel grateful for the camaraderie and all the help I received from my officemates and coworkers, especially during the lockdown in 2022: Lis Kerr, Qing Yang, Tingting Zheng, Zhuoyi Luo, Jiaqi Wang, Ruoyu Shi, Jian Sun, Ruixue Wu, Yiran Ding, Fei Bai and Priscilla Lam, as well as the rest of them. I thank Timothy Hadjah and Ahmed Sosal for co-organizing the talk series “This Time For Africa”. I cherish the memories with my roommates in Den Haag and Leiden, Jeremy Liu and Feng Xu, for all the enjoyable moments we shared. I also extend my gratitude to Hang Cheng, Han Hu, Menghui Shi, Chen Ran, and Dan Yuan for their excellent examples as senior researchers at the very beginning of my PhD. I also sincerely thank everyone I have met in talks and courses at LUCL or just randomly in the building, for the conversations between us and the enjoyable working atmosphere we created together. Crucially, I thank my former teachers in Beijing and London:

Anshan Li, whose expertise has consistently encouraged me to explore Africa, and Lutz Marten, whose Bantu course gave me a solid foundation for delving into this language family.

Lastly, I want to express my deepest gratitude to my friends and family who have provided me with unwavering support in China. I am thankful to Guohao Yang for his friendship since our childhood, and to Chao Jiang for his companionship throughout these years. The love and support from my parents have always been immeasurable and enshrined in my heart.

Abbreviations

Glosses

AGR	agreement marker
APPL	applicative
ASP	aspect marker
AUG	augment
AUX	auxiliary
CAUS	causative
CF	counterfactual
COMP	complementiser
COND	conditional
CONJ	conjunctive
CONN	connective
COP	copular
DEM	demonstrative
DEP	dependent tense
DET	determiner
DIM	diminutive
DJ	disjoint form
DL	dual
EMP	emphatic
F	feminine
FUT	future

FV	final vowel
GEN	genitive
HAB	habitual aspect
IMP	imperative
IMPF	imperfective
INF	infinitive
LOC	locative
N	neuter
NARR	narrative
NEG	negative
NOM	nominative
PASS	passive
PERF	perfective
PL	plural
POSS	possessive
PRD	predicative
PREP	preposition
PRO	pronoun
PROG	progressive
PROS	prospective
PRS	present
PST	past
REL	relative marker
RFL	reflexive
RPST	remote past
SBJV	subjunctive
SG	singular
SM	subject marker
WH	<i>wh</i> -phrase

CHAPTER 1

Introduction

This chapter first introduces the background and research questions of the thesis. Then I provide the geographic and demographic information of the Teke-Kukuya language, and I also introduce the people who speak the language. The methodology for fieldwork is discussed, and the conventions in the presentation of the data are also mentioned. The overview of each remaining chapter and the scope of the thesis are given in the last section.

1.1 Research questions

Earlier studies on the grammar of Bantu languages tend to focus on the noun class system and verbal morphology. In recent years, syntax has been a growing field in the studies of Bantu languages and the importance of information structure for the analysis of Bantu morphosyntax has been highlighted (van der Wal 2015; Downing and Marten 2019). Information structure (or information packaging) is the way in which speakers structure

2 *Word order, information structure and agreement in Teke-Kukuya*

the speech information they give so that the information can be more easily adapted into the current knowledge of the addressees. Different languages make use of various linguistic strategies for the sake of expressing what is given/known, what is new and/or what is contrasted. In many recently published grammars of Bantu languages, the expressions of information structure of the language have been described in separate chapters, e.g. Crane et al. (2011) and Guérois (2015). Information structure can affect many aspects of the morphosyntax of Bantu languages. The agglutinative verbal morphology such as subject/object marking and conjoint/disjoint alternation in Bantu is always on the boundary with syntax, which is often affected by information packaging. In addition to verbal inflection, word order variation is also commonly associated with discourse functions. In many Bantu languages, focused information and given information can (or tends to) occur in some particular positions in the clause or verb phrase, which is known as topic fronting and the use of dedicated focus positions. The studies on the interaction between syntax and information structure in Bantu have also influenced some theoretical debates on syntactic phenomena such as movement and agreement.

As for dedicated focus positions, the use of an immediate-after-verb (IAV) focus position has been reported for example in Aghem (Watters 1979), Bemba (Costa and Kula 2008), Makhuwa (van der Wal 2009), Matengo (Yoneda 2011) and Zulu (Buell 2006, 2009); and some formal analyses on the structural position of the IAV element have been proposed (Buell 2006, 2009; van der Wal 2006; Cheng and Downing 2012). A sentence-final focus position is attested in Kirundi (Sabimana 1986; Ndayiragije 1999) and Kinyarwanda (Gibson et al. 2017). Notably, there are also some West-Coastal Bantu languages such as Mbuun (Bostoen and Mundeke 2012), Kisikongo (De Kind 2014; De Kind et al. 2015) and Nsong (Koni Muluwa and Bostoen 2014) that have been claimed to have an immediate-before-verb (IBV) focus position. However, the exact formal and functional properties of the IBV focus position and its historical origin still merit more detailed research. The aim of this research is to concentrate on one understudied language in the West-Coastal area, Teke-Kukuya, which also displays this IBV focus position, investigating its functions and origin as well as providing a structural analysis of the IBV focus construction, in order to better understand this focus strategy as an areal feature in a broader sense.

This study is part of the Bantu Syntax and Information Structure (BaSIS) research project¹ which systematically studies the expression of information structure and its influence on nominal licensing in selected Bantu languages. Teke was chosen for the project and for my research, since in previous research the IBV focus position and an associated intriguing class 1 subject marking alternation are reported in some closely related West-Coastal Bantu languages such as Mbuun and Kisikongo, so at the start of the project we were curious about whether similar phenomena can be attested in Teke and how syntax and information structure interact in the presence of the IBV focus strategy. The existence of IBV focus in Teke and the class 1 subject agreement alternation were borne out at the very beginning of my first fieldwork period with elicitation on focus expressions such as interrogatives, and were investigated in detail throughout the rest of the fieldwork. Teke is also a yet insufficiently described Bantu language, so the project also aims at contributing to the description of the language.

Therefore, the goals in this thesis are first to provide a brief grammar description of the language, as well as to address the research questions mentioned above. The research focuses on the expression of information structure in Teke, paying particular attention to the word order variation and the use of the IBV focus strategy. In the major part of the thesis, I investigate the functions of the IBV focus construction, its diachronic development and its structural representation. I also provide a plausible explanation on the class 1 subject agreement alternation that conspires with IBV focus.

1.2 Teke-Kukuya: language and people

Teke is a group of Bantu languages which are spoken in Congo, DRC and Gabon, and is coded as B70 in Guthrie (1948)'s classification. Teke is a Kongo exonym for “the populations who live on the plateaus north of

¹My research was mainly supported by the China Scholarship Council and was also partly funded by the BaSIS project which was supported by NWO Vidi grant 276-78-001.

4 Word order, information structure and agreement in Teke-Kukuya

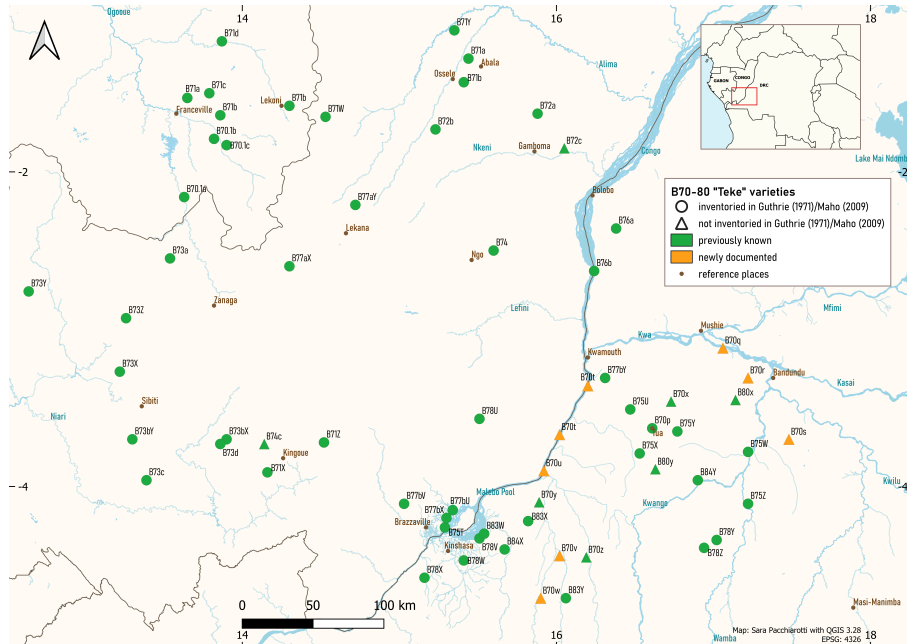


Figure 1.1: Map of the Plateaux region and location of Teke varieties (Kouarata et al. 2023)

Malebo Pool on both sides of the Congo River as far north as the mouth of Nkeni” (Vansina 1966: 102). Teke people usually name their subgroups after the plateau where they live (Pacchiarotti et al. 2019: 16). Since the internal classification of Teke and the labeling of its different varieties still meet a lot of confusion, here I do not list all the Teke varieties in the literature but I refer the readers to Pacchiarotti et al. (2019) and Kouarata et al. (2023) on the internal classification of Teke. A map of the geographic distribution of the known Teke varieties is illustrated in Figure 1.1 (Kouarata et al. 2023), which includes some newly documented varieties/dialects.

I chose the Kukuya variety which is coded as B77a (including B77aX and B77aY in the map above) for my research, because it shows less phonological reduction compared to other Teke varieties, and there has been a detailed study on its phonology (Paulian 1975) and some papers and notes by the same author on this language (Paulian 1997, 1998, 2001), which

allowed me to get familiarised with the language much more easily at the start of my research. During my fieldwork I also found that the speakers are quite aware of how Kukuya differs from other Teke languages on the grammatical features. In the rest of the thesis, I only use “Kukuya” to refer to the Teke-Kukuya language. In my description and analyses on the grammar of Kukuya, I don’t imply that the same holds for other Teke varieties but only make it explicit where I have solid data to prove.

The Kukuya language is mainly spoken around the Lékana district of the Plateau Department in Congo, to the east of the Lékéti river. The language was also referred to as Kukwa, Kikukuya, Koukouya, Kukwa, Kukwa or Küküia in some literature. Kukuya was reported to have about 38,800 native speakers according to Ethnologue (2000) (900,000 speakers of all Teke languages in Congo), but I don’t have a solid number of its current speakers. In Congo, French is the official language of the country, while Lingala and Kituba serve as the *lingua franca* in the northern and southern parts of the country respectively. According to my observation, each person that I made acquaintance with in Congo, no matter in the village or in urban areas, can speak at least three languages, namely French, Lingala or/and Kituba, and their native indigenous languages.

Kukuya is still used actively in daily life in the Lékana district, which is the primary communication language among its community members. It is currently considered as a language in vigorous status (6a*) according to Ethnologue (2024). However, there is an obvious tendency that the younger speakers are losing vocabulary and grammatical knowledge of the language. When I checked some words that are listed in Paulian’s (1975) work, younger speakers were less familiar with many of the words. In Brazzaville, young people are often reluctant to show that they are Teke speakers. Some speakers who have lived in Brazzaville for years also tend to mix the indigenous languages that they know, many words that I consulted with them turned out to be Lingala or Kikongo. So apparently there is risk for Kukuya to be used less and less in the future. The SIL Congo branch has made much effort on documenting the Congolese indigenous languages including Teke, and has worked on and published several orthography guide books, dictionaries, textbooks with exercises on various Teke varieties, but until now there has not been any textbook or grammar on Kukuya yet. For missionary

purposes, three Gospels of the New Testament were translated into Kukuya as “Mandaka Mabve ma Yezu Kristo” by Pierre Loubier and published in June 1979. The book is a good reference on what the language looked like a half century ago, but the transcription was rough and is not always reliable.

There were also some very detailed anthropological works in the 1970s-1980s on the history and society of the Kukuya people; to concentrate on the linguistic part of the thesis I don't present and discuss these here. I refer the readers to the studies of Bonnafé (1973, 1979, 1987).

1.3 Methodology and data

The data presented in this thesis are based on my two fieldwork periods in Lékana and Brazzaville in the Republic of Congo, from May until the end of August 2019, and from April until end of August 2021. The time I spent in Brazzaville and Lékana is almost equal. In Lékana I worked intensively with three local speakers and kept communicating with people who live in the village; in Brazzaville I worked with speakers who were all born in Lékana and were brought up there. My main language of communication in these sites was French. During these periods I collected almost 4000 sentences with grammaticality judgements and I transcribed 6 stories with the help of four main speakers who speak Kukuya as their first language and can speak French very fluently.

Gilbert Mbou was my main consultant in Lékana, with whom I worked for the most of the time in the summer of 2019 as well as in May 2021. He was born in 1967 and was brought up in Lékana where he spent all his previous years. He lives as a farmer and has a family with four children. Gabriel Ntsiebele who was born in the 1950s is his best friend in the village and they had known each other since their childhood. I worked with the two of them every evening after they finish their work in the field. Gilbert also often drove me out to see the village and visit the community members with his motorbike and often treated me with hearty dinners. Zacharie Ngouloubi (1949) and Alain Mbiambourou (1967) were the ones who I worked with

everyday in Brazzaville in July and August 2021. They were introduced to me by the spokesperson of the King of Teke. We worked in a café in the center of Brazzaville with beer and peanuts in the afternoons. They were both born and brought up in Lékana and have a very good sense of the Kukuya language.

My corpus consists of two different types of data: elicited and spontaneous data. For writing the grammar sketch part, I relied much on the elicited data. I checked different grammatical features of the language by asking the speakers to translate French sentences into Kukuya. With the help of the methodological guide “The BaSIS basics of information structure” (Van der Wal 2021) which was developed as part of the BaSIS project, I have collected both elicited and spontaneous data during the fieldwork, but since the marked information-structural expressions such as the use of the dedicated focus position are more obviously attested in elicited sentences, the examples presented in this thesis are largely extracted from the elicited data. I often get generalisations from the speakers’ judgement on the grammaticality and appropriateness of the elicited sentences when I intentionally change the word order and agreement patterns. I also included data from natural speech of different genres such as dialogues and storytelling, but I leave more elaborated investigation on the information structure of spontaneous speech for further research. Moreover, since the speakers are all literate, I also made some written tasks such as question-answer pairs, idioms and letters for them to write down, especially for Zacharie and Alain. All the data mentioned above are transcribed and stored in the OLD database which is accessed via the platform Dative, and will be archived in the Language Archive.

1.4 Outline of the thesis

The thesis is organised as two main parts. The first part is a grammar sketch of the Teke-Kukuya language (chapter 2), and the second part discusses word order variation and expressions of information structure of this language, with particular interests in a dedicated immediate-before-verb

8 *Word order, information structure and agreement in Teke-Kukuya*

(IBV) focus strategy in this language (chapter 3). I provide a diachronic account (chapter 4) as well as a synchronic analysis (chapter 5) of this IBV focus strategy, also discussing the agreement mechanism of the class 1 subject marking alternation that is associated with the syntax of IBV focus.

The grammar sketch in chapter 2 covers the basic grammatical properties of the Teke-Kukuya language, including its segmental phonology and prosodic system, noun classes and noun phrases, verbal morphology and TAM conjugations, as well as some syntactic issues such as complementisers and embedded clauses. The aim of this part is to provide a description on the structure of this language in theoretically neutral terms, for the readers to refer to while reading the thesis and to better understand the other chapters. It also allows readers who are interested in the typology of (West-Coastal) Bantu languages or the other varieties of Teke to use the data to conduct comparative research. Since there is no previous reference grammar of the Kukuya language, this part also contributes to the documentation of the language for the community.

The second part of the thesis consists of three chapters. Chapter 3 presents the influence of information structure on the word order of the language. Some notions of information structure are defined in this chapter. I describe different types of focus and topic expressions, showing how they deviate from the canonical SVO order. The main aim of this chapter is to illustrate the function and interpretation of the IBV focus position which is available for different kinds of argument and adjunct focus. In the rest of the chapter I also introduce different types of topical elements in the pre-verbal domain, functional passive constructions, and cleft constructions for focus expressions. This chapter also allows the readers who are interested in the expressions of information structure of (Bantu) languages to learn about the specific properties of Teke-Kukuya and the IBV focus position for comparative work.

In chapter 4, I investigate the diachronic origin of the IBV focus strategy. I compare the IBV focus construction with clefts, showing that they share many grammatical properties, both segmentally and tonally. I make the hypothesis that the IBV focus originates from a basic cleft, and it has been grammaticalised towards a monoclausal focus construction. I propose

a grammaticalisation pathway of the IBV focus strategy, and introduce its distribution as an areal innovation in some other West-Coastal Bantu languages. This chapter may also offer inspiration to the studies on the IBV focus position in other Bantu languages, especially with the solid tonal data on the connection between IBV focus and relative constructions, which was not taken into account in previous research.

Chapter 5 aims to provide analyses on the structural representation of the IBV focus construction and on the class 1 subject marking alternation. I first illustrate the structural representation of the canonical SVO order and look into the structural position of the IBV focused element. To account for the class 1 subject marking alternation, I start by discussing the subject agreement asymmetry in subject and non-subject relatives, and I propose that the class 1 subject marking allomorphy in IBV focus is connected to the agreement with a clause-internal ϕ P that is co-referential with an initial base-generated topic. This chapter contributes to a structural analysis of the IBV focus strategy in Bantu.

Chapter 6 concludes this thesis and highlights some further research questions. In the appendix I include one glossed and translated Kukuya texts, which is a dialogue between husband and wife.

CHAPTER 2

A grammar sketch of Teke-Kukuya

This chapter is dedicated to an overall grammar sketch of the Teke-Kukuya language. Since previously the language has not been described in detail except for its phonology and noun classes, with this chapter I would also like to contribute to a reference grammar for the Kukuya people and society. The chapter is organised as follows. In section 2.1 I present the segmental phonology of the language, namely the inventories of consonants and vowels, as well as the syllable structure and some phonotactic rules. I describe the prosody system of the language in section 2.2, in which different phenomena associated with tone and accent are discussed. Section 2.3 concentrates on the nominal morphology. The noun class system and some nominal derivation rules are presented first and then the different components of noun phrases. Verbal morphology and the TAM conjugations are sketched in section 2.4. In section 2.5, I illustrate the clause structure of the language, including complement clauses and complementisers, non-verbal predication, adverbial clauses, relative constructions, as well as other word categories such as prepositions and conjunctions. Since word order variation influenced by information structure is described in detail in chapter 3, I will not

discuss word order and topic/focus expressions in this section. The syntax of relative constructions is analysed in chapter 5, so in this section I only display different functions and uses of the relative/adverbial clauses.

2.1 Segmental phonology

This section gives an overview of the sound system in Teke-Kukuya. I first introduce the consonant and vowel inventories, their distribution on different positions of a phonological word, the phonetic variation in their realisation, as well as some phonotactic rules. Then I discuss the syllable structure. The tone and prosodic system of the language are described in the next section. Here and throughout the thesis high tones (H) are marked by an acute accent, while low tones (L) remain unmarked and are highlighted by a grave accent only where useful.

The description on the phonology of Kukuya relies much on the pioneering and careful work done by Christiane Paulian, which was published in 1975 as a detailed presentation on the phonology of the language, and later observations and analyses of the prosodic domain in Kukuya by Larry Hyman in 1987 also provide crucial reference for my description in this part. The work in Paulian (1975) was based on her intensive fieldwork in Brazzaville, Lagué, Ngoulonkila and Lékana on the Kukuya variant of Teke, and Hyman (1987)'s analysis was dependent on the same corpus and personal communication with Paulian.

Since my main aim of the thesis is to investigate the morphosyntax and information structure of the language, I was not ambitious on contributing a more fine-grained description on its phonology. Therefore, the minimal pairs established by Paulian to investigate the phonemes and the calculation of the frequency of different phonemes and syllable structures based on her large lexicon database are taken as a baseline in my description. However, in the fieldwork I also checked most of the words she collected and the transcription in her monograph, and I also added the description on vowel nasalisation and coalescence, and some discussion on the re-

maining problems on the postlexical prosodic domain in the verb phrase that were not discussed in Paulian (1975) and Hyman (1987). Some of the previous findings and generalisations are re-interpreted in the light of the fieldwork carried out by myself as well as of the enlarging investigation and knowledge on the Teke cluster conducted by the linguists in recent years.

2.1.1 Consonants

As we will see, the distribution of consonants in Kukuya is quite unbalanced in the initial and intervocalic positions of the stem. Here a stem can refer to a nominal stem or a verb stem. Kukuya as a Bantu language is characterised by different noun classes, which are comparable to grammatical genders and are usually encoded by a CV-shape nominal prefix (see section 2.3.1). A nominal is usually formed by a noun class prefix and a lexical stem, and a verb stem consists of a lexical root, derivational extensions and a final vowel. Since there are no productive verbal extensions in Kukuya, the verb stem is simply the lexical root plus the final vowel. In this subsection I first present the stem-initial consonants which cover the full consonant inventory, and then introduce the extremely reduced intervocalic distribution, a special phenomenon on the “constricted” articulation of the consonants is described last.

2.1.1.1 Stem-initial consonants

		Bilabial	Labial-dental	Alveolar	Palatal	Velar	Glottal
	Nasal	m	ɱ	n	ɲ	ŋ	
Plosive	Prenasalized	mp	mb	nt	nd	ŋk	ŋg
	Plain	p	b	t	d	k	
Affricate	Prenasalized		mpf	mbv	nts	ndz	
	Plain		pf	bv	ts	dz	
	Fricative		f	s	(ʃ)		(h)
	Approximant			l	j	w	

Table 2.1: The stem-initial phonemic consonants in Teke-Kukuya

As we can observe in Table 2.1, in Kukuya there is no voiceless/voiced con-

trast for non-prenasalised velar plosives, and this merger may have already happened in Proto-West-Coastal Bantu (Pacchiarotti and Bostoen 2020); no labial-dental and alveolar fricatives, i.e. no /g/, /v/ and /z/ as an initial consonant, the absence of which is also attested in other Teke varieties (Kristensen et al. 1986; Abandzounou 2012, Raharimanantsoa 2012, 2020); while all the prenasalised consonants have this contrast. All the voiceless consonants are slightly aspirated, but this can vary depending on the region and speaker. /h/ is attested only in one word which means ‘also’ [hé] so it is hard to justify identifying it as a phoneme. The same word is pronounced as [kɛ] in Teke-Eboo (B74), so perhaps the [h] is just a variant of /k/. In the transcription and orthography, the affricates /mpf/, /mbv/, /nts/ and /ndz/ are transcribed as <mf>, <mv>, <ns> and <nz> for simplicity. The palatal fricative /ʃ/ (transcribed as <s> in orthography) which was not placed in the consonant inventory in Paulian’s description is rarely attested. See comparison of /ʃ/ with /s/ in (1). The boundary between the prefix and the nominal stem is marked by a hyphen.

- (1) se /ʃe/ ‘also, regardless of’
 ki-sébe /ki-sébe/ ‘to laugh’
 séneme /séneme/ ‘six’

In Paulian (1975) some notes were made on the conditions of some of these initial consonants: /w/ is realised as labio-velar when followed by back vowels and /a/, and as a labio-palatal glide [ɥ] when followed by a front vowel, for example in the words ‘to give’ [ki-wâ] and ‘to lack’ [ki-ɥ^wéènà]. /m/, /n/ and /l/ can be geminated when occurring at the stem-initial accentuated position (see more detail in 2.2.2). /n/ is never attested before /u/. /f/, /mpf/ and /mbv/ never occurred before /o/ and are only rarely seen before /e/. /pf/ is never attested before the low vowels /e/ and /o/, which may be due to the fact that it originated from a voiceless plosive, and a process of assibilation (also known as Bantu spirantisation) occurs before high vowels for it to become a fricative and then an affricate. However, its voiced counterpart /bv/ is attested before /e/. /d/ is very rarely observed (absent in some other Teke varieties) and is never attested before /o/. Many stems in Proto-Bantu (PB) with this voiced dental plosive as initial consonant have been realized as the lateral /l/ in Teke, which may account for the rarity of /d/. /s/, /ts/

and /dz/ are realised as apico-prepalatal sibilants [s], [ts] and [dz] (as well as when prenasalised) before most vowels but are palatalised as [ʃ], [tʃ] and [dʒ] respectively before <wV> (see example (11b) below), /dz/ is also realised as [dʒ] before /e/ as in *dze* [dʒe] ‘step’.

/ɱ/ as a phonemic labio-dental nasal, is reported as *only* occurring in the Kukuya language but very rarely as a phoneme in other known languages (but only phonetically) in the world (Ladefoged and Maddieson 1996; also see Hajek 2006, 2009 on this phoneme in Tibeto-Burman). According to Paulian (1975), this plosive phoneme is realised by the release of airflow between the top teeth and the lower lip and is accompanied by “a strong protrusion of both lips”, which was exactly what I observed on the manner of articulation of the speakers when they tried to contrast this phoneme with /m/. However, this contrast seems to be only observed with elder speakers and is being neutralised among younger generations. This phoneme is never attested before /u/ and /o/ and is realised as [ɱ^w] before /a/, thus is transcribed as <mw>. As we will see shortly in the discussion of glides, the vowel in the sequence CwV has the same restrictions, so it seems that the /ɱ/ can be analysed as a variant of /m/ before the sequence wV. Here I follow Paulian (1975) to analyse the /ɱ/ as an independent phoneme, as all other bilabial consonants cannot be followed by the wV sequence, and the fact that it cannot be followed by /o/ reflects the same restriction for all other labio-dentals (/pf/, /bv/, /f/). The [ɱ^w] realisation before /a/ could be due to the labialisation triggered by the strong protrusion of the lips conflicting with the spread front vowels /i/ and /e/ and can only be fully realised before /a/. Some minimal pairs from Paulian (1975) distinguishing /m/ and /ɱ/ are illustrated in (2).

- (2) mwáana /ɱjáana/ ‘child’
 mwíi /ɱjii/ ‘eyes’
 míi /mii/ ‘urine’
 ki-mwaala /ki-ɱjaala/ ‘to laugh at’
 ki-maala /ki-maala/ ‘to complete the rest of...’ (Paulian 1975: 36)

Another question arising from Table 2.1 is whether we should treat the *prenasalised consonants* as phonemes or consonant clusters. In some stud-

ies on other Teke variants, the prenasalised consonants are analysed as sequences of nasal plus consonants (Raharimanantsoa 2012, 2020). According to Paulian (1975) the prenasalised consonants in Kukuya should be considered as mono-phonemes, for the following reasons: the nasal part of the complex consonant always assimilates with the oral part in the place of articulation; the prenasalised consonants have a clear contrast between voiceless and voiced ones, while some voiceless oral consonants do not have their voiced counterparts, e.g. there exists the prenasalised consonant /ŋg/ that contrasts with /ŋk/ while /g/ is absent, so /ŋg/ should necessarily be a mono-phoneme, otherwise we have to postulate that /g/ only occurs when prenasalised; in addition, according to Paulian's calculation on the frequency of the consonants, she showed that the frequency of /nd/ even surpasses that of /d/, if we treat the prenasalised consonants as consonant clusters, this difference in frequency is not expected. Here I object her last account since the rarity of /d/ may be due to the diachronic shift from /d/ in PB to /l/ and the larger occurrence of /nd/ sequence may also result from neutralisation of the historical /n-l/ cluster. However, I follow Paulian to analyse (at least some) NC sequences as phonemes, while some of them are unambiguously consonant clusters. Here at least three types of NC sequences should be distinguished:

i) NC sequences that are mono-phonemic consonants, which always occur at the stem-initial position: see examples in (3). Here I interpret the prenasalised consonants as originated from an NV-prefix in which the vowel may have been lost and the nasal becomes attached to the stem and inseparable, new class prefixes can be added as the nasal is reanalysed as part of the stem. This can also account for the fact that a verb root never begins with a prenasalised consonant in Kukuya.

- (3) /baa-ntaba/ 'goats'
 /kii-mbúli/ 'lion'
 /nziimi/ 'many, much'

ii) separable nasal prefix followed by a stem-initial oral consonant: these cases are attested with class 9/10 nominal and concord prefixes as shown in (4), and more frequently, with the 1SG subject marker N- in (5). In these ex-

amples the duration of the NC cluster in pronunciation is noticeably longer than in (i) above. When the NC sequence occurs, a neutralisation rule on the contrast between stop and continuant is always applied as in (6), which Hyman (2019: 139) summarises as “Bantu languages prefer that post-nasal consonants be [+voice] rather than [-voice] and [-continuant] rather than [+continuant]”. This rule can also be interpreted as a prohibition of the change of C after a nasal, for example the weakening of *p to [w] and *d to [l] as in the word *ki-lóbo* (*dób-) ‘to catch fish’ is blocked after a nasal as in *m-lóbo* [mdóbo] ‘I catch fish’. Post-nasal affrication is also attested, also shown in (6). When a nasal prefix is followed by a stem that also starts with a nasal, the sequence is realised as a geminated nasal, but as we will see shortly, stem-initial nasals are themselves geminated, so the nasal prefix may be just unpronounced.

- (4) li-yǔ [li-yǔ] (cl.5) ‘a peanut’
 n-yǔ [n-dzǔ] / (cl.10) ‘peanuts’
 mpólo n-kima [mpólo ŋ-kima] (cl.9) ‘another chance’

- (5) a. Me m-wééná mí-para.
 [mempéénámípata]
 ‘I lack money.’
 b. Me n-yuk-í bó-ri we ya má-lúa.
 [mendzukíbóri we ya málúa]
 ‘I heard that you were sick.’

- (6) *Neutralisation rule of NC cluster*
 N+/p/, N+/w/ > [mp]
 N+/pf/, N+/f/ > [mpf]
 N+/ts/, N+/s/ > [nts]
 N+/dz/, N+/j/ > [ndz]
 N+/l/ > [nd]

iii) there are also cases in which the underlying structure of NC sequence is ambiguous, especially for the class 9 nouns that have their plural counterparts in class 6. Class 9/10 nominals start exclusively with a nasal, however,

as shown in (7), the class 6 plural prefix *ma-* does not substitute the nasal prefix of the singular form, in contrast with the class 10 nasal prefix that replaces the singular class 5 prefix in (4) above. Also shown in (7), in the formation of diminutives, the first syllable of a class 9 noun stem combined with the nasal prefix is reduplicated (see more on the diminutives in 2.3.1) as a whole unit, while for other noun classes the prefix is never reduplicated, which again shows that the nasal is inseparable from the stem. This may also suggest that the original nasal prefix has undergone homorganic assimilation with the initial oral consonant and becomes available for noun class change by adding other class prefixes. What slightly differs from (i) is that the nasal prefix here can still contribute to the characterisation of noun class and is associated with the concord morphology (see section 2.3.1 below).

- (7) /ntséke/ (cl.9) '(small) field'
 /ma-ntséke/ (cl.6) '(group of) fields'
 /nzó/ (cl.9) 'house'
 /ki-nzó-nzo/ 'small house' (diminutive)
 /mu-bilí/ (cl.3) 'cola tree'
 /ki-bĩ-bilí/ 'small cola tree' (diminutive)
 /nzó yi-m-bvé/ (cl.9) 'a nice house'

2.1.1.2 Glides

The phonological status of glides is not ambiguous in Kukuya. There are clearly cases in which an initial glide is itself consonantal and is not preceded by any other consonants, as shown in (8a). This can be proven by the facts that in these cases the /w/ and /y/ are not a tone-bearing unit, and they can be followed by vowels of all qualities, which shows that they function as the initial consonants (see in Table 2.3 below that vowel successions are quite restricted), though they can be originally derived from vowels.

- (8) a. ki-woo 'to help'
 ki-yaala 'to peel'

- b. ntswíi ‘fish (sg.)’
 ki-byáala ‘to command’

For the sequences CwV and CyV as in (8b), the glides should be considered as derived from underlying vowels (i>y; u>w) but no longer function as vowels, as they cannot bear tones and there are oppositions between CwV/CyV and CuV/CiV which also differ in duration, see the examples in (9a). The sequences Cwe and CwVCV are allowed, but we will see in section 2.1.2 that the vowel succession u-e is not allowed and only identical V₁V₂ succession occurs in CVVCV, which set the glides apart from the true vowels. The glides in these sequences can neither be equated to the consonants /w/ and /j/, as the latter can be followed by any vowels, while in the CwV sequence, the V is limited to /i/, /e/ and /a/, and the C cannot be anterior consonants (/n/, /t/, /d/); for the CyV sequence, the choice of C is more free but the V can only be /a/ and /e/ (Paulian 1975: 59). Due to the limited distribution of intervocalic consonants and constraints on the syllable structure (see momentarily in this section), the sequence VGV is not attested.

The CwV can in some cases labialise and palatalise the initial consonant, /s/, /ts/ and /dz/ are realised as [ʃ^w], [tʃ^w] and [dʒ^w] respectively when followed by the glide /w/, as illustrated in (9b). Moreover, as we will see shortly in section 2.1.2, in CwV the glide can trigger labialisation on the following vowel /i/, which is realised as [y]. The stem-initial sequences Cw and Cy are often considered to be derived from a merge of a Proto-Bantu nominal prefix and a root, in which case the vowel is usually lengthened (see section 2.3.1).

- (9) a. ki-kwâ ‘to die’
 ki-kúa ‘to bring up a child’
 li-pya ‘gamble’
 pia ‘fetish’
- b. ki-sweeke [ki-ʃwεεgε] ‘to hide’
 ki-swaaka [ki-ʃwaaga] ‘to wash’
 tswaala(-tswaala) [tʃwaala] ‘quickly, fast’
 dzwiá [dʒwá] ‘hole’

2.1.1.3 Constricted articulation ‘ChVV’

Next I introduce a somehow special articulation of the initial CVV-shape syllable, which is commonly attested in Kukuya words. I follow Paulian’s orthography to put an **h** between the initial C and the VV sequence to symbolise this “marked” articulation. According to the calculation in Paulian (1975), **ChVV** has almost the same frequency of occurrence as the non-marked CVV syllable. In the fieldwork I observed that while the elder speakers can easily distinguish **ChVV** from CVV, this contrast is being lost among younger speakers who also know less vocabularies of the language. A similar phenomenon is reported in Kristensen et al. (1986) in their phonological sketch of Teke-Eboo (B74), in which they refer to this modified articulation as “glottalised consonants” that are realised with strong aspiration but they did not provide solid generalisations.

This modified articulation can be described as the narrowing of the passage of airflow between the initial consonant and the VV sequence, and this stricture triggers a fricative-like offglide (Hyman 1987: 326), not only influencing just one segment but affecting both the realisation of the consonant and the vowels. This “constricted” /Ch/ only occurs when a stem-initial consonant is followed by a VV sequence and can contrast with the non-modified initial consonant in the same syllable structure CVV, see the minimal pairs in (10a). The consonant in this ChVV can be any of the initial consonants except the alveolar stops, and the first vowel in the VV sequence cannot be /a/.

- (10) a. ki-kée [ki-kée] ‘to be dry’
 ki-khée [ki-k^híe] ‘to try’
 ki-kúa [ki-kúa] ‘to raise (a child)’
 ki-khúa [ki-k^húá] ‘to weed the grass’
- b. mu-pfhíili [mu-p^híili] ‘widower’
 ki-bhiima [ki-b^hi:ma] ‘corpse’
 bhooko [b^wo:ko] ‘fear’
 nghuunú [ŋg^wu:nu] ‘forest field’ (Paulian 1975: 76)

This marked articulation can influence the realisation of the consonant,

which also depends on the quality of the following vowels. From the examples in (10) we see that when the passage of the airflow is narrowed, a friction occurs between C and VV and C is realised as [Cj] or [Cw] when followed by a front vowel or back vowel respectively. The constriction can also trigger affrication of the plosive consonants, as summarised in the Table 2.2. [p^f] and [b^v] never occur before back vowels, possibly because of the existence of the phonemes /pf/ and /bv/. More examples are given in (11a).

Ch	before front vowel	before back vowel	Example
/k/ (/nk/ and /ng/)	[k ^{sj}]	[k ^{fw}]	ki-khúula [ki-k ^{fw} úula] 'to depend'
/p/, /b/ (/mb/)	[p ^{sj}], [b ^{zj}]	–	bhiima [ki-b ^{zj} i:ma] 'corpse'
/j/	[z ^j]	[ʒ ^w]	yhoomi [ʒwo:mi] 'grasshopper'
/s/, /ts/, /dz/	[ʃ], [tʃ], [dʒ]		ki-tshuuma [ki-tʃu:mu] 'tree rat'

Table 2.2: The realisation of Ch before different vowel qualities

The modified articulation can also cause a tension on the following vowels. According to Paulian's transcription, while the non-modified VV sequences are realised as [i:], [ɛ:], [u:] and [ɔ:] (also see section 2.2.2), <hii> and <huu> are realised as [yⁱi:] and [w^uu:] when the VV sequence bears the same tone and [yⁱii] and [w^uuu] with a contour tone; <hee> and <hoo> are realised as [yⁱie] and [w^uuo] in ChVV, while as [ye:] and [wo:] in ChVVCV when the vowels bear the same tone, see some examples in (11b).

- (11) a. ki-tshuumu [ki-tʃu:mu] 'tree rat'
 ki-yhiila [ki-z^yi:la] 'to last'
 yhoomi [ʒ^wo:mi] 'grasshopper'
 ngheemé [ŋg^{zy}e:mé] 'dogfish'
- b. ki-khúu [ki-k^{fw}úu] 'axe'
 ki-khúina [ki-k^{sy}úina] 'get blocked'
 ki-lheeme [ki-lye:mɛ] 'become clear'
 ki-khóolo [ki-k^wúolɔ] 'crow (cock)' (Paulian 1975: 78)

The <h> in this environment does not seem to be a segmental phoneme, as its presence can influence both the consonant and the vowels in different ways, but its own features are difficult to schematise, as it is not realised as simple aspiration or palatalisation. It seems more appropriate to describe

this <h> as a modified articulation that influences the initial CVV-unit as a whole. Diachronically, Paulian noticed that most of the stems ChVV(CV) in Kukuya with this modified articulation originated from the Proto-Bantu stems that had intervocalic NC consonants, and the occurrence of <h> thus is due to the loss of a stem-medial consonant and the compensatory lengthening of the V₁; when the stem vowel is /a/, <h> does not occur, see some comparisons in (12).

(12)	<i>Proto-Bantu</i>	<i>Teke-Kukuya</i>
	*-bimba ‘corpse’	/ki-bhiima/
	*-cimb- ‘to catch’	/ki-shiima/
	*-dond- ‘to follow’	/ki-lhoono/
	*-banjí ‘arrow’	/li-bái/

Paulian also proposed an alternative interpretation of this constricted articulation that it could be triggered by two very high vowels, which cause the friction between the initial consonant and these vowels.

2.1.1.4 Intervocalic consonants

In the intervocalic positions in a Kukuya word, consonants have a very limited distribution compared to the initial position. Only six consonant phonemes can occur stem-internally, as summarised in Table 2.3 below. While all the six phonemes can occur at C₂ position, only /p/, /m/ and /k/ can occur at C₃. The loss of consonants after C₁ is commonly seen in many Zone A and B languages.

Consonant lenition is often attested in the intervocalic positions. /p/ and /k/ are never aspirated intervocalically. /p/ is often realised intervocalically as voiced [b] or approximant [β], while /k/ is realised as [g] or [ɣ], which vary among speakers and depend on the speed of utterance. In rapid speech, intervocalic /m/ and /n/ can often be dropped and nasalise the following vowel (see vowel nasalisation in section 2.1.2). A stem-internal /t/ is always pronounced as an apical tap [ɾ] as in the word *taara* /taata/ ‘father’, which is never attested in the stem-initial position. Given the fact that intervocalic

	Bilabial	Dental	Velar
Oral	p	t	k
Nasal	m	n	
Lateral		l	

Table 2.3: Intervocalic consonants in Teke-Kukuya (Paulian 1975: 88)

positions have a too restricted consonant system to include a phoneme that is distinctive from all the consonant inventories in the initial position, I treat [r] as an variant of /t/ while in orthography I always transcribe it as <r>. According to Paulian’s calculation, /l/ is the most frequently occurring intervocalic consonant.

C ₁	C ₂
/n/	/l/
/ŋ/	/k/
/d/	/t/
/mv/,/m/	/m/
/bv/, /m/,/ŋ/,/ɲ/	/p/

Table 2.4: Prohibited C₁-C₂ combinations in Kukuya

There are some restrictions on the combination of consonants in different positions. For the CV(V)CV shape, the possibilities of C₁-C₂ combinations are relatively free, so I just list in Table 2.4 those which are prohibited. The combinations of C₂-C₃ in CVCVCV are much more limited so I list all the possibilities according to Paulian (1975)’s calculation in Table 2.5. We can notice that in CVCVCV shape the choice of C₂ is more restricted than in CVCV, anterior consonants are not allowed in C₂ while central consonants are not allowed in C₃, only /k/ can occur in both positions. Nasal and oral consonants exclude each other in C₂-C₃ combinations, and C₂ and C₃ cannot be identical.

As for the combinations of C₁-C₂-C₃ in CVCVCV, only the initial /w/, /t/, /l/, /s/, /k/ allow all the six possible C₂-C₃ combinations above. See some

C ₂	C ₃
/t/	/p/,/k/
/l/	/p/,/k/
/k/	/p/
/n/	/m/

Table 2.5: C₂-C₃ combinations in Kukuya

illustrative examples in (13).

- (13) ki-téleke /ki-téleke/ 'to prepare, to cook'
séneme /séneme/ 'six'
ki-kokobo /ki-kokopo/ 'to calm down'
ki-téreke /ki-téteke/ 'to tremble'

To summarise, in this subsection we see that while the full inventory of consonants are allowed in the initial position, the intervocalic positions have only a very reduced consonant system and even the same phoneme has different realisations when occurring initially and intervocalically. The reduced consonant system in the intermediate positions may be due to a neutralised realisation such as voiceless and voicing of the initial consonants. Next I introduce the vowel system of the language.

2.1.2 Vowels

Kukuya employs a five-vowel system as summarised in Table 2.6. The phonemes /i/, /e/, /o/ and /u/ can have variants with regard to different degrees of openness. /i/ is often realised as a high front vowel but as [ɪ] for some speakers, /e/ is realised as an intermediate vowel between the [e] and [ɛ], /o/ can vary between [o] and [ɔ], /u/ can vary from [u] to [ʊ], which depend on different speakers from various regions and may be influenced by the vowel qualities of the vowels in other Teke varieties that they have contact with.

Orthography	Phoneme	Description	Example
i	[i]	high front unrounded	ki-síka [ki-síka] ‘to make excited’
e	[e]~[ɛ]	(low-)mid front unrounded	mu-yeni [mu-jeni] ‘visitor’
a	[a]	low front unrounded	mbaá [-mbaá] ‘fire, heat’
o	[o]~[ɔ]	(low-)mid back unrounded	ki-tô [ki-tô] ‘to arrive’
u	[u]	high back rounded	bu-bui [bu-bui] ‘youth’

Table 2.6: Phonemic vowels in Teke-Kukuya

There is no close front rounded vowel phoneme /y/ but we can often encounter it as an allophone, as in the language name ‘Ki-kukuya’ /ki-kwíkwíá/ [ki-kykya]. This allophone [y] occurs when the stem-initial syllable starts with Cw and the following vowel successions are /i/, /ii/, /ee/ or /ia/, and is triggered by the glide /w/ that labialises the following front vowel, as in the words *tswíi* /tswíi/ [-tʃyi] ‘ear’, *ki-dzwídzwáali* /ki-dzwídzwáali/ [ki-dʒydzwáali] ‘(small) injury’ and *dzwíá* /dzwíá/ [dʒyá] ‘hole’.

2.1.2.1 Vowel combination and coalescence

All the five vowel phonemes can occur in different positions on the stem but there are restrictions on their combinations. For the V₁-V₂ combination in CVCV shape, V₁ and V₂ can be any of the five vowels when the two vowels are identical; /i/ as V₂ can combine with any V₁; and only high vowels can combine with /a/ as V₂. These possible combinations are summarised by Paulian in Table 2.7 below, and I don’t find exceptional cases in my data. The V₁ and V₂ combination patterns in CVCV are exactly the V₁-V₃ combinations that are allowed in CVCVCV, the V₁-V₂ successions allowed in CVV, and the V₁-V₃ combinations in CVVCV. Based on these facts, I distinguish only the initial vowel (V_i) and final vowel (V_f) in the table.

The intermediate vowel (V_m) in CVCVCV behaves like a “neutral” vowel and does not affect the occurrence of V_i-V_f, while the realisation of V_m, on the contrary, depends on the V_i-V_f combination. V_m is the same as V_i and V_f when these are identical; V_m is realised as the same as V_i when V_i is a high vowel and V_f is /a/; V_m is realised as V_f when V_i is not a high vowel and V_f is /i/; when V_i and V_f are both high vowels, V_m can be realised as either V_i

$V_i \backslash V_f$	i	e	a	o	u
i	+		+		
e	+	+			
a	+		+		
o	+			+	
u	+		+		+

Table 2.7: The combination patterns of V_i - V_f
(adapted from Paulian 1975: 104)

or V_f . Some examples of the realisation of V_m in CVCVCV are given in (14). In spite of these rules, V_m is often realised simply as a schwa [ə] in natural speech.

- (14) ki-feréke /ki-fetéke/ 'small hole'
 mu-nkonomo /mu-nkonomo/ 'punch'
 ki-binima /ki-binima/ 'go to search'
 mu-sáliko /mu-sáliko/ 'shout, scream'
 ki-dzulibi /ki-dzulipi/ 'door'

In $CV_1V_2CV_3$, the V_2 is always realised as identical to V_1 , and successions of different vowels are not allowed. I suppose that there is no phonological contrast between short and long vowels. The V_2 in CVVCV should be treated as independent of V_1 but is not part of a long vowel for the following reasons. First, as we will see shortly on the tonal system (see 2.2.1) that in CVVCV with LHL tone pattern, the H tone is distributed to V_2 , so the tones on V_1 and V_2 are not necessarily the same; the tone on V_1 is always fixed while V_2 is subject to tone change rules such as H tone plateauing (see the rule application in (24) below), which also indicates that V_1 - V_2 is not a long vowel but vowel succession, as seen in (15b). In addition, in the derivation of diminutive nominals (see section 2.3.1) in which the first syllable is reduplicated, only CV_1 can be doubled even if the stem has an initial CVV syllable, as shown in (15c).

- (15) a. ki-tene [ki-tɛnɛ] ‘to knead’
 ki-teene [ki-tɛ:nɛ] ‘to respond’
 ki-pura [ki-pura] ‘to plow’
 ki-puura [ki-pu:ra] ‘to hide behind’ (Paulian 1975: 111)
- b. ki-baáama [ki-baáama] ‘to wake up (suddenly)’
 li-báani ‘the beginning’
 li-báání líí ‘the beginning of’
- c. mbaa ‘fire’
 ki-mbu-mbaa ‘small fire’
 nkwiíri ‘medicine (by witchcraft)’
 nkwi-nkwiíri ‘ineffective medicine’

Based on these facts, I treat the identical VV sequence in CVVCV as two vowels rather than one long vowel. It is noteworthy that in CVV the vowels are not necessarily identical as in CVVCV and the V_1 - V_2 combination actually corresponds to V_i - V_f which is subject to the occurrence restrictions in Table 2.7.

Vowel coalescence is a phonological phenomenon in which two adjacent vowels cause each other to change (Harford 1997). This can occur across the boundary of nominal or verbal prefixes within and across words, and is commonly attested with a prefix and a vowel-initial stem. Some examples are illustrated in (16) and (17). In (16a) the vowel in the word *nzó* ‘house’ corresponds to the V_1 on a stem and is maintained when followed by the class 9 connective marker *aa*, which also results from the stem-initial prominence (see section 2.2.2); in (16b) the final vowel of the word *ngúku* ‘mother’ corresponds to the V_2 which is non-accentuated, thus it disappears when combined with the connective marker.

- (16) *nzó aa me* [nzóɔ-mɛ] ‘my house’
ngúku aa we [ngúga:-ʉɛ] ‘your mother’

In rapid speech and also depending on individual speakers, the contact of adjacent vowels can change the quality of both of them, some general rules would be /u/+a/>/o/ and /i/+a/>/e/. In (17a) we see that the bilabial stop

/b/ is dropped when it is not stem-initial, and the contact between adjacent /a/ and /u/ across morpheme boundaries triggers the realisation of /o/; in (17b) the class 7 relative marker *ki-* is merged with the remote past tense marker *â-* and is realised as *keê-* (see the formation of relative clauses in section 2.5).

- (17) a. *yě ŋa bu-lá búú baarí* [yě-ŋo-lló:ba:rĩ]
 go 16.LOC 14-village 14.CONN 2.people
 ‘go to the village of the people’ (Paulian 1975: 146)
- b. *ki-ko keê-li â-n-ték-í me*
 7-clothes 7REL.RPST-COP RPST-1SG.SM-sell-PST 1SG.PRO
 ‘the clothes that I had sold’

2.1.2.2 Nasalised vowels

Nasalisation is not a contrastive feature of vowels in Kukuya but only appears in certain phonetic environments. The nasalised vowels have been reported as occurring frequently in the Teke group since Guthrie’s fieldwork in the 1950s (Guthrie 1960, 1971), and some preliminary comparative works and description on the nasalised vowels in different Teke varieties were conducted in Hombert (1986, 1987) and Paulian (1994) and later in several manuals by SIL Congo (2013, 2016, 2020). Compared to other Teke varieties such as Ngungwel (B72a), Ndzindziu (B74a) and Ibali (B75), Kukuya (B77a), Fuumu (B77b), Iyaa (B73) and Tege (B71) seem to be located in an earlier stage of vowel nasalisation, which may reflect their relatively isolated linguistic situation. In Fuumu, the vowel nasalisation has not even started (at the moment of Hombert’s studies). According to my observation, in Kukuya the nasalised vowels as the realisation of nasal stops are just optional but is commonly attested in casual speech, which remains almost the same situation as forty years ago in the first documentation on this language.

The presence of nasalised vowels is one characteristic of the Teke cluster, and some universal tendencies of the vowel nasalisation process are observed in different varieties (Hombert 1986, 1987; Raharimanantsoa 2012). The nasalisation happens first with the stems that have $*-V_1mbV_2$ or $*-$

Underlying	Nasalised	Translation
ki-yíma	[ki-yó̃]̃	'to sing'
ki-yéeme	[ki-yé̃]̃	'to sleep'
ki-téneme	[ki-té̃]̃	'to stand'
ki-tshoomó	[ki-tsõ]̃	'to think'
bi-táámí	[bi-tá̃]̃	'fun (pl.)'
mpfúúmú	[-mpfú̃]̃	'chief, king'

Table 2.8: Examples of vowel nasalisation in Kukuya

V_1mV_2 reflexes in Proto-Bantu. More concretely, the nasalisation occurs first when the V_1 preceding a labial nasal is low, and at a later stage the nasalisation process happens with high vowels and with dental nasals. Velar nasals in Proto-Bantu disappeared without any traces of nasalisation on its adjacent vowels in Teke (Pacchiarotti et al. forthcoming). For the reflexes $*-V_1mbV_2$ and $*-V_1mV_2$, the nasalisation process happens earlier in the former, as the prenasalised stops often become plain nasals with compensatory lengthening of the preceding vowel and long vowels are affected before short vowels in nasalisation. The quality of the nasalised vowel is more often conditioned solely by the V_1 independent of the V_2 . These generalisations are also borne out in Kukuya, as illustrated in the examples in Table 2.8.

In Kukuya, the vowel nasalisation is attested with all qualities of V_1 and happens with intervocalic nasal consonants /m/ and /n/, but never triggered by stem-initial nasal consonants. The intervocalic labial nasal can influence all the adjacent vowels but nasalise only the V_2 and V_3 . As we see in the table above, the final vowels are often rounded as [ɔ̃] when being nasalised and the labial nasal can also the labialise V_1 as in the word *ki-yíma* 'to sing'. Vowel nasalisation can also be triggered by prefixes that starts with a labial nasal, as show in the class 1 adjectival concord prefix *mu-* in (18a). There is another optional casual speech rule by which a bilabial [b] may be deleted if it is not stem-initial, as shown in (18b). In this case the final vowel is also rounded by the dropped labial stop, it remains to be studied whether and how this is related to the vowel nasalisation process.

- (18) a. mwáana wu-mu-kái ‘a girl, *lit.* a female child’ [mʷáanawɔ́kái]
 b. ki-yáaba [ki-jáao] ‘to know’
 ki-kaba [ki-kao] ‘to share’

2.1.3 Syllable structure

All the possible syllable shapes in Kukuya are listed in Table 2.9 with examples. In the table, C stands for an oral consonant and is distinguished from a nasal consonant N. A morpheme consisting of one nasal is only attested for 1SG subject marker and class 9/10 concord prefix which must be attached to the stem. The syllable shape of V(V) is only allowed for functional categories such as possessive markers. It should be noted that the prenasalised consonants and glides only occur in the stem *initial* syllable, so if we unify the oral C, NC and CG as one unique onset C, Kukuya has in fact only two syllable types, namely CVV and CV. I include the CVV (NCVV, CGVV) shape here as a syllable type, but it only occurs in the stem structure (N)C(G)VVCV in which V₂ is always identical to V₁, and when the CVV occurs as a stem by itself, it should be divided into two syllables, i.e. CV-V.

Syllable	Examples
VV	áá ‘of’ (cl.3/10)
CV	ko ‘banana’
CVV	ki-yéeme ‘to sleep’
NCV	ndé ‘she/he/it’
NCVV	mbaa ‘fire’ nkaaka ‘grandmother’
CGV(V)	dzwiá ‘hole’ myáaka ‘hands, arms’
(N)CGV(V)	nkwí ‘firewood’ nkweéme ‘praying mantis’

Table 2.9: The syllable shapes in Kukuya

A nominal or verbal stem in Kukuya can have maximally three moras, and

if we consider the oral and prenasalised consonants and the glides as the stem-initial onset, Kukuya only distinguishes five types of stem structures, namely CV, CVV, CVCV, CVVCV and CVCVCV, as exemplified in Table 2.10 below. This phonological constraint on the syllable structure may be associated with the loss of productive verb derivational extensions in this language.

Most loan words are adapted to the syllable structures allowed in Kukuya, while there are many exceptions, see some examples in (19). We see cases in which an epenthetic vowel does not need to occur to create a CV syllable, as in the word *doktóori* ‘doctor’, and a non-initial syllable can surface in the CVV shape, as in the word *mu-políisi* ‘policeman’.

- (19) doktóori ‘docteur’ (Fr.)
 tabúlu ‘table’ (Fr.)
 mu-políisi ‘police’ (Fr.)
 li-kóoli ‘école’ (Fr.)
 búku ‘book’ (Eng.)
 lopitáali ‘hospital’ (Fr.)
 mangúlu ‘mango’ (Fr.)

Stem structure	Examples
CV	ki-wá ‘to give’
CVV	li-meé ‘stone’ dzwiá ‘hole’
CVCV	bu-ndúku ‘friendship’
CVVCV	ki-fúula ‘to ask’ myáaka ‘hands, arms’
CVCVCV	ki-bólolo ‘to be broken’ ki-kwereke ‘to punch’

Table 2.10: The stem types in Kukuya

In this section I have provided an overview of the segmental phonology of Kukuya, including the consonant and vowel inventories, their distribution on a stem and phonetic modifications in various conditions. We see that

while the stem-initial position can have full inventory of consonants and vowels, the intervocalic positions have a very reduced consonant system. So far we have an impression that in a Kukuya word stem, various types of fixed combinations are exploited. For example, exactly six combinations are allowed for the C_2 - C_3 in CVCVCV (see Table 2.5), eleven possible combinations of V_i - V_f (see Table 2.7) are attested, and five types of stem syllable structures are observed. We will see in the next section that this preference on fixed combinations also exists in the tonal system of this language.

2.2 Prosody

Kukuya is a tonal language and the tone patterns can encode lexical and grammatical distinctions. What is of particular interest here is the fact that in Kukuya there are five **fixed tonal melodies** which are L, H, LH, HL and LHL. These tone patterns are assigned to certain prosodic domains, and we will see evidence that there is both a lexical and a postlexical prosodic domain in this language. This postlexical domain is somehow unusual in that a prefix-stem sequence, which is the most common shape of a grammatical word, does not stand for a phonological word. Instead, all the prefixes join the **preceding stem** to form a prosodic domain. In addition, a stem-initial prominence interacts with the tonal system as well as the phonetic realisation of consonants and vowels. These prosodic properties of Kukuya were first described in Paulian (1975) and were further discussed in Hyman (1987). In this section, I first introduce the tonal system and then I discuss the stem-initial prominence.

2.2.1 Tone

Kukuya distinguishes two contrastive pitch levels that are transcribed as high tone (H) and low tone (L). The tonal system in Kukuya is intriguing in that a certain phonological domain is assigned one of five fixed tone patterns independently of the number of the tone-bearing units (TBU) that it contains. In other words, the H tone and L tone in this language have combined into different schemes which are assigned to a certain phonological unit in regular ways without exceptions, which is comparable to the four tone schemes in Mandarin Chinese that are distributed on the syllable level. As we will see shortly, the realisation of the tone patterns also depends on whether they occur before a pause or are followed by other elements, and the phonological domain that a tone pattern is assigned to is also fixed but can be extended beyond the syllable and stem level to cross over the word boundaries. In this subsection, I first present the distribution of the tone patterns on lexical stems, then I discuss the prosodic domain beyond the lexical word level. I will also introduce different types of grammatical tones

in this language.

2.2.1.1 Lexical tone pattern realisation

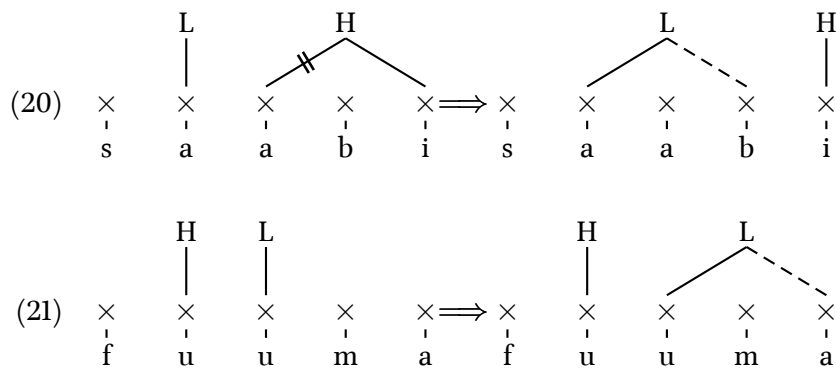
The realisation of the five tone patterns on different syllable structures of lexical stems are presented in Table 2.11, when the stems are followed by other stems without any prosodic break. We notice that in all these five tone patterns a H tone can occur only once, which may be accounted for by an economy principle as H tone is often considered as marked while L tone is default in many Bantu languages (Marlo and Odden 2019).

Tone pattern	Realisation on different stems	Examples
H	ĆV, ĆV́, ĆVĆV, ĆV́VĆV, ĆVĆV́ĆV	<i>mi-táámí</i> 'fun'
L	CV, CVV, CVCV, CVVCV, CVCVCV	<i>ki-lili</i> 'tear'
HL	ĆV̂, ĆV́, ĆVĆV, ĆV́VĆV, ĆVĆV́ĆV	<i>ki-fúuma</i> 'to buy'
LH	ĈV, CV́, CVĆV, CVVĆV, CVCVĆV	<i>ki-saabí</i> 'roof'
LHL	ĈV, CV̂, CVĈV, CV̂VCV, CVĆV́CVCV	<i>ki-palâ</i> 'to go out'

Table 2.11: The tone patterns in Kukuya and their realisation on lexical stems (when followed by another stem)

We start by looking from the table how these tone patterns are mapped onto the lexical stems of different numbers of syllables. Here I refer to a stem as the verb or noun root plus the final vowel excluding any prefixes. Hyman (1987) observed that the mapping fashion of the five tone patterns to the stem is left-to-right and one-to-one which follows exactly the association conventions proposed in Goldsmith's (1976) autosegmental phonology. When the number of tones and moras are the same, each tone is just distributed on the respective mora, e.g. the LH pattern on a CVĆV stem. When there are more moras than tones, then copy the L tone (if present) to the next syllable until the whole tone pattern is distributed to all moras, e.g. the HL pattern on a ĆVCVCV stem and the LH pattern on a CVCVĆV stem; when there are more tones than moras, then assign all the remaining tones to the rightmost mora, e.g. the LHL pattern on a CVĈV stem.

To illustrate, for example the LH tone pattern on the noun *ki-saabí* ‘the roof’. When LH is assigned to a stem with three moras, it cannot be realised as LHH but instead LLH is attested, whereby a L tone is copied to the next syllable and the remaining H tone is assigned to the last syllable, as shown in (20). For the HL tone pattern on the verb stem *ki-fúuma* ‘to buy’, the same L tone copying rule is applied as in (21).



When a stem occurs before a prosodic break, i.e. not immediately followed by any other element, a final H tone is always lowered to a mid tone, as illustrated in Table 2.12. A L tone remains unchanged when occurring utterance-finally and when the number of moras are not fewer than tones. The avoidance of H tone on a vowel that is adjacent to a pause is also a general tendency in Bantu languages (Hyman 2001; Downing 2019). This utterance-final H tone lowering rule also provides strong evidence on establishing the five autosegmental tone schemes that are independent of their TBUs.

As illustrated in the table, when H tone lowering occurs before a pause, all and only the moras of the H part in a tone pattern are affected, as shown in the words *mi-tāāmī* ‘fun’ and *ki-saabí* ‘roof’ which have underlying H and LH tone pattern respectively (see their tonal realisation before another stem in Table 2.11). This fact supports that we are dealing with a H autosegment as (part of) a tone melody that affects on the whole phonological domain rather than multiple successive H tones, in which case only the H tone of the final mora would be lowered.

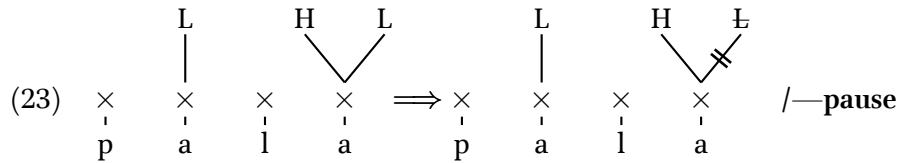
Tone pattern	Realisations before pause	Examples
H	CV̄, CV̄V̄, CV̄CV̄, CV̄VCV̄, CV̄CVCV̄	<i>mi-tāāmī</i> ‘fun’
L	CV, CVV, CVCV, CVVCV, CVCVCV	<i>ki-lili</i> ‘tear’
HL	CV̄, CV̄V̄, CV̄CV̄, CV̄VCV̄, CV̄CVCV̄	<i>ki-fúuma</i> ‘to buy’
LH	CV̄, CV̄V̄, CVCV̄, CVVCV̄, CVCVCV̄	<i>ki-saabī</i> ‘roof’
LHL	CV̄, CV̄V̄, CVCV̄, CV̄VCV̄, CVCV̄CV̄	<i>ki-palâ</i> ‘to go out’

Table 2.12: Realisation of tone patterns in Kukuya lexical stems (before a prosodic break)

An example of the utterance-final H tone lowering is given in (22) on the verb *ki-fúuma* ‘to buy’. In (22a), when there is an element being focused in an immediate preverbal position, a verb-final H tone occurs (also see chapter 4 section 4.3) and raises the HL tone pattern on the verb as H, due to the H tone plateauing effect. This verb-final H tone can only be detected when there is a following nominal prefix on its object complement, onto which this H tone can spread. When the verb appears utterance-finally as in (22b), this final H tone undergoes lowering which bleeds the H tone plateauing rule, and the verb just surfaces with its original HL tone pattern.

- (22) a. Taará ku dzándú ká-fúúm-í má-lí.
 1.father 17.LOC 5.market 1SM.PST-buy-PST 6-wine
 ‘Father bought the wine AT THE MARKET.’
- b. Taará ma-lí ku dzándú ká-fúum-i.
 1.father 6-wine 17.LOC 5.market 1SM.PST-buy-PST
 ‘Father bought the wine AT THE MARKET.’ [*ki-fúuma* ‘to buy’]

A noteworthy exception in the table above is the realisation of HL and LHL tone patterns before a pause. We see that when there are fewer moras than tones in a stem before a pause (as the stems in bold in the table), the final L tone in HL and LHL patterns is disassociated, as in the word *ki-palâ* ‘to go out’ with the underlying LHL tone. This final L tone deletion rule is illustrated in (23). When this rule is applied and renders a H tone exposed on the final mora, we see that the H tone lowering rule is not applicable and the stem surfaces with a final H tone before the pause.



2.2.1.2 Postlexical prosodic domain

We now move on to see how the tone schemes are realised on a lexical stem when it is followed by a *prefix*. I present the analysis proposed in Paulian (1975) and Hyman (1987) that a stem together with the prefix on the following word form a prosodic domain to which the five tone patterns are assigned. When a stem is followed by a word with a H tone prefix or any H tone functional markers (in bold), the realisation of the tone patterns on the preceding stem is summarised in Table 2.13.

Tone pattern	Realisations before a H prefix	Examples
H	ĆV́, ĆV́V́, ĆVĆV́, ĆV́VĆV́, ĆVĆV́ĆV́	<i>bi-táámí bú báana</i> ‘fun of children’
L	CV, CVV, CVCV, CVVCV, CVCVCV	<i>ki-lili kú me</i> ‘my tear’
HL	ĆV́, ĆV́V́, ĆVĆV́, ĆV́VĆV́, ĆVĆV́ĆV́	<i>ki-fiúúmá báa-ntaba</i> ‘to buy goats’
LH	CV, CVV, CVCV, CVVCV, CVCVCV	<i>ki-saabi kú nzó</i> ‘roof of the house’
LHL	CV, CVV, CVCV, CVVCV, CVCVCV	<i>ki-beneme kú me</i> ‘my new born’

Table 2.13: The realisation of different tone patterns in Kukuya on lexical stems (when followed by a H tone prefix)

From the table we notice that the stems with HL, LH and LHL tone patterns, regardless of their syllable structure, are realised as H, L and L respectively when followed by a H prefix. Paulian (1975) and Hyman (1987) accounted for these tone changes by assuming that the sequence *stem+prefix* constitutes a postlexical prosodic domain in Kukuya, which the five fixed tone patterns are mapped onto. The tone modifications on the surface is to avoid some tone occurrences that are not allowed within the five tone pattern. Recall that an autosegmental H tone can occur only once in a tone scheme and successions of more than one segmental H tones in a contour tone pattern are prohibited, we exclude the surface tone realisations *HLH,

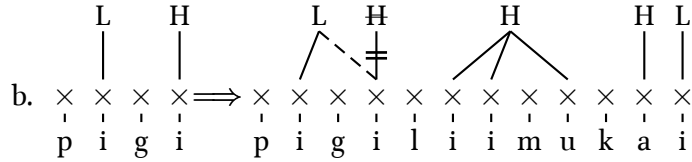
*LHH and *LHLH in a postlexical prosodic domain, which can be triggered by a H tone prefix following the stem. See the illustrations below which are adapted from Hyman (1987), where ‘|’ indicates the boundary of a prosodic domain.

In (24) we find that when followed by a H tone connective marker *líí-*, the lexical stem with an underlying HL tone pattern is realised as H-H. This tone modification is due to the fact that a HLH tone pattern is not allowed in a prosodic domain of the language, thus the L tone on the mora is deleted and then either the preceding H tone spreads to this mora or the L tone is raised to H by the adjacent H tones. This is also known as H tone plateauing (Hyman and Katamba 2010) which is also the case in example (24a) above.

- (24) a. ma-téme(HL) \implies ma-|témé líí (H) | me ‘my hoes’
- b. $\begin{array}{cccccccccccccccc} & & L & & H & & L & & L & & H & L & & H & & L \\ & & | & & | & & | & & | & & / \backslash & \# & & / \backslash & & | \\ \times & \times & \times & \times & \times & \times & \times & \implies & \times & \times & \times & \times & \times & \times & \times & \times \\ | & | & | & | & | & | & | & & | & | & | & | & | & | & | & | \\ m & a & t & e & m & e & & & m & a & t & e & m & e & l & í & í & m & e \end{array}$

The tonal realisation of a LH stem followed by a H tone prefix is illustrated in (25). We see that in this case the L tone copying rule as in (21) is applied, which may be due to the Obligatory Contour Principle (OCP) in which two consecutive H tones are avoided. What is unexpected here is that the L tone does not replace the first H tone on the connective marker and a surface L-L-H-H tone sequence is allowed in the postlexical domain. In (25) we also notice that the H tone on the connective marker spreads to the following nominal prefix *mu-* that has a default L tone, which results multiply linked H tones to which the L tone on the stem cannot spread. Given this exception, Hyman (1987) proposed that the L tone spreading rule just does not apply in the postlexical domain and interpreted the tonal realisations in (24) and (25) as any tone must be deleted when preceded by any tone and followed by a H tone in the postlexical domain, thus collapsing the H tone plateauing and L tone copying rules, as schematised in (26).

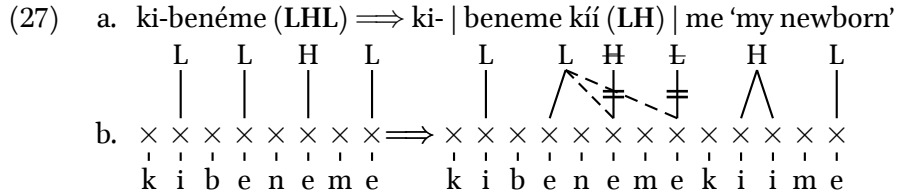
- (25) a. pigí (LH) \implies pigí líí mú- (LH) | kái ‘raphia palm of the woman’



- (26) a. HL + H → H ... H
 LH + H → L ... H
 b. *Tone deletion in the postlexical domain* (Hyman 1987: 323)

$$T \rightarrow [\emptyset] / T_H$$

As for a LHL stem followed by a H tone prefix, which is shown in (27), according to Hyman (1987) its realisation as LH can be interpreted as the rule in (26) applying iteratively, deleting the final L tone and the H tone on the stem, then the initial L tone may spread to these moras the tones of which are deleted. It does not seem to be case that the LHL tone pattern is assigned to the whole postlexical prosodic domain, otherwise we would expect the surface L-H-L-L or L-L-H-L tone sequences on this domain.



From the examples above we also notice that the prefix of a nominal stem has nothing to do with the tonal realisation of this stem. For example in the possessive construction in (24) above, the prefix plus the stem surface a L-H-H sequence but the L tone of the prefix never spreads rightwards. This extra-prosodic status of the initial prefix further supports that the postlexical prosodic domain in Kukuya consists of a stem and a following prefix, while the prefix of the stem itself does not join into the prosodic domain of the stem. Since the prefix cannot spread its tone to the right, and the tone spreading rule is assumed to be always in a left-to-right fashion, we find that the tone on the first mora of the stem is stable (except when it bears a H tone and occurs utterance-finally, which undergoes lowering). I

will discuss further on the particularity of the stem-initial mora in section 2.2.2.

In Paulian (1975) and Hyman (1987; 2007)'s data and analysis of the postlexical prosodic domain, they mainly took into account the nominal phrases, for example a nominal stem followed by a prefix such as an connective marker or a demonstrative marker, while the prosodic domain within VP was not illustrated in much detail. Only one relevant example (28a) is found in Paulian (1975), in which she demarcated three prosodic domains within the sentence (the domain boundary is marked by '|'): the subject together with the subject marker, the verb stem, and the object (with no segmental prefix). The LH tone pattern on the subject is modified as L-L because of the rule in (26) above. Example (28b) is extracted from Hyman (2007) and the HL subject *téme* followed by the H tone subject marker is subject to H tone plateauing.

- (28) a. |pigi lí:|pî mba:
 5.raphia 5SM.PST-get.PST 9.fire
 'the raphia palm was on fire' (Paulian 1975: 141)
- b. Mà-témé máá-bvì.
 6-hoe 6SM.PST-fell.PST
 'The hoes fell.' (Hyman 2007: 208)

I suppose that the prefix on an object complement also forms a prosodic domain with the preceding verb stem in the way subject and subject prefix do, and this is borne out in (29) and (30). In (29) the HL tone pattern on the preverbal object is changed to H, when followed by a H tone subject prefix on the verb. In (30) the tone pattern on the relative verb also shifts from HL to H, due to the emergence of a verb-final floating H tone in non-subject relatives, and the H tone is also carried onto the following nominal prefix of the object.

- (29) Ndé má-láálá káá-fúum-i.
 1.PRO 6-orange 1SM.PST-buy-PST
 'S/He bought the ORANGES.' [ma-láala 'oranges']

- (30) mi-fémé mi-kíí-fúúm-í mú-kái
 4-pig 4REL-7SM.PST-buy-PST 1-woman
 ‘the pigs that the woman bought’ [*ki-fúuma* ‘to buy’]

So far I have presented the assignment of the five autosegmental tone schemes to stems of different syllable structures in both the lexical and postlexical prosodic domains, as well as some associated tone distribution and modification rules. Following Paulian (1975) and Hyman (1987), I have shown that in Kukuya a postlexical prosodic domain consists of a stem plus a following prefix. In the previous subsections I mainly introduced the representation of lexical tones, but we have also encountered grammatical tones such as the final H tone on a non-subject relative verb. Next I present more cases of different types of grammatical tones.

2.2.1.3 Grammatical tones

In section 2.1.3 I have shown that a nominal or verbal stem in Kukuya can contain maximally three moras, which may lead to the loss of some TAM affixes and all the verbal extensions. This result of shortened shape of the stem and morphological attrition could enhance the role of tones, since some grammatical categories carried by the lost morphemes need to be marked in an alternative way (Nurse 2008). In this subsection, I introduce the application of grammatical tones in Kukuya, most of which are floating tones. I will present the metatony effect, grammatical tones for TAM marking, predicative H tone and tone patterns in relative clauses.

Let us first consider the so-called metatony effect in Kukuya. Metatony can be defined as a tonal process whereby in certain TA forms a H tone replaces a L or falling (HL) tone on post-radical syllables (which corresponds to the final vowel of a verb stem in Kukuya) as well as on a following non-accentuated L tone syllable, if and only if the verb is followed by other material such as object or adverbial (Meeussen 1967; Nurse 2008; Hyman and Lionnet 2011). In (31), we find that when immediately followed by an object, the tone on the final vowel of the infinitive verb is substituted by a H tone, and the nominal prefix of the object which has a default L tone

distinctions are usually placed on the subject marker as well as verb-finally. In the recent past tense example (33a), we see that the class 2 subject marker takes a H tone. In the remote past sentence (33b), in addition to the use of the auxiliary *âli*, the tone on the subject marker shifts to HL. In these two past tense sentences we notice that the metatony effect does not show up, which may indicate that there is a floating grammatical L tone which replaces the metatonic H tone. In the future tense (33c) the tone on the subject marker is also HL, while a verb-final H tone appears and spreads onto the following prefix, which may be just the metatonic H tone. From these examples we see that both grammatical H and L tones are involved in tense marking. Mood distinctions such as imperative and subjunctive are also tonally marked, see a detailed description in section 2.4.2.

- (33) a. Bó **báá-fúum-i** ma-ko.
 2.PRO 2SM.PST-buy-PST 6-banana
 ‘They bought some bananas.’ [recent past]
- b. Bó *âli* **báá-fúum-i** ma-keé.
 2.PRO RPST 2SM.RPST-buy-PST 6-tobacco
 ‘They had bought some tobacco.’ [remote past]
- c. Bó lí **báá-fúúm-á** má-nzó.
 2.PRO FUT 2SM.FUT-buy-FV 6-house
 ‘They will buy some houses.’ [far future]

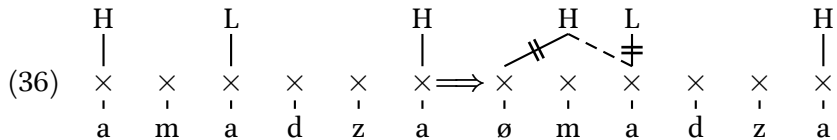
Relative clauses are also marked by means of grammatical tones. In both subject and non-subject relatives, the REL-SM sequence on the verb always surfaces with a LH tone scheme. Subject relatives differ from non-subject relatives in the verbal tone. We see that in the non-subject relative (34a) which is repeated from (30), there is an emergent verb-final H tone which triggers H tone plateauing on the verbal tone and spreads onto the following nominal prefix; while in the subject relative (34b) this H tone does not appear. The tone pattern of a negative sentence is also discerned from an affirmative, on which I refer to more detail in section 2.5.

- (34) a. mi-féme mi-kíí-fúúm-í mú-kái
 4-pig 4REL-7SM.PST-buy-PST 1-woman
 ‘the pigs that the woman bought’
 b. mu-kái wǔ-fúum-i mi-féme
 1-woman 1REL-buy-PST 4-pig
 ‘the woman who bought the pigs’

Grammatical tones are also attested on nominal prefixes which have default L tone in the citation form. The nominal prefix of predicative NP and the concord prefix of a predicative adjective bear a H tone, as shown in the examples in (35a). A preverbal focused NP also takes a H tone prefix, as shown in (35b) (see their connections in chapter 4).

- (35) a. mu-tí ‘a tree’
 Mú-tí. ‘It is a tree.’
 ma-téme ma-bvé ‘good hoes’
 Ma-téme má-bvé. ‘The hoes are good.’
 Ndé mú-tsúli. ‘S/He is a goldsmith.’
 b. (*Did the woman give the fish to the DOGS?*)
 Mu-kái baa-ntsúli báa-ndzulí káá-wí.
 1-woman 2-fish 2-cat 1SM.PST-give.PST
 ‘The woman gave the fish to the CATS.’

The H tone on a predicative NP may have its origin in a H tone augment marking predication which was lost later, while the H tone was retained and was attached to the nominal prefix. The potential derivation of the predicative H tone on the word *ma-dzá* ‘water’, which may have an augment *a- in an earlier stage, is illustrated in (36).



To summarise, in this subsection I have discussed the lexical and postlexical prosodic domains onto which the five tone patterns are distributed, as well

as different types of grammatical tones in Kukuya. We have already noticed that there is no interaction between a prefix and a consonant-initial stem in this language, and that the first mora on a stem is never subject to tone changes. In the next subsection we consider in detail how this stability of the stem-initial mora characterises the prosodic system of the language.

2.2.2 Stem-initial accent

In the introduction of the segmental phonology, we have seen that the initial syllable of a stem shows greater licensing capacity for both onset consonants and vowels than other syllables, and tone realisation presented above also shows that the tone on the stem-initial mora is always stable. In this subsection I describe this interesting phenomenon in Kukuya, which is labeled as the “positional prominence” or “stem-initial accent” (Hyman 2007, 2015; Lionnet 2017). Here I adopt the term “accent” and refer it to an abstract property that can influence the licensing potential of consonants and vowels as well as the realisation of tone patterns. There are many other Northwest Bantu languages that have been reported as having a stem-initial accent, for example in many zone A languages an “exaggeration” of the duration of the initial consonants is attested (Idiatov and Van de Velde 2016), which can influence the whole realisation of the initial mora. Since Northwest Bantu languages often impose size constraints on their prosodic stems, the role of the stem-initial syllable may be enhanced.

Here I summarise some phonological properties of the stem-initial mora in Kukuya, some of which we have already encountered in the description above:

- the full inventory of consonants (thirty-one phonemic consonants as in Table 2.1) can occur as the onset of the initial syllable, as opposed to only six “archi-phonemes” that are allowed in other positions in the stem; while the voiceless initial consonants are often realised with aspiration, it is not attested with intervocalic consonants; consonant lenition occurs in intervocalic positions but never in the initial position;

- the “constricted articulation” which can influence the realisation of the whole syllable is only attested within the stem-initial position;
- the initial consonants /m/, /n/ and /l/ are always geminated in deliberate speech as in (37), though this gemination could become hard to notice in rapid utterance; for other initial oral consonants, according to Paulian (1975: 137) there is a very slight pause preceding them in preparation for the accentuated initial syllable, which means that the duration of initial consonants is relatively lengthened;

(37) ki-lóbo [ki-llóbo] ‘to fish’
 ki-na [ki-nna] ‘moment’

- a nasal consonant in the stem-initial position never triggers nasalisation on its neighbouring vowels, while intervocalic nasals can do; the deletion of bilabial consonants /m/ and /b/ as a casual speech rule is never applied to the initial consonant (see section 2.1.2);
- the domain-initial tone is always immutable: it is never affected by any tone that occurs to its left or right; the predicative H tone cannot be imposed to the stem-initial mora, as shown in (38) ; it is realised as a floating H tone by a movement of the larynx before the articulation when the initial consonant is oral, and is realised on a geminated consonant when it is nasal or lateral; the only exception is that a stem with underlying H tone pattern is subject to the utterance-final lowering rule which is also applied to the initial mora;

(38) papí [‘papí] ‘it is a wing’
 meé [méé] ‘it is a forging stone’ (Paulian 1975: 137)

- the vowel reduction operates on V₂ which is an immediately post-accentual unaccentuated mora, such that /a/ in this position is often realised as a schwa, which is also similar to that typically found in stress languages;

A prefix, as an unaccentuated unit, does not fall into the same phonological domain with the stem that it selects, though they constitute one grammatical word. Some segmental evidence could suggest the prefix to be phonologically autonomous from the stem. Here we consider the prefixes including

noun class prefixes, adnominal concord prefixes, connective markers and subject markers.

- the possible shapes of prefixes in Kukuya are CV and CVV, while VV shape is only attested in connective markers, which may be due to diachronic phonological reduction; interestingly, the C in all the prefixes is restricted to /b/, /k/, /m/ and /l/ which is a subset of the six consonants allowed in the intervocalic positions of the stem; this may just be coincidental considering that these prefix shapes are attested cross-Bantu and have been reconstructed, so we cannot ascribe the limited consonant inventory on the prefixes to the same restrictions on the intervocalic consonants; however this fact still clearly sets the prefix apart from the stem, as the constraints on consonant combinations within the stem do not apply to prefixes;
- the vowels in a prefix may be /i/, /a/ and /u/; according to the restricted vowel combinations on the stem (see Table 2.7), the vowel combination **u-o-o** in the words in (39) is not allowed stem-internally, which means that the prefix must be domain-external;

(39) bu-nóno [bu-nnóno] 'selfishness'
mu-kokó [mu-kokó] 'king'

- the /m/ and /l/ on the prefixes are never geminated, which shows that they are never assigned with prosodic accent; the prefixes with nasal consonant /m/ can nasalise the adjacent vowels in the utterance; and the prefixes with /b/ are often reduced to V in casual speech with the bilabial stop being dropped; the prefix /k/ is subject to lenition and is often realised as [g] or sometimes [ɣ] when preceded by other elements; these phenomena clearly reflect domain-internal properties of the prefix rather than the initial accent;

Based on these facts presented above, the phonological prominence of the stem-initial syllable in a prosodic domain and the autonomous domain-internal characters of the prefix are clearly demarcated. The accentuated stem-initial mora functions as the core of the prosodic domain, attracting all the non-accentuated elements that follow it until the next accentuated

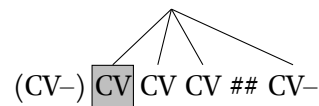
mora. The non-accentuated prosodic status of the prefix leads it to join the preceding stem to form a prosodic domain. It is not clear from the diachronic point of view how the maximality constraint on the syllables of a stem is associated with the stem-initial accent and the analytic tendency of Northwest Bantu, which may need to be carried out in future research.

To conclude this section on the prosody system of the language, some key rules and principles are summarised in (40). I have shown that there are five independent tone melodies that are regularly assigned onto a certain prosodic domain under some distribution rules. This prosodic domain can be lexical which is just the lexical stem of a nominal or a verb, or postlexical which consists of the stem with the following prefix of the next word. Like in many other Northwest Bantu languages, there is always a stem-initial accent delimiting the left boundary of a prosodic domain in Kukuya, which reflects the phonological strengthening of the accented material, both segmentally and tonally.

- (40) a. *Autosegmental tone patterns in Kukuya*
 L, H, LH, HL, LHL
- b. *Low tone copying rule*
 Copy the L tone (if present) to the next syllables until the whole tone pattern is distributed to all moras.
- c. *Tone deletion in the postlexical domain* (Hyman 1987: 323)

$$T \rightarrow [\emptyset] / T_H$$

- d. *Stem-initial accent and postlexical prosodic domain:*
 stem + prefix(es) on the next word



2.3 Nominal morphology

This section describes the nominal morphology of Kukuya. I first provide an overview on the noun class system of the language, as well as the rules in nominal derivation. Then I look into the structure of noun phrases, discussing the form and use of the connective, demonstrative, adjective, quantifiers, numerals, interrogatives and pronouns. A table on the concord relations within noun phrases is given at the end of this section.

2.3.1 Noun classes

Nominals in Kukuya have the common property of Bantu languages as belonging to different noun classes which are morphologically distinguished by the nominal prefixes, verbal agreement and concord with modifiers in the noun phrase. In the generative tradition, noun classes in Bantu are considered as a grammatical gender system with gender-specific spell-out of number features (Carstens 2008, 2011). Throughout the thesis, I refer to the noun classes using the numbering system established in Meinhof and Van Warmelo (1906/1932). In this subsection, I first provide a brief overview of each noun class, and how they are grouped into different singular/plural gender pairs. Then I present different types of nominal derivation, such as verb-to-noun derivation, compound nouns and the formation of diminutives. Since Paulian (1975) has conducted a very comprehensive description on the noun class system of this language, my presentation here is also based on her analysis, especially on her calculation of the frequency of each class, with my own observations and new data collected in the fieldwork.

2.3.1.1 Noun classes overview

The phonological shape of a noun class prefix in Kukuya can be CV-, Ø- or a nasal, and the underlying tone on it is always L. The vowel on the prefix is lengthened before a prenasalised consonant NC, thus CV- becomes CVV-, which can be interpreted as a compensatory lengthening due to the

moraicity of the nasal which contributes to a unit of length and transfers it to the preceding vowel (Clements 1986; Hyman 2018: 139). A list of all the noun classes with examples is given in Table 2.14. Next I present the morphological and semantic properties of each noun class.

Class	Prefix	Examples
1	∅-/mu-	<i>mpúku</i> 'rat'; <i>mu-ndziá</i> 'foreigner'
2	<i>ba-</i>	<i>ba-bhii</i> 'hunters'; <i>baa-mbvi</i> 'grey hair'
3	<i>mu-</i>	<i>mu-tíma</i> 'heart'; <i>muu-ɲaa</i> 'whistle'
4	<i>mi-</i>	<i>mi-dza</i> 'roots'; <i>mhiilí</i> 'feet, legs'
5	∅-/li-	<i>dzukí</i> 'frog'; <i>li-fukí</i> 'smell of an animal'
6	<i>ma-</i>	<i>ma-kilá</i> 'blood'; <i>maa-mphiini</i> 'magic power'
7	<i>ki-</i>	<i>ki-sáani</i> 'orphan'; <i>kii-ndomó</i> 'sheep'
8	<i>bi-</i>	<i>bviila</i> 'possessions, food'; <i>bi-bu</i> 'covers, shells'
9	N-	<i>ngwaalí</i> 'morning'; <i>mfúra</i> 'envy'
10	N-	<i>nsaá</i> 'seeds'; <i>mpeé</i> 'fish scales, shells'
14	<i>bu-</i>	<i>bu-sú</i> 'face'; <i>buu-pfúru</i> 'shame'

Table 2.14: Noun classes in Kukuya

Class 1

The class 1 nominals are characterised by taking either the prefix *mu-*, or starting with a nasal with no separable prefix, see example words in (41).

- (41) a. *mu-bhii* 'hunter' (*ki-bhia* 'to hunt')
mu-pfúuru 'elder person'
mu-ndziá 'foreigner'
mu-tsúli 'goldsmith' (*ki-tsúla* 'to forge')
- b. *mbuurú* 'person'
nzálí 'river'
ngampíka 'younger twin'
nkaná 'acne'

The shape in (41b) without the prefix *mu-* is quantitatively much more attested (85% according to Paulian) in class 1 nouns, while the prefix *mu-*

mainly occurs with derived nouns. These are plausibly former class 9 nouns which have been reclassified in class 1 mainly on the basis of animacy as in other Teke varieties (cf. Hyman et al. 2019; Bollaert et al.), but Kukuya seems to have gone a step further by also reclassifying inanimate nouns of class 9.

Class 2

The prefix of a class 2 nominal is typically *ba-*, which can be realised as *baa-* when the stem begins with N or NC, i.e. when the prefix is used to pluralise a class 1 noun in group (41b). The shorter prefix *ba-* alternates with the class 1 prefix *mu-*. Some examples of class 2 nouns are shown below:

- (42) a. *ba-bhii* 'hunters'
ba-tsúli 'blacksmiths'
ba-boli 'lazy persons'
ba-téléké 'grandsons'
- b. *baari* 'persons'
baa-nkéle 'siblings'
baa-mbvi 'grey hair'
baa-nzáli 'rivers'
baa-mpúri 'eye boogers'

From (41) and (42) we can also see that while class 1 and 2 in Bantu typically refer to humans and some animals, it is also the case in Kukuya but not restricted to it.

Class 3

A class 3 noun takes the prefix *mu-* which is of the same shape as the class 1 prefix, which is realised as *muu-* before prenasalised NC and nasals /m/ and /ŋ/. The short form *mu-* is usually attested when the noun is derived, even the stem starts with a nasal. Some examples are given in (43):

- (43) a. *mu-tíma* 'heart'
mu-bilí 'kola tree'

- mu-ŋaani* ‘threat’ (*ki-ŋaani* ‘to threat’)
mu-ŋono ‘sound of rain’ (*ki-ŋono* ‘to rumble’)
mu-miká ‘trial’ (*ki-miká* ‘to try’)
- b. *muu-ŋaa* ‘whistle’
muu-mhuuku ‘pestle’
muu-nkáání ‘letter’

In many Teke varieties, class 1 and class 3 have merged, which may be due to their identical nominal prefix shape (Hyman et al. 2019), and this tendency is also observed in Kukuya. However, class 3 can still be distinguished from class 1 by its plural class 4 counterpart and the H tone connective marker that is different from class 1 (see section 3.2.1).

Class 4

The prefix of class 4 nouns takes the short shape *mi-* when alternating with the class 3 prefix *mu-*, and is lengthened as *mii-* when replacing *muu-*. See examples in (44):

- (44) a. *mi-tí* ‘trees’
mi-dza ‘roots’
mi-bvî ‘maggots’
mi-lá ‘intestines’
myáaka ‘hands, arms’
- b. *mhiílí* ‘feet, legs’
mii-mfi ‘farts’
mii-ŋeé ‘grasshoppers’

Semantically, class 3 and 4 nominals in Kukuya can have miscellaneous referents, most of which are plants (32.7% according to Paulian’s calculation), insects and small animals, things of vegetable origin or related to soil, as well as body parts.

Class 5

A class 5 noun takes the prefix *li-*, which is realised as *lii-* when the initial

consonant of the stem is a prenasalised NC or nasal, and more commonly it surfaces with no prefix when the stem begins with an oral consonant. See the examples in (45):

- (45) a. *li-nsá* 'tear'
li-fukí 'smell of an animal'
li-yǔ 'peanut'
lii-mpí 'kidney'
- b. *beemé* 'pigeon'
dzukí 'frog'
fungúla 'key'
búku 'portion'

As we will see shortly, the class 5 nouns without prefix have their plural counterparts in class 6, while the nouns with the prefix *li-* form a gender pair with class 10.

Class 6

The prefix of class 6 nouns typically has the shape *ma-*, which again is realised as *maa-* before a prenasalised or nasal consonant, and as *ma-* before an oral consonant. Some examples are illustrated in (46):

- (46) a. *ma-táki* 'buttocks'
ma-lá 'villages'
ma-kilá 'blood'
ma-fungúla 'keys'
mhúni 'teeth'
- b. *maa-ndoko* 'misfortune'
maa-júru 'bodies'
maa-mphiini 'magic power'
maa-mpópi 'judgements' (*ki-wópo* 'to judge')

The class 5 and 6 nominals can form a singular/plural pair, whose semantic content is very diverse: they can refer to plant names, animals and body

parts. I will show later that a class 6 noun can also form a gender pair with many other noun classes (see Table 2.15 below), or it can refer to mass or abstract nouns and under that reading does not have any singular counterpart.

Class 7

The prefix of class 7 nouns is *ki-*, which is again lengthened as *kii-* before prenasalised and nasal consonants as well as in derived nouns. See some examples in (47):

- (47) a. *ki-sáani* ‘orphan’
ki-tuura ‘spit’
ki-fúla ‘wind’
ki-ku-ko ‘small banana’ (diminutive)
- b. *kii-mfíli* ‘nail’
kii-ŋee ‘sour fruit’
kii-ndomó ‘sheep’
kii-taará ‘paternal family’ (*taará* ‘father’)

The class 7 nouns in Kukuya usually refer to miscellaneous “things” and “objects”, the words *ki-lóko* ‘thing’ and *ki-má* ‘what’ also belong to class 7. Diminutives and some other derived nominals also take the class 7 prefix, as will be shown later in the nominal derivation rules. We will see in the rest chapters, a class 7 subject marker can function as the default subject marking strategy in many cases when the subject of a sentence is unknown but inferably non-human. An infinitive verb in Kukuya also takes the prefix *ki-* which is formally identical to the class 7 nominal prefix. I will show in section 2.4 that infinitives does not behave like typical nominals in Kukuya.

Class 8

The prefix of class 8 nouns is realised as *bi-* or *bii-*, which alternate with the class 7 prefixes *ki-* and *kii-* respectively. Note that the class 8 prefix *bi-* is realised as *bvi-* when it is inseparable from the stem. Some examples are given below:

- (48) a. *bi-bu* 'covers, shells'
bi-pee 'seats'
bi-bínu 'heels'
bviila 'possessions, food'
bvhiibi 'weed'
- b. *bii-mú* 'bladders'
bii-na 'moments'
bii-bu-báana 'small children'

The class 8 nominals in Kukuya form a gender pair only with class 7, which can refer to objects, plants, small animals and especially birds. This is exceptional to the other plural noun classes which can match with multiple classes, as seen in Table 2.15.

Class 9

Class 9 nouns systematically lack a separable prefix and overwhelmingly start with a prenasalised consonant. Some examples are given in (49):

- (49) *mpáki* 'branch'
mpúru 'anger'
ngwaalí 'morning'
mbuka 'place'
mfáribá 'armpit'
ɲkúúlá 'road'
nsilí 'curse'
mfúra 'envy'

Except for some class 9 nouns that do not have number opposition (see the discussion below), the plural counterparts of all other class 9 nouns are in class 6. A semantic generalisation of the class 9/6 gender pair is that it does *not* include animals, plant names or humans, some abstract notions related to supernatural powers are included in this pair.

Class 10

The nouns of class 10 also begin with a prenasalised consonant, just similar to the class 9 nouns. Some example are given below in (50):

- (50) *mfú* 'hair (pl.)'
nzála 'claws, nails'
nzǔ 'peanuts'
mpúpí 'peanut skins'
nsaá 'seeds'
nsísálá 'eyelashes'
mpeé 'fish scales, shells'

Class 10 is a plural class, and all the plural nominals of this class have their singular counterparts in class 5 nouns which take the prefix *li-*. This class 5/10 gender pair mainly contains nominals of fruits as well as objects of long and thin shape. We notice that in this gender pair, the prefix *li-* on the singular nouns in class 5 seems to be replaced by a nasal prefix in the plural forms in class 10, as seen from the oppositions in *li-yǔ* (cl.5)/*nzǔ* (cl.10) 'peanut(s)', *li-yála* (cl.5)/*nzála* (cl.10) 'claw(s), nail(s)' and *li-saá* (cl.5)/*nsaá* (cl.10) 'seed(s)'. It is likely that class 10 nouns have a nasal prefix [N]- which has fused with the noun stem and becomes inseparable, also triggering the neutralisation on the initial continuant oral consonant (see the rules in (6) in section 2.1).

Class 14

The prefix of class 14 nouns is *bu-*, which is commonly attested as *buu-* when the noun is derived from other classes and has no separable prefix in its original form (see the discussion on the lengthening of noun class prefixes below). Again, some examples of this noun class are given here:

- (51) a. *bu-lá* 'village, city'
bu-ka 'cassava'
bu-tá 'gun'
bu-sú 'face'

bu-tí ‘mascot’

- b. *buu-ɲama* ‘animacy’
- buu-mpalá* ‘adultery’
- buu-pfúru* ‘shame’
- buu-tsúku* ‘night’
- buu-koló* ‘this evening’

Most class 14 nominals express abstract meaning (without magic connotation) or are related to time expressions, thus do not have plural forms, while this class also includes some nouns such as names of mushrooms which have their plural counterparts in class 6.

There are two points that I want to discuss a bit further here on the shape of the nominal prefixes. The first is the occurring condition of the long prefix CVV-. We have seen above that CVV- is realised before prenasalised initial consonants, which can be explained by the compensatory lengthening due to the integration of the nasal which counted as a unit of length into the stem. To explain, for example the class 9 nouns all start with an NC, we can propose that there exists a structural prefix for class 9 which is |N|, and the initial NC can be interpreted as the result of the fusion of |N| and a noun stem that begins with an oral consonant. The prefix |N| cannot be separated from the initial NC, since the plural class 6 prefix *maa-* cannot substitute |N| but is attached to it. The class 6 prefix is realised as *maa-* because the prosodic length that the |N| should contribute is compensated by lengthening the preceding vowel.

Under this analysis, the noun stems that take a CVV- shape prefix and start with NC should be assumed to have this structural prefix |N|, thus unifying the occurring condition of the long prefix. However, we have also encountered cases in class 7 and 14 in which the CVV- prefix appears when the nominal is derived from other classes. To account for this, Paulian (1975) further suggested that the derivation process may involve two intermediate stages in which the original noun takes the nasal prefix |N| and then takes the class 7 and 14 derivational prefixes *ki-* and *bu-*, and compensatory lengthening happens, resulting the long realisation of the prefixes. Since there is no clear clue on the ever existence of the prefix |N| in these cases,

and we also see that in class 1 and 3 the derived nouns instead take the short form of the prefix, it remains a puzzle why long prefix is used in class 7 and 14, I leave this for further research.

The other question is how to deal with the nouns whose nominal prefix CV- seems to be merged with the stem and becomes inseparable, such as the words in (52) which are all realised as CVVCV shape.

- (52) *mhéele* (cl.1) ‘moonlight’
báana (cl.2) ‘children’
mhiilí (cl.4) ‘legs, feet’
maalí (cl.6) ‘oil’
mhúini (cl.6) ‘teeth’
kyaali (cl.7) ‘chicken nest’
bvhiila (cl.8) ‘things, possessions’

Paulian (1975) proposed an abstract morphophoneme |Y| for all the nouns of this type and supposed that the noun stems that start with |Y| are always fused with the class prefix CV-, for example *bvhiila* is underlyingly |bi-Yila| and *maalí* is |ma-Yalí|. The |Y| is realised phonologically by lengthening the V₁ of the noun stem and the constricted articulation of the first syllable (see section 2.1.1). Here I just present this hypothesis but leave more analysis for future studies.

In the above I have briefly displayed the forms and semantic content of noun classes (not including the locative classes) in Kukuya. I also mentioned that as typical in Bantu, noun classes in Kukuya are characterised by a series of binary complementary oppositions that are always realised as singular-plural pairs, which is also referred to as grammatical genders in Bantu. We have also observed that singular and plural noun classes are not paired in an “one-to-one” fashion, i.e. a noun class can form a gender pair with multiple other classes respectively, which may be explained by the historical loss and merging of different noun classes. A certain noun class in different gender pairs may also encode distinctive semantic properties. I summarise the main gender pairs in Kukuya in Table 2.15. I also include isolated noun classes that do not have number distinction, which mainly

Gender	Prefix	Examples	Semantics
1/2	∅-, <i>mu-/ba-</i>	<i>mpúku</i> <i>baa-mpúku</i> 'rat(s)'	mainly humans, animals
3/4	<i>mu-/mi-</i>	<i>mu-bílí</i> <i>mi-bílí</i> 'kola tree(s)'	plants, insects, small animals
5/6	∅-/ma-	<i>taabi</i> <i>ma-taabi</i> 'puddle(s)'	plants, animals, miscellaneous
7/8	<i>ki-/bi-</i>	<i>ki-sáani</i> <i>bi-sáani</i> 'orphan(s)'	objects, birds, plants, body parts
9/6	N -/ma-	<i>ndoko</i> <i>maa-ndoko</i> 'misfortune(s)'	abstract, magic, non-animal/plant
5/10	<i>li-/ N -</i>	<i>li-saá</i> <i>nsaá</i> 'seed(s)'	fruits, long and thin objects
14/6	<i>bu-/ma-</i>	<i>bu-sú</i> <i>ma-sú</i> 'face(s)'	miscellaneous, mushroom names
5	∅-, <i>li-</i>	<i>li-fukí</i> 'smell of animal'	abstract, miscellaneous
6	<i>ma-</i>	<i>ma-kilá</i> 'blood'	liquid
9	N -	<i>nsílí</i> 'curse'	abstract, magic
14	<i>bu-</i>	<i>buu-ndúku</i> 'friendship'	derived abstraction

Table 2.15: Singular/plural pairs in Kukuya

include nominals encoding abstract meaning or referring to uncountable things. Some marginal pairs that only have one or two entries are not listed.

Locative classes 16, 17 and 18

There are three locative class-like morphemes *ɲa*, *ku* and *mu* in Kukuya which correspond to the Proto-Bantu locative prefixes **pa-*, **kú-* and **mú-* of classes 16, 17 and 18. Nevertheless, in Kukuya these three morphemes are most commonly attested as preposition-like elements which are placed before a noun phrase where necessary, while they can also appear as prefixes in the verbal and clausal domains, but never as real locative class prefixes in the nominal domain. In this part, I will show that while these three morphemes have maintained some typical uses of locative classes (thus I still gloss them with the numbers 16, 17 and 18), some extended functions are innovated. Simple examples of their prepositional use are first given in (53).

- (53) *ɲa nzó* 'in the house'
ɲa ngwaalí 'in the morning'

ŋa bulá ba ‘in this village’
ku kú me ‘in my place’
ku dzáandu ‘at the market’
mu ntsié ‘on foot’
mu mu-súru ‘in the forest’
mu nkólónkóló ‘slowly’
ki-mwáálá má-dzá mu nsakí ‘to water (on) the fruits’

Semantically, the functions of these three locative markers are not restricted to distinguishing the positions of references as their typical use in Bantu, but they have developed other functions as prepositions. Some examples on the use of the class 16 locative marker *ŋa* are given in (54). *ŋa* is the only one of the three that can be used pronominally, meaning “here” as in (54a). In most other cases, *ŋa* functions as a proximal locative marker, or a temporal preposition. We will also see in section 2.5 that *ŋa* can be cliticised onto a complementiser and select conditional or temporal clauses.

- (54) a. We á-síil-i **ŋa** mu ma-tsúku
 2SG.PRO 2SG.SM.PST-stay-PST 16.PRO 18.LOC 6-day
 kwê?
 how.many
 ‘You stayed here for how many days?’
- b. Mwáana áá-bvĩ ntsína **ŋa** kaarí aa
 1.child 1SM.PST-fell.PST 9.ground 16.LOC 9.middle 9.CONN
 maa-ntsieme máá nká ma-máa-sî kingá mu-tí.
 6-antler 6.CONN 1.deer 6REL-6SM.PST-do.PST like 3-tree
 ‘The child sat between the antlers of the deer which look like a tree.’
- c. Gilbert â-kalá **ŋa** mbuká yi-báa-dzhiik-á
 Gilbert 1SM.FUT-stay 16.LOC 9.place 9REL-2SM.FUT-bury-FV
 ku mfai bu-kyá.
 17.LOC 9.Brazzaville 14-tomorrow
 ‘Gilbert will stay at the place where (people) bury in Brazzaville tomorrow.’

The class 17 *ku* usually has a prepositional use for distal referents according to the speakers. Some examples with different placements of the class 17 *ku* are given in (55). The *ku-* as a verbal prefix, when occurring on a relative verb, is the only of the three which retains the locative reading, as in (55b, c). I will present more innovated functions of *ku-* as a verbal prefix in the description of relative clauses.

- (55) a. We **ku** ndúku aa we kí-má ké
 2SG.PRO 17.LOC 1.friend 1.CONN 2SG.PRO 7-what 7.PRO
 á-fúum-i?
 2SG.SM.PST-buy-PST
 ‘What did you buy from your friend?’
- b. **ku** nzó kǔ-lí we
 17.LOC 9.house 17REL-be 2SG.PRO
 ‘in the house where you are’
- c. Kǔ-dzií me píri líi-kab-a ɲama
 17REL-please 1SG.PRO 1.SG.COMP 1PL.SM.FUT-share-FV 1.animal
ku nzó aa we.
 17.LOC 9.house 9.CONN 2SG.PRO
 ‘Where I like is to share the animal in your house.’

The class 18 *mu* has the most diverse functions among the three, notably expressing cause and reason. Some examples on its use as a reason preposition are illustrated in (56). We notice that the *mu* must select a nominal or a nominalised element, as seen from the relativisation strategy in (56c), which shows that it is indeed used as a preposition.

- (56) a. Mi-pará mi-kíí-wí taará míí-bol-í
 4-money 4REL-7SM.PST-give.PST 1.father 4SM.PST-rot-PST
mu mvúla.
 18.LOC 3.rain
 ‘The money that father gave rotted because of the rain.’

- b. Ndé áá-síik-i me **mu** ma-kâ-n-sá
 1.PRO 1SM.PST-thank-PST 1SG.PRO 18.LOC 6REL-IMPF-1SG.SM-do
 me ndé.
 1SG.PRO 1.PRO
 ‘He thanked me for what I (usually) do for him.’
- c. Mwáana li ya kí-sáábí **mu** taará wŭ-fúum-i
 1.child COP with 7-happiness 18.LOC 1.father 1REL-buy-PST
 ndé ntaba.
 1.PRO 1.goat
 ‘The child is happy that father bought him/her a goat.’

To express “can, be able to”, a verb-like element *-kuki* is always used and it must be followed by *mu*, which then selects an infinitive. The exact category of *-kuki* is still not clear to me. Two example sentences are given in (57).

- (57) a. Me wu-m-pééná mí-pará ka-n-kukí **mu**
 1SG.PRO 1REL-1SG.SM-lack 4-money NEG-1SG.SM-can 18.LOC
 ki-tsúa nzó ni.
 INF-build 9.house NEG
 ‘I who lack money cannot build a house.’
- b. Ki-kí-kukí mvá **mu** ki-sâ kí-má ke?
 7REL-7SM-can 1.dog 18.LOC INF-do 7-what 7.PRO
 ‘What can the dog do?’

There are also other functions of the class 18 *mu*, two of which are illustrated in (58). In (58a), *mu* is used to express comparison, indicating in which aspect the two elements are compared. In (58b), the “*mu*+infinitive’ structure is used as a common way of expressing progressive in this language (Bastin 1989; De Kind et al. 2015).

- (58) a. Ma-yéle má-lur-á mphiini **mu** mpáani.
 6-intelligence 6SM-surpass-FV 9.magic.force 18.LOC 9.power
 ‘Intelligence is more powerful than magic force.’

- b. Mwáana wína **mu** **kâ-lil-á** tíí
 1.child only 18.LOC 1SM.IMPf-cry-FV until
 ku-kí-yi-bvúruk-a ngúku.
 17REL-7SM-IMPf-return-FV 1.mother
 ‘The child kept crying until mother returned.’

Syntactically, *ya-*, *ku-* and *mu-* can all occur as verbal prefixes, either being attached to a complementiser or appearing as a relative pronoun, as seen from examples in (55) above and in (59). However, a prepositional phrase headed by these three morphemes can never trigger subject marking on the verb, which is also shown in (59) where the phrase headed by *mu* is only indexed on the relative verb but not on the matrix verb. In the next subsection, I will show that these three locative morphemes also show concord in demonstratives, thus on the one hand they behave like prepositions, while in some situation they can also function as the prefixes of locative DPs.

- (59) **Mu** Marí **mu-káá-túr-í** ndzií **kí-bi**.
 18.LOC 1.Mary 18REL-1SM.PST-steal-PST 9.money 7SM-bad
 ‘That Mary stole the money is bad.’

In this subsection I have provided an overview of the noun class system in Kukuya, also having addressed some relevant questions. Next I introduce the nominal derivation rules.

2.3.1.2 Nominal derivation

In this part I describe the process of different types of nominal derivation in Kukuya. First I introduce derivation between nominals and verb-to-noun derivation, then I describe the formation of diminutives and compound nouns.

The singular-plural opposition is not the only correlation among different noun classes. One noun stem can be selected by different class prefixes to derive different semantic fields. The examples in (60) show that the gender

pair class 3/4 that includes plants and trees can derive the names of their fruits by replacing the class prefixes by those of class 5/10.

- (60) a. *mu-sakí* (cl.3) 'safou tree'
 li-sakí (cl.5) 'safou fruit'
- b. *mi-bónomo* (cl.4) 'aubergine trees'
 li-bónomo (cl.5) 'aubergine fruit'
 mbónomo (cl.10) 'aubergine fruits'

In (61) the noun roots *-wéle* and *-boli* may encode qualitative meanings 'poor' and 'lazy'. When they are selected by class 1/2 prefixes, the derived nouns refer to the people with particular characters; when they take the class 14 prefix, the resulted nouns express the abstract concepts of these qualities.

- (61) a. *mu-wéle* (cl.1) 'poor person'
 bu-wéle (cl.14) 'poverty'
- b. *ba-boli* (cl.2) 'lazy persons'
 bu-boli (cl.14) 'laziness'

As in many other Bantu languages, nominals in Kukuya can also be derived from verbs. In (62) and many examples above, we see that nouns of different classes are derived from verbs by class prefix shifting. This verb-to-noun derivation is also encoded by shifting the final vowel *-a* to a suffix *-i*, which according to the vowel combination restrictions (see Table 2.7 above) can be combined with all the stem-initial vowels. Here I treat the infinitives as derived in the way that the verb roots being nominalised by the class 7 prefix *ki-*.

- (62) a. *mu-tsúli* (cl.1) 'blacksmith'
 ki-tsúla (cl.7) 'to forge'
- b. *mu-ŋaani* (cl.3) 'threat'
 ki-ŋaana (cl.7) 'to threaten'
- c. *ma-kúia* (cl.6) 'slander words'
 ki-kúia (cl.7) 'to malign'

- d. *mpópi* (cl.9) 'judgement'
ki-wópo (cl.7) 'to judge'
- e. *lí-búru* (cl.5) 'seed, fruit'
ki-búra (cl.7) 'to engender, to bear'

Instead of maintaining diminutive classes such as class 12 and 19 of Proto-Bantu, Kukuya employs an alternative interesting way of deriving diminutive nominals. The diminutives all take the class 7/8 prefixes *ki-* and *bi-* which substitute the prefixes of the original noun. The form of a diminutive always surfaces with partial doubling of the noun stem, in which the first syllable of the stem is reduplicated (see Gibson et al. 2017; Goes and Bostoen 2021 on similar ways of diminutive formation in the Kikongo Language Cluster). Some examples are given below:

- (63) *kii-ku-ko* 'small banana'
- bii-bú-baana* 'small children'
- ki-bí-bilí* 'small lola tree'
- ki-sí-saka* 'small gourd'
- ki-mbu-mbaa* 'little fire'
- ki-nzú-nzo* 'small house'
- ki-yǔ-yǔ* 'small peanut'

When the class 7/8 prefixes are attached to a noun with no separable prefix or the |N|, they are realised as long form as in the first words in (63). We also observe that the vowel in the reduplicated first syllable of the stem is in some cases modified, for example the diminutive form of *mbaa* 'fire' is *ki-mbu-mbaa*. The tone pattern of the reduplicated syllable and the original stem also co-vary in diminutives, as the diminutive form of *báana* 'children' is realised as *bii-bú-baana*. The relevant vowel modification rules are illustrated in (64), while the tone change rules are summarised by Paulian (1975: 193) in Table 2.15, which I also checked and confirmed in the fieldwork.

- (64) *Vowel change in the reduplicated syllable of diminutives*
- i, e > i**
- u, o > u**
- a > i or u**

Original tone	Example	Derived tone	Example
L	<i>pama</i> 'animal'	L-L	<i>kii-pi-pama</i> 'small animal'
LH	<i>bilí</i> 'kola tree'	LH-LH	<i>kii-bí-bilí</i> 'small kola tree'
H	<i>li-bái</i> 'latte'	H-H	<i>ki-bú-bái</i> 'a little latte'
HL	<i>mu-lúmi</i> 'husband'	HL-L	<i>ki-lù-lumi</i> 'small husband'
H	<i>nzó</i> 'house'	HL-L	<i>ki-nzù-nzo</i> 'small house'

Table 2.16: Tone patterns in diminutives

We can see from the table that the tone on the reduplicated syllable is just copied from the tone pattern of the whole original stem. The stem just keeps its original tone pattern in the diminutive form, except when the stem is HL or when it is monosyllabic and H, in both cases the stem is realised as L in the diminutive form.

As is common for diminutive derivation, on the semantic level the derived diminutives do not always have the reading of the “smaller” counterparts of the original nominals but can have various interpretation, which includes derision as well as slight appreciation, as shown in (65).

- (65) a. *mpúru* 'anger'
ki-mpû-mpuru 'ineffective anger'
- b. *mu-ngwa* 'salt'
ki-ngu-ngwa 'a little salt'
- c. *mbaa* 'fire'
ki-mbu-mbaa 'fire that burns well'
- d. *nkwírí* 'medicine'
ki-nkwí-nkwírí 'effective medicine'
- e. *ki-se* 'good flavour'
ki-si-se 'very good flavour'
- f. *mwáana* 'child'
kii-mû-mwaana 'insupportable child'

In many Teke varieties including Kukuya, there are also compound nouns that are formed by the combination of independent words, which can be

written together with a hyphen as one word. Most of the compound nouns are animals, plant names and birds. Common ways of deriving a compound noun are the repetition of noun stem or the connection of two stems by a connective marker *-a-*. Some examples are given in (66).

- (66) *ngu-balaka* ‘uncle’ *nzala-nsa* ‘pride’
 musali-ntsie ‘farmer’ *ki-mpala-mpala* ‘lizard’
 kolo-kolo ‘sorry’ *poro-poro* ‘porridge’
 dongo-dongo ‘ocra’ *nzala-ngonomo* ‘black centipede’
 bhii-a-we ‘we and you’ *ntsie-a-tswii* ‘ear-drum’
 nga-nzó ‘owner of a/the house’ *ba-ŋómo-a-ŋómo* ‘stars’

Having established the noun class system of the language and the composition of nominals, in the next subsection I extend the description to noun phrases.

2.3.2 Noun phrase

In this subsection, I provide a description on the structure of noun phrases. I present all the elements that can modify the head noun within a noun phrase, which include connective markers, demonstratives, adjectives, quantifiers, numerals and interrogatives. I show the functions of each adnominal modifier respectively and their concord patterns with the head noun. I also introduce the use of independent pronouns.

2.3.2.1 Connective

The connective (also referred to as “associative”) marker is a linker that syntactically and semantically connects two nominals in which the former functions as the head and the latter functions as a modifier. The connective element always agrees with the head noun in noun class. Some simple examples of connective constructions are given in (67). When the connect-

ive element bears H tone, this H tone always spreads onto the following nominal prefix.

- (67) a. ki-mpfúúmú kíí mú-kokó
 7-chief 7.CONN 1-king
 ‘power of a/the king’
- b. mi-ɲwa míí yúlu
 4-mouth 4.CONN 5.nose
 ‘nostrils’
- c. li-bakí líí baarí
 5-network 5.CONN 2.people
 ‘connection of people’
- d. bi-bu bíí mfúlu
 8-cover 8.CONN 1.tortoise
 ‘shells of tortoise’
- e. mbona aa ɲginíma
 1.tire 1.CONN 9.bicycle
 ‘tire of a/the bicycle’

Many Teke languages make use of two sets of two connective markers, which is also the case in Kukuya. I follow the grammar guide by SIL Congo on Teke-Eboo (B74) (2016, 2020) to distinguish the two sets as “basic form” and “emphatic form”. The concord morphology of the two sets of connective markers with each nouns class is shown in Table 2.16. There are no connective markers for locative classes. It is interesting to see that formally the basic set corresponds to the shape of subject markers of each class, while the emphatic set is identical to relative markers (see section 2.4.2). Note that all the connective markers bear H tone except for class 1 and 9, and class 1 and 3, class 9 and 10 connective markers are distinguished tonally.

In most cases the basic and emphatic forms of connective markers are interchangeable, but there are clear cases in which the emphatic form is required when the connective construction is focused. In (68) we see that in a possessive construction, the basic form cannot be used when the possessor is the inherently focal pronoun *ná* ‘who’, instead the emphatic form

Class	Basic	Emphatic
1	aa	wuu
2	báá	ba báá
3	áá	wuu
4	míí	mi míí
5	líí	li líí
6	máá	ma máá
7	kíí	ki kíí
8	bíí	bi bíí
9	aa	yii
10	áá	yii
14	búú	ba búú

Table 2.17: The connective markers in Kukuya

should be used. Sometimes the speakers pronounce the class 1 *wuu* as *wu-á* or *wá*, which indicates that this emphatic use in fact involves relative marking (see section 2.5.4). In example (69), the class 9 emphatic form *yii* is used instead of the basic form *aa*, since the connective phrases are focused in the alternative question.

- (68) a. mu-lúmi aa me
 1-husband 1.CONN 1SG.PRO
 ‘my husband’
- b. *mu-lúmi aa ná
 1-husband 1.CONN 1.who
Int: ‘whose husband’
- c. mu-lúmi wuu ná
 1-husband 1.CONN.EMP 1.who
 ‘whose husband’
- d. kii-ndomo ki-kíí ná
 7-sheep 7-7.CONN.EMP 1.who
 ‘whose sheep’

e. mfulá yii bu-ní
 9.calabash 9.CONN.EMP 14-which
 ‘which (type) of calabash’

(69) Ndé nzó yii taará wo yii mu-káli
 1.PRO 9.house 9.CONN.EMP 1.father or 9.CONN.EMP 1-wife
 káá-tsú-i?
 1SM.PST-build-PST
 ‘Did he build a house for father or for the wife?’

Paulian (1997: 219) noted that the basic connective may encode a closer relation of the two nominals than the emphatic form, as shown in example (70), in which the use of the two sets of connective markers triggers different interpretation. More research should be conducted on the occurring conditions of emphatic connective markers.

(70) a. mbǒ aa ntaali
 9.ring 9.CONN 1.snake
 ‘ring of a/the snake’
 b. mbǒ yii ntaali
 9.ring 9.CONN.EMP 1.snake
 ‘trace of a/the snake’

When the connective marker takes the shape VV (class 1/3/9/10), it is prosodically merged with the preceding head noun in casual speech, as shown in the examples below. In (71), the VV connective marker is realised by lengthening the final vowel of the head noun (citation forms in brackets), whose tone is also influenced by that of the connective element.

(71) a. ku nzóo we
 17.LOC 9.house.CONN 2SG.PRO
 ‘in your house’ (nzó ‘house’)

(73). In (73a) the noun *mu-kái* ‘woman’ adds specifying information to the head *ki-puní* ‘old person’; in (73b) the head noun functions as the unit of measuring the quantity of the other referent; in (73c) the modifying noun *mu-tí* ‘tree’ brings to light the material of which the bowl is made; the connective in (73d) introduces a purpose or usage; and in (73e) the “connective+numeral” construction is the equivalent of an ordinal number.

- (73) a. *ki-puní kíí mú-kái*
 7-old.person 7.CONN 1-woman
 ‘old lady’
- b. *beke líí nselé*
 5.bundle 5.CONN 10.straw
 ‘a bundle of straw’
- c. *nkhíe yii mu-tí*
 9.bowl 9.CONN 3-tree
 ‘wooden bowl’
- d. *nkwî aa mbaá*
 9.firewood 9.CONN 3.fire
 ‘firewood (for fire)’
- e. *tsúku líí mwóolo*
 5.day 5.CONN two
 ‘the second day’

The connective maker is also frequently seen to be used by itself, without an overt head noun. In (74) the class 10 connective construction *yii me* has the reading of “my words”, in which the head noun *ndaka* (cl.10) “words” is elided; and in (75) the class 7 connective phrase *kíí ndé* is another common expression which means “someone’s own home”.

- (74) *Kalí we yúk-i [yii me], kéné we*
 COND 2SG.PRO listen-PST 9.CONN 1SG.PRO EMP 2SG.PRO
ka-wéén-á bí-lokó bvíí we ni.
 NEG-lack-FV 8-thing 8.CONN 2SG.PRO NEG
 ‘If you listened to my words, you would not lose your properties.’

- (75) Mu-pfúuru [kíí ndé] káá-siir-i ŋáa.
 1-uncle 7.CONN 1.PRO 1SM.PST-stay-PST 16.DEM.II
 ‘The uncle stayed here at his home.’

2.3.2.2 Demonstratives

Kukuya employs three series of demonstrative markers which can be used to distinguish spatial and temporal deixis according to the point of reference and the direct availability of the speaker, which are summarised in Table 2.18. The first set corresponds to English “this”, the second type is a medial demonstrative which is translational equivalent to English “that”, and the third type can be translated as “that (over there)” in English. The demonstrative markers always follow the noun that they modify in a noun phrase. In glossing, I use I, II and III to distinguish the three.

Class	I	II	III
1	wu	wúa	wuúra
2	ba	báa	baára
3	wu	wúa	wuúra
4	mi	mía	miúra
5	li	lía	liúra
6	ma	máa	maára
7	ki	kía	kiúra
8	bvi	bvía	bviúra
9	yi	yía	yíúra
10	yi	yía	yíúra
14	ba	búa/báa	baára
16	ŋa	ŋáa	ŋaára
17	ku	kúa	kuúra
18	mu	múa	muúra

Table 2.18: The demonstrative markers in Kukuya

The first type of demonstratives is used to denote a referent that is spatially

or temporally proximal to the speakers' surroundings or that is the current topic of conversation, referring to some entity that is given in the discourse. As shown in the table, this type has the shortest form of the three and always bear a L tone. Some sentences in which this type of demonstrative is used are shown in (76) and (77).

- (76) Mwáana wu ná áá-yii ndé?
 1.child 1.DEM.I 1.who 1SM.PST-bring.PST 1.PRO
 'Who brought this child?'

- (77) Ma-dzá ma ka-má-li tsítsie ni.
 6-water 6.DEM.I NEG-6SM-COP clean NEG
 'This water is not clean.'

The second type is used to refer to an entity which is not necessarily remote from the speakers, but is not directly accessible in the discourse context. This type of demonstratives can also have discourse-linking anaphoric function referring to the referent that is already mentioned in the context. Two examples on its anaphoric use are given in (78) and (79).

- (78) (*Will father send the parcel to Brazzaville?*)
 Ambú, taará mu-dzilá wúa Ngo káa-wéek-e.
 no 1.father 3-parcel 3.DEM.II Ngo 1SM.FUT-send-FV
 'No, father will send the parcel to NGO.'

- (79) Me ki-n-wí me ma-táli ma-bííbí, kéne
 1SG.PRO COND-1SG.SM-give.PST 2SG.PRO 6-sunshine 6-few EMP
 we tsúuk-i ndáká yía.
 2SG.PRO solve-SBJV 9.problem 9.DEM.II
 'If I had given you much time, you should have solved that problem.'

This type II demonstratives can also be used pronominally, as in example (80), in which the class 14 demonstrative pronoun refers to the fact/manner that s/he has eaten a plate of rice.

- (80) Ndé á-maá-dzá saaní líí lóoso, ndé kukí mu ki-dzá
 1.PRO 1SM.PERF-eat 5.plate 5.CONN 5.rice 1.PRO can 18.LOC INF-eat
 má-lúr-a búá.
 6REL-surpass-FV 14.DEM.II
 ‘S/He has already eaten a plate of rice, and s/he can eat more than that.’

The third type of demonstratives is least commonly seen and refers to entities or locations that are distant from the speakers. The most frequently attested one is the class 16 demonstrative *kuúra*, whose final vowel can be lengthened if the speaker would like to emphasise the far distance. Two examples on this type of distal demonstratives are given in (81) and (82).

- (81) Yá baarí **kuúra** ba-bá-kâ-wo-ó mvá mu
 with 2.people 17.DEM.III 2REL-2SM-IMPF-help-FV 1.dog 18.LOC
 ki-búr-a.
 INF-give.birth-FV
 ‘There are people over there who are helping the dog bearing babies.’
- (82) Báana **baára** mu ki-má bá-kâ-bvúnum-a?
 2.children 2.DEM.III 18.LOC 7-what 2SM-IMPF-gather-FV
 ‘Why are the children over there gathering together?’

2.3.2.3 Adjectives

There is only a limited set of adnominal modifiers that can function as true adjectives in Kukuya, which can be further divided into two types: the invariant adjectives and the ones that can show concord with the head nouns. In (83) I list all the invariant adjectives in Kukuya that I have in my corpus, whose forms never change and never show concord with the head noun, for example in *mu-ŋwa bvaí* ‘open mouth’ and *baarí mpima* ‘crowded people’.

- (83) *bvaí* ‘open’ *ngoli* ‘free’
mpima ‘tight’ *mwe* ‘alive’
nziimi ‘many, much’ *paála* ‘flat’
pi ‘calm, quiet’ *tsítsie* ‘clear’

The other closed set of adjectives follow the head nouns they modify and show concord with them in noun class. These agreeing adjectives and some examples are shown in (84), while the concord prefixes of each noun class are summarised in Table 2.19. These agreeing adjectives can derive their corresponding abstract notions by taking the class 14 prefix *bu-*, for example *bu-níni* ‘bigness’ and *bu-bvé* ‘beauty’.

- (84) a. *-bvé* ‘good’ *-bí* ‘bad’
-níni ‘big’ *-la* ‘long’
-pfii ‘short’ *-kima* ‘other’
-ɲaríki ‘new’ *-tsiina* ‘whole’
-bíibi ‘little, few’ *-khiele* ‘small’
-bui ‘yellow, immature’ *-kulu* ‘old’
-mfe ‘cold’ *-ke* ‘several’
-lílakâ ‘numerous’ *-bíi* ‘green, fresh’
- b. *tsúku li-tsiina* ‘whole day’
nzũ n-kima ‘(the) other peanuts’
mu-siá wu-pfii ‘short rope’
nzalí wu-bíibi ‘small river’
má-lúa ma-níni ‘big disease’

From the table, we see that the adnominal concord prefixes are formally identical to the proximal demonstrative markers, except for the class 9/10, whose concord prefix is always a nasal which assimilates with the following consonant in the place of articulation. Though the concord prefixes usually occur as a single morpheme as in the table, it is also common to see a complex concord form, in which the head noun is indexed twice on the modifier and formally looks like a subject relative (see section 2.5.3), as shown in the examples in (85). This was elsewhere described as the CONN NPrefix-adjective construction for Kikongo languages (Baka 2000; Bouka 1998).

Class	Concord	Class	Concord
1	wu-/mu-	2	ba-
3	wu-/mu-	4	mi-
5	li-	6	ma-
7	ki-	8	bi-
9	N-	10	N-
14	bu-		

Table 2.19: The adnominal concord prefixes of adjectives in Kukuya

- (85) *mbuurú wu-mu-bvé* 'a/the good person'
nzó yi-m-bvé 'a/the good house'
ki-dzulibí ki-kí-níni 'a/the big door'
búku li-lí-kúlú 'a/the old book'

Beyond the adjectives in the two limited sets in (84), descriptive meanings can be expressed by various other means in Kukuya. In (86) the translational adjectives are expressed by using a subject relative clause.

- (86) a. *mvúla wǔ-ya*
 1.year 1REL-come
 'next year'
- b. *mwáana wǔ-wéén-á má-yéle*
 1.child 1REL-lack-FV 6-intelligence
 'stupid child'
- c. *ki-ko ki-kí-baak-i*
 7-loincloth 7REL-7SM.PST-be.torn-PST
 'torn clothes'
- d. *ki-baká ki-kí ya má-mpirí nziimi*
 7-wall 7REL-7SM with 6-colour many
 'colourful wall'

Another common strategy to express specifying meaning is to use a relative marker plus the preposition *yǎ* 'with' whose underlying tone pattern is LH

but can donate the H tone to the following prefix. For example, the concept “a deep river” can be literally expressed as “a river that has/is with depth” and “a tall person” as “a person who has/is with height”. Some examples are given in (87).

- (87) a. *ɲama wũ yǎ mpúru*
 1.animal IREL with 9.anger
 ‘dangerous animal’
- b. *nzáli wũ yǎ dzika*
 1.river IREL with 9.depth
 ‘a deep river’
- c. *mbuka yii ya kí-dziimi*
 9.place 9REL with 9.shade
 ‘a shady place’
- d. *mbuurú wũ ya mú-téle ndziimi*
 1.person IREL with 3-height much
 ‘a very tall person’
- e. *tshoo li-lí yǎ ɲkali*
 5.food IREL-5AGR with 9.bitterness
 ‘bitter food’

To express comparison on a certain quality between two entities, the word *ki-lura* ‘to surpass’ is frequently used, and the class 18 preposition *mu* brings into the aspect that A surpasses B, as illustrated in examples (88) and (89).

- (88) *Ngo lur-í kii-mbúli mu ntsiini.*
 1.leopard surpass-PST 7-lion 18.LOC 9.speed
 ‘A leopard (runs) faster than a lion.’
- (89) *Mu-kúlu lur-í me mu mvúla tíri.*
 1-elder surpass-PST 1SG.PRO 18.LOC 1.year three
 ‘My brother is three years older than me.’

To express superlative, a similar construction is used, except that it does not

need to show comparison with another element. In (90), “the longest one” is expressed as “the one that surpasses the length”.

- (90) Me ya má-téme mpúomo, me m-wî
ISG.PRO with 6-hoe eight ISG.PRO ISG.SM-give.PST
li-lí-lur-i **bu-la** ndúku.
5REL-5SM.PST-surpass-PST 14-length 1.friend
'I have eight hoes, I gave the longest one to a friend.'

2.3.2.4 Quantifiers

Kukuya does not employ a large amount of quantifiers. In this part I introduce the form and use of the universal quantifier “all” and the expressions of “each/every”. Quantifiers are a subset of adjectives in Kukuya, which refer to certain amounts of entities, and we have already seen the expressions of “many/much” and “few/little” above in (84) as adjectives, I don’t discuss them further here.

The universal quantifier “all” always shows concord with the nominal that it modifies, resulting an amalgamated element whose concord prefix is not separable, as summarised in Table 2.20. According to Paulian’s (1975: 203) analysis, this quantifier has the underlying root as -|Yóî| in which the abstract morphophoneme |Y| is represented by the constricted articulation **Ch** (see section 2.1.1.3) and the initial consonant is determined by the noun class prefix. Vowel coalescence also happens when the amalgamated word is formed. I don’t discuss the formation of the words in the paradigm in detail here. Some examples of the universal quantifier are given in (91) and (92). To emphasise the universal reading, this quantifier can also be formally reduplicated, as in (92). As seen from the table, this quantifier can modify not only the countable nouns in the plural classes, but it can also modify singular nouns, meaning “the whole”.

- (91) Mpéle wu-kíi-yáab-i baarí **bhoî** áá-puum-i
 1.priest IREL-7SM.PST-know-PST 2.people 2.all 1SM.PST-get.lose-PST
 mfúúlá ku ntsá mu-kééné.
 9.way 17.LOC inside 3-desert
 ‘The pastor that all the people knew lost his way in the desert.’
- (92) Wu-kâ-bvúúr-á bí-lokó **bvheîbvheî** ngampu
 IREL-IMPF-return-FV 8-thing 8.all 1.God
 ndé-me-nkúlu.
 1.PRO-1SG.PRO-oneself
 ‘The one who reimburses all the things is the God himself.’

Class	‘All’	Class	‘All’
1	whoî	2	bhoî
3	whoî	4	mheî
5	lhoî	6	mhoî
7	kheî	8	bvheî
9	yhoî	10	yhoî
14	bhoî		

Table 2.20: The concord pattern of ‘all’

To express “every/each”, the quantifier *ná*, which has the same shape as the *wh*-word “who” (see the description on interrogatives shortly) is used and is placed before the noun it modifies. See two examples in (93) and (94).

- (93) Mbuurú wu-kâ-tsúk-a maá-bí **ná** tsúku lii-nkíla
 1.person IREL-IMPF-talk-FV 6REL-bad every 5.day 5-tongue
 líi-tsírik-i.
 5SM-cut-PST
 ‘The man who speaks bad words everyday was cut the tongue.’

- (94) Taará ka-káá-ték-i ná ntsúí ni.
 1.father NEG-ISM.PST-sell-PST every 1.fish NEG
 ‘Father did not sell each fish.’

This quantifier *ná* has an invariable shape and never shows concord with the head noun. It can also follow the head noun if followed by a class 1 pronoun *ndé* which may have an emphatic function, as shown in (95).

- (95) Ntaba ná ndé ŋa ntsá bu-lá ba baa-ték-e.
 1.goat every 1.PRO 16.LOC inside 14-village 14.DEM.I 2SM.FUT-sell-FV
 ‘Every goat in this village will be sold.’

“Other” or “another” is expressed by the adjectival quantifier *-kima* which shows concord with the head noun. As shown in (96), the modified noun *balaka* ‘male person’ does not form a closed set with the precedent subject.

- (96) Mbuurú wu-kí-yéni mu-káli yă balaka mu-kima mí-kíí
 1.person IREL-7SM-go.away 1-wife with 1.man 1-other 4-curse
 kâ-ló ná tsúku.
 1SM.IMPF-curse every 5.day
 ‘The person whose wife went away with another man curses every-day.’

When expressing “the other” element that is contrasted with a different entity in the same set, the quantifier *-mó* is used. In the context of example (97), two chicken are involved: one was caught, and the other survived.

- (97) Ntsúú wu-mó wu-báá-bí siiba ndíri, ká we
 1.chick 1-other IREL-2SM.PST-refuse catch 1.COMP.say EMP 2SG.PRO
 ŋa yi-seb-e.
 16.PRO IMPF-laugh-FV
 ‘The other chick which was not caught said, ‘it was you who were laughing there!’

2.3.2.5 Numerals

The citation forms of cardinal numerals in Kukuya are listed in Table 2.21, and an example of the formation of a complex number is given in (98).

Numeral	Translation
nguumó	one
mwóolo	two
tíri	three
na	four
taaní	five
séneme	six
nsaama	seven
mpúomo	eight
wá	nine
khúumi	ten
ma-khúumi mwóolo	twenty
ma-khúumi ma-taaní	fifty
ma-khúumi wá yă nsaama	ninety-seven
nkámá	hundred
pfuuná	thousand

Table 2.21: The cardinal numerals in Kukuya

- (98) pfuuná yă nkámá wá ya má-khúumi wá yă nsaama
 thousand and hundred nine and 6-ten nine and seven
 '1997'

Only the numbers from 1 to 6 can show concord with the head noun by taking the same set of concord prefixes as adjectives (see Table 2.18), for 7 to 10 the shapes of numerals are invariant, but the concord pattern maintains for the 1 to 6 single digits in a number bigger than 10, as shown in the (99). The concord patterns of “one” and “two” are exceptional and they are displayed in Table 2.22. It remains to be investigated what the *nga-* stands for in the various forms of numeral “one”.

Class	One	Class	Two
1	nguumó	2	bhóolo
3	nguumó	4	mhéele
5	ngalimó	6	mhóolo
7	ngakimó	8	bvhéele
9	ngamó	10	yhóolo

Table 2.22: The nominal agreement pattern of ‘one’ and ‘two’

- (99) *báana ba-séneme* ‘six children’
mi-tí mi-na ‘four trees’
ma-ko nsaama ‘seven bananas’
bi-lóko ma-khúumi mhóolo yă bvhéele ‘twenty two objects’

Ordinal numerals consist of a connective marker and a cardinal number that displays concord with the class of the *plural* counterpart of the head noun, as shown in (100).

- (100) a. tsúku líí mhóolo
5.day 5.CONN 6.two
‘the second day’
b. mu-tí wuu mí-tíri
3-tree 1.CONN.EMP 4-three
‘the third tree’

So far I have introduced all the adnominal modifiers including connectives, demonstratives, adjectives, quantifiers and numerals. According to the speakers, the order of them can be highly flexible, despite that the connective marker has to occur immediately after the head noun or lastly, but not in between, as shown in (101).

- (101) a. *m̥iibi áá-túr-i ma-téme máá me ma-tíri*
 1.thief 1SM.PST-steal-PST 6-hoe 6.CONN 1SG.PRO 6-three
ma-ɲwa ma-níni.
 6-red 6-big
 ‘The thief stole my three big red hoes.’
- b. *m̥iibi áá-túr-i ma-téme ma-tíri ma-ɲwa ma-níni*
 1.thief 1SM.PST-steal-PST 6-hoe 6-three 6-red 6-big
máá me.
 6.CONN 1SG.PRO
 ‘The thief stole my three big red hoes.’
- c. **m̥iibi áá-túr-i ma-téme ma-tíri máá me*
 1.thief 1SM.PST-steal-PST 6-hoe 6-three 6.CONN 1SG.PRO
ma-ɲwa ma-níni.
 6-red 6-big
Int: ‘The thief stole my three big red hoes.’

2.3.2.6 Interrogatives

Next I present the interrogative words in Kukuya, which are listed in Table 2.22.

Interrogatives	English counterpart
<i>ná</i>	who
<i>(kí)-má</i>	what
<i>ku-ní</i>	where
<i>bu-ní</i>	how
<i>-ní</i>	which
<i>munkí</i>	when
<i>mu ki-má</i>	why
<i>kwê</i>	how many/much

Table 2.23: The interrogative words in Kukuya

To ease the presentation, in Table 2.23 I match the Kukuya interrogatives

with English *wh*-words. In fact, many interrogatives words have the same root but take different concord prefixes, which determines their interpretation. For example *ku-ní* ‘where’ and *bu-ní* ‘how’ are composed of the class 17 and class 14 concord markers *ku-* and *bu-* along with the interrogative morpheme *-ní* ‘which’, thus *ku-ní* and *bu-ní* literally mean ‘which place’ (where) and ‘in which manner’ (how).

Most commonly, the interrogative words are linearly placed immediately **before** the verb in a content question, which corresponds to a dedicated focus position in the Kukuya syntax, since *wh*-words are inherently focused. See more description and discussion on this in chapter 3, which are not the main aim of this section.

Next I present the forms and uses of the interrogative words in turn.

ná ‘who’

When asking about a person (or sometimes an animal), the interrogative word *ná* is used. This *wh*-word itself is in class 1 and has a plural counterpart *baa-ná* in class 2. Some examples on its occurrences are provided in (102)-(105). The sentence in (102) is an object question, while the embedded clause in (103) is a pseudo-cleft questioning on the subject.

- (102) We **ná** á-mún-i ku mu-súru?
 2SG.PRO 1.who 2SG.SM.PST-see-PST 17.LOC 3-forest
 ‘Who did you see in the forest?’

- (103) Ndé á-fúul-i me [ba-báá-dzwí nǵíibi **baa-na**].
 1.PRO 1SM-ask-PST 1SG.PRO 2REL-2SM.PST-kill.PST 1.thief 2-who
 ‘S/He asked me who (pl.) killed the thief.’

There is often a class 1 pronoun *ndé* immediately following *ná* when questioning an object, whose exact function is not yet clear to me. This is reminiscent of the quantifier *ná* ‘every/each’ which also co-occurs with a class 1 pronoun, as introduced above in (95). It is interesting to investigate whether

the quantifier *ná* and the interrogative *ná* are related or they are underlyingly the same.

- (104) Nkaaká ma-désu ná ndé káá-wî?
 1.grandmother 6-bean 1.who 1.PRO 1SM.PST-give.PST
 ‘To whom did grandmother give the beans?’

- (105) Ndé ki-bhiimá kíí ngúku ya ná ndé
 1.PRO 7-corpse 7.CONN 1.mother with 1.who 1.PRO
 kâ-dziik-a?
 1SM.FUT-bury-FV
 ‘With whom will s/he will bury the corpse of mother?’

“Whose” is expressed by an emphatic connective marker and *ná*, see examples in (68) above.

kí-má ‘what’ and *mu ki-má* ‘why’

The interrogative word for ‘what’ in Kukuya most commonly appears as the class 7 *ki-má*, which is used to question objects, animals as well as general events. Two examples are given in (106) and (107).

- (106) We kí-má kâ-tsuom-ó ku ntsá mu-tíma áá
 2SG.PRO 7-what IMPF-think-FV 17.LOC inside 3-heart 3.CONN
 we?
 2SG.PRO
 ‘What are you thinking in your heart?’

- (107) Taará ku ndúku aa ndé kí-má káá-fúum-i?
 1.father 17.LOC 1.friend 1.CONN 1.PRO 7-what 1SM.PST-buy-PST
 ‘What did father buy for his friend?’

Similar to the class 1 pronoun *ndé* that is placed after *ná* ‘who’, there is often a class 7 pronoun *ké* occurring immediately after *ki-má*, as shown in (108).

- (108) Maamá kí-má ké kâ-dzií kí-yáab-a?
 1.mother 7-what 7.PRO 1SM.IMPf-like INF-know-FV
 ‘What does mother want to know?’

In fact, the interrogative *ki-má* consists of a root *-má* plus a class 7 prefix, and this root can *modify* nominals of various noun classes and shows concord with them, an agreeing pronoun also appears after this interrogative modifier. However, only the class 7 *ki-má* and its plural counterpart in class 8 (sometimes also class 6, depending on the content it refers to) are most commonly used as an interrogative *pronoun*. For other noun classes, when they are modified by *-má*, usually have the reading of ‘what kind/type’, for example *kíri lí-má ló* (cl.5) means ‘what kind of chair’, and *mu-ŋwa mú-má ndé* (cl.3) means ‘what language’. Some examples of its pronominal use are given in (109) and (110), and the concord patterns are illustrated in Table 2.23. It should be noted that the concordial pronouns only occur after the interrogatives *ná* ‘who’ and *-má* ‘what’, but never with other *wh*-words or focused elements in a sentence.

- (109) Bó má-má mó bá-kâ-tsúk-a?
 2.PRO 6-what 6.PRO 2SM-IMPf-talk-FV
 ‘What are they talking about?’

- (110) Báana bí-má bvé báá-kil-a mu mbára bí-baka?
 2.children 8-what 8.PRO 2SM.PST-write-PST 18.LOC against 8-wall
 ‘What did the children write on the wall?’

When asking about reasons, ‘why’ is expressed by a combination of the class 18 locative preposition *mu*, which is often used to introduce reason, and the interrogative word *ki-má*, so *mu ki-má* ‘why’ in Kukuya is literally ‘for what’. Some examples are given below. We may also notice that the tone on the prefix of (*mu*) *ki-má* are realised differently when it is used as ‘what’ and ‘why’, I refer the readers to chapter 4 on the tone of the preverbal focused elements.

Class	'What'	Class	'What'
1	wu-má ndé	2	ba-má bó
3	wu-má ndé	4	mi-má njé
5	li-má ló	6	ma-má mó
7	ki-má ké	8	bi-má bvé
9	yi-má yó	10	yi-má yó
14	bu-má bó		

Table 2.24: The concord patterns of 'what' in Kukuya

- (111) Bá-kái **mu** **ki-má** bá-li yă buokó?
 2-woman 18.LOC 7-what 2SM-be with 14.fear
 'Why are the women frightened?'
- (112) **Mu** **ki-má** ké kí-li yă nkelé ku mbali?
 18.LOC 7-what 7.PRO 7SM-be with 9.noise 17.LOC 9.outside
 'Why is there noise outside?'

ku-ní 'where'

When asking about a place, the interrogative word *ku-ní* is always used, which is composed of the locative class 17 prefix *ku-* and the interrogative root *-ní* 'which', so literally it means 'which place'. Three examples on the use of *ku-ní* are shown in (113)-(115). Note that in the matrix clause, the interpretation of *ku-ní* is always associated with the place of the event related to the verb that it immediately precedes, so in (114) *ku-ní* is used to inquiring after the place where the addressee heard the news, rather than where the man under discussion died. In example (115), *ku-ní* functions as the head NP of a relative clause (see section 2.5.2) that expresses an embedded question.

- (113) Joní mhiilí **ku-ní** mí-buok-i?
 1.John 4.leg 17-which 4SM.PST-hurt-PST
 'Where did John hurt the legs?'

- (114) We maa-ndáká máa ku-ní á-yúk-i
 2SG.PRO 6-problem 6.DEM.II 17-which 2SG.SM.PST-hear-PST
 bóri mbuurú áá-kwî mu mbaá?
 2.COMP 1.person 1SM.PST-die.PST 18.LOC 9.fire
 ‘Where (place of hearing) did you hear that the man died of fire?’
- (115) Me ka-n-yáab-i [ku-ní ku-m-bak-i
 1SG.PRO NEG-1SG.SM-know-PST 17-which 17REL-1SG.SM-get-PST
 me ma-dzá] ni.
 1SG.PRO 6-water NEG
 ‘I did not know where I (could) get water.’

***bu-ní* ‘how’**

The interrogative word *bu-ní* can be translated as ‘how’ and is used to ask about the manner in which an action is accomplished, or the state of a person or an entity. This *wh*-word consists of the class 14 prefix, which is commonly associated with the manner of doing something, and the root *-ní*, so literally it means ‘in which manner’. In this sense it should be grouped together with *ku-ní* as the ‘which’ interrogatives. Two examples are given in (116) and (117).

- (116) Baa-nziá bu-ká bun-í bá-dzá?
 2-foreigner 14-cassava 14-which 2SM-eat
 ‘How do the foreigners eat the cassava?’
- (117) Ndé nká bu-ní ká-siib-i?
 1.PRO 1.antelope 14-which 1SM.PST-catch-PST
 ‘How did s/he catch the antelope?’

***-ní* ‘which one’**

The interrogative root *-ní* is used for discourse-linking questions ‘which (one)’ or ‘what kind of’. It occurs after the noun that it modifies and takes the concordial prefixes. Two examples are given below.

- (118) Mu-kái **wu-ní** wũ-lak-í ndíri áá-mún-i Gilbert?
 1-woman 1-which 1REL-say-PST 1.COMP 1SM.PST-see-PST Gilbert
 ‘Which woman said that she saw Gilbert?’

- (119) Baa-mángulu ba-káá-fúúm-í bó **ba-ni**?
 2-mango 2REL-1SM.PST-buy-PST 1.PRO 2-which
 ‘What kind of mangoes did s/he buy?’

munkí ‘when’

The interrogative word *munkí* is used for inquiring general time, and has an invariant form. Two examples on its use are given below. It would be interesting to investigate the origin of this temporal interrogative word, since in some other Teke varieties ‘when’ is expressed by the word *li-ní* ‘which time’ that seems to belong to the same set of interrogatives as *ku-ní* and *bu-ní*.

- (120) Ngúku mu-kiike **munkí** káá-yók-í má-dzá?
 1.mother 1-baby when 1SM.PST-bath-PST 6-water
 ‘When did mother bath the baby?’

- (121) We ña tíí **munkí** â-kál-a?
 2SG.PRO 16.PRO until when 2SG.SM.FUT-stay-FV
 ‘Until when will you will stay here?’

kwê ‘how many’

The interrogative word *kwê* is used for asking the quantity of a entity, either countable or uncountable. Like *munkí* ‘when’, this *wh*-element also has an invariant form and always follows the noun that it modifies. See the examples below.

- (122) Mii-ndéle ma-tsúku kwê míi-sál-a ŋa ntsá
 4-foreigner 6-day how.many 4SM.FUT-work-FV 16.LOC inside
 bu-lá ba?
 14-village 14.DEM.I

‘For how many days will the foreigners work in this village?’

- (123) We mii-nkáání kwê â-kil-a?
 2SG.PRO 4-letter how.many 2SG.SM.FUT-write-FV

‘How many letters will you write?’

***ka* + interrogatives**

An intriguing fact on the interrogatives in Kukuya is when they occur in a non-preverbal position in an embedded clause, they must be preceded by a *ka* marker whose exact function is still unclear to me (I gloss it as emphatic ad hoc), as shown in the examples (124)-(127) below. It is also unknown if this *ka* can be categorised with any other homophonic *ka* markers in this language. It somehow seems that *wh*-word needs some additional licensing when it occurs in an A-position in the embedded context, which I leave for future research.

- (124) Ba-ntsúú ba-kíi-ká-í *(ka) ná bá-yiká bá-bí?
 2-chicken 2REL-7SM.PST-grill-PST EMP 1.who 2SM-IMP 2-bad

‘The chicken that who grilled is getting bad?’

- (125) Ndé áá-som-í mu ki-ná kǐ-yi-tsúk-í
 1.PRO 1SM.PST-enter-PST 18.LOC 7-moment 7REL-IMP speak-PST
 we *(ka) ma-má?
 2SG.PRO EMP 6-what

‘He entered when you were saying what?’

- (126) Mu-kái wǔ-fúum-í *(ka) ki-má ké báá-pfur-í?
 1-woman 1REL-buy-PST EMP 7-what 7.PRO 2SM.PST-cheat-PST

‘The woman who bought what was cheated?’

- (127) Me mún-i mbuurú wu-kí-tí me
 1SG.PRO 1SG.SM.see-PST 1.person 1-COMP '1SG.PRO
 ka-n-yáab-i ndé *(ka) ná ni.
 NEG-1SG.SM-know-PST 1.PRO EMP 1.who NEG
 'I saw the person that I don't know who he was.'

2.3.2.7 Person and class pronouns

The independent pronouns including speech participants and noun class pronouns in Kukuya are listed in Table 2.24. The most common function of these pronouns is to serve as the pronominal arguments of the verb, referring back to a certain person or entity that has been given in the discourse, either in the previous or the same sentence. The class pronouns are also attested after the interrogatives *ná* 'who' and *-má* 'what', which I have shown above. Here I give two examples on the resumptive use of the pronouns.

Class/Person	Pronoun	Class	Pronoun
1st SG	me	2nd SG	we
1st PL	bhii	2nd PL	bé
1	ndé	2	bó
3	ndé	4	ɲé
5	ló	6	mó
7	ké	8	bvé
9	yó	10	yó
14	bó	16	ɲa

Table 2.25: The independent pronouns in Kukuya

- (128) Me ka-n-dzií ki-yé ku kíí ngúbalaka ni
1SG.PRO NEG-1SG.SM-like INF-go 17.LOC 7.CONN 1.uncle NEG
mu bu-kítí ndé kâ-nywí-a me ma-lí ndziimi.
18.LOC 14-COMP 1.PRO IMPF-make.drink-FV 1SG.PRO 6-wine much
'I don't like going to my uncle's house because he (usually) makes
me to drink much wine.'
- (129) Mi-fémé mí-kâ-tsuom-ó bóri me njé
4-pig 4SM-IMPF-think-FV 2.COMP 1SG.PRO 4.PRO
â-n-siib-a.
FUT-1SG.SM-catch-FV
'The pigs think that I will catch them.'

To summarise, in this section I have introduced the nominal morphology, including noun classes and noun phrases. We see that adnominal modifiers show concord relations with the nominals in noun classes. As is well known, noun classes in Bantu are also indexed in the verbal domain via agreement. Next I present the verbal morphology and inflectional categories in Kukuya.

2.4 Verbal morphology and TAM

In this section, I describe the verbal morphology and TAM conjugations in Kukuya. I first display the structure of the verb, showing the inventories of the morphemes that can occur in each slot on a verb complex, namely the negative and relative prefixes, subject markers, tense and aspect markers and the final vowel. Then I present the tense/aspect/mood markings on the verb.

The structure of an agglutinating verb is very commonly attested across Bantu languages, and a formal template of different slots on the verb complex was proposed by Meeussen (1967: 108), which “exhibits a clear structure with definable elements occurring in a fixed order”. A verb in Kukuya is not so “complex” as it makes use of a very restricted subset of the slots in Meeussen’s template. The scheme of a Kukuya verb is illustrated in (130). I don’t have an example that exploits all these slots, but I will show later that they can be clearly distinguished.

(130)	Pre-initial	Initial	Formative	Limitative	Root	Final
	REL, NEG	SM	ASP	other prefixes	verb	FV

As seen from (130), there are only three pre-stem slots in Kukuya which correspond to the Pre-initial, Initial and the Formative positions in Meeussen (1967: 108) (also see Nurse 2008). There is no object marking on the verb and no productive verb derivational extensions. However, we can still recognise some traces of the ancient derivational suffixes which have been lexicalised and become inseparable.

In (131) I have listed some verb pairs, between which we can observe a systematic correlation in both meaning and form. I suppose that the verb forms in the right column can reflect some remnants of the ancient derivational extensions (also see Paulian 1998). For example, the *l/k* opposition in some of the pairs may reflect the Proto-Bantu *-*ud* and *-*uk* suffixes that encode the transitive/intransitive distinction. I don’t discuss the reconstruction into detail here but just provide examples.

- | | | |
|-------|---------------------------------------|---|
| (131) | <i>ki-múna</i> 'to see' | <i>ki-múono</i> 'to meet' |
| | <i>ki-pálaba</i> 'to make crack' | <i>ki-pákaba</i> 'to crack' |
| | <i>ki-kolobo</i> 'to make calm' | <i>ki-kokobo</i> 'to calm' |
| | <i>ki-lelebe</i> 'to loosen (a rope)' | <i>ki-lekebe</i> 'to become loose' |
| | <i>ki-dziika</i> 'to bury' | <i>ki-dzuula</i> 'to get out of the ground' |
| | <i>ki-yíka</i> 'to learn' | <i>ki-yúka</i> 'to teach' |
| | <i>ki-khúula</i> 'to disconnect' | <i>ki-khúuka</i> 'to disengage' |

Some verbal meanings are expressed by using a complex VO structure, most notably with the light verb *ki-tá* 'to launch'. Several examples are given in (132).

- (132) *ki-tá bví* 'to throw'
ki-tá mí-táami 'to enjoy oneself'
ki-tá nzaamí 'to pray'
ki-tá mí-tári 'to witness'
ki-tá nzunimí 'to sweat'
ki-tá mú-dzéle 'to scream'
ki-múná má-kinima 'to suffer'
ki-múná bú-pfúru 'to feel ashamed'
ki-bvâ nsiina 'to sit down'

After some brief words on the verb roots and complex verbs, next I present the verb affixes in different slots in the template (130) in turn, following their linear order.

2.4.1 Verbal affixes

2.4.1.1 Pre-initial: negative and relative markers

The Pre-initial slot on a Kukuya verb can be filled by two kinds of elements, which are the negative prefix *ka-* and the relative pronouns, but they *cannot* co-occur simultaneously. The negative prefix *ka-* is often prefixed to the verb, preceding the subject marker, and it is always used along with a clause-final

negative particle *ni*. Two examples of negative sentences are shown in (133) and (134). It is interesting to see in (133) that the particle *ni* occurs twice, one appears immediately when the matrix clause ends, and the other in the final position of the whole sentence.

- (133) Mwáana **ka**-ká-lí ya kí-sáabi **ni** mu
 1.child NEG-ISM-be with 7-happiness NEG 18.LOC
 bu-kíí-sí-i ngúku ndé mu ntsá nzó **ni**.
 14REL-7SM.PST-leave-PST 1.mother 1.PRO 18.LOC inside 9.house NEG
 ‘The child is not happy that mother left him alone in the house.’

- (134) Me **ka**-n-kukí mu ki-sâ ndé kólokóli **ni** se
 1SG.PRO NEG-1SG.SM-can 18.LOC INF-do 1.PRO excuse NEG even.if
 káá-sak-í kólokóli ndziimá-ndziimi.
 1SM.PST-search-PST excuse much-much
 ‘I cannot forgive him even if he apologised several times.’

This negative prefix *ka-* is not always attached to the verb, but can also be placed before an NP constituent that it negates and has an independent prosodic status. In (135), the scope of negation is the preverbal subject NP, which is preceded by the negative marker *ka*. I would analyze the *ka* in (135) as a homophonous negative marker with a different morphological status, i.e., independent word instead of prefix.

- (135) Ngo **ka** mú-bhií áá-dwî **ni**, nzokó
 1.leopard NEG 1-hunter 1SM.PST-kill.PST NEG 1.elephant
 á-dwî ndé.
 1SM.PST-kill.PST 1.PRO
 ‘The leopard was not killed by the hunter, an elephant killed it.’

The Pre-initial slot can also be occupied by a relative pronoun, which always agrees with the adjacent relativised NP in noun class and is prefixed to the verb. The paradigm of relative pronominal prefixes of each noun class is given in Table 2.25. We see that the shape of relative prefixes are identical to

the proximal demonstrative marker (see Table 2.17), with the only exception for class 14, whose relative prefix is *bu-* instead of *ba-*. There are no relative prefixes for person pronouns. Since the syntax of relative constructions will be discussed in more detail in section 2.5 as well as in chapter 5, here I only provide two examples to show the use of relative prefixes.

Class	Relative	Class	Relative
1	wu-	2	ba-
3	wu-	4	mi-
5	li-	6	ma-
7	ki-	8	bi-
9	yi-	10	yi-
14	bu-	16	ŋa-
17	ku-	18	mu-

Table 2.26: The relative prefixes in Kukuya

- (136) Me ŋa m-bvî nsiina mu báana
 1SG.PRO 16.PRO 1SG.SM-fall.PST 9.ground 18.LOC 2.children
 ba-lí-maá-búr-a bhií.
 2REL-1PL.SM-PERF-give.birth-FV 1PL.PRO
 ‘I stay here for the children that we have given birth to.’

- (137) Mwáana wu-mu-kái wu-kíí-kwí ngúku á-yika
 1.child 1-1AGR-female 1REL-7SM.PST-die.PST 1.mother 1SM-IMPF
 kí-lilá kâ-líl-a.
 INF-cry 1SM-IMPF-cry-FV
 ‘The girl whose mother died is crying.’

2.4.1.2 Subject (and tense) markers

The subject in Kukuya is always indexed by a subject marker in the Initial slot of the verb, except for verbs in the infinitive, narrative, and imperative

conjugations. A list of subject markers for all noun classes and persons is given in Table 2.26. We see that all the subject markers have underlying H tone, except for the 1st/2nd person plural subject markers with underlying L tone. The 1st person singular subject marker is always realised as a nasal that assimilates with the following consonant in the place of articulation; the 2nd person singular subject marker has a null form; and the subject markers for 1st and 2nd person plural are identical. The class 1 subject marker has two variants *á-* and *ká-*, the alternation of which will be discussed in chapter 5. The class 1 subject prefix *á-* is often omitted on the verb when the subject is the class 1 pronoun *ndé* ‘s/he’, but is compensated by the lengthening of the final vowel of the latter.

Class/Person	Subject marker	Class	Subject marker
1st SG	N-	2nd SG	∅-
1st PL	li-	2nd PL	li-
1	á-/ká-	2	bá-
3	á-	4	mí-
5	lí-	6	má-
7	kí-	8	bí-
9	á-	10	á-
14	bú-		

Table 2.27: The subject markers in Kukuya

We also see from the table that compared to adnominal concord prefixes and relative prefixes, the subject markers in Kukuya seem to be more phonologically reduced, and class 9/10 subject markers do not show correlation with any of their concord markers in the nominal domain but seem to be merged with those of class 1/3.

As indicated in the glossing of many examples throughout the thesis, the subject marker is often realised as a combined subject and tense marker, in which the tense information is encoded by lengthening the vowel of the subject prefix and by grammatical tones, as shown in two additional examples in (138) and (139). We also see that though subject and tense marking can be combined, they can bear different tones. In (139) the

prefix *báa-* consists of a H tone subject marker *bá-* and a lengthened vowel with a grammatical L tone that encodes general future tense. This vowel lengthening strategy functions together with the alternation of final vowel on the verb to mark tense distinctions (see next subsection).

- (138) Ndé kí-má káá-fúúm-í ki-kíí-fi nsáa?
 1.PRO 7-what 1SM.PST-buy-PST 7REL-7SM.PST-end.PST 9.value
 ‘What did he buy that has expired?’

- (139) Munkí we kâ-tsuom-ó wurí bó báa-tô?
 when 2SG.PRO IMPF-think-FV 2SG.COMP 2.PRO 2SM.FUT-arrive
 ‘When do you think that they will arrive?’

This combined marker is also attested in some other Teke varieties, for example in Raharimanantsoa’s study (2017) on the tense marking in Teke-Eboo (B74), she labeled this marker as the “subject-tense marker” (STM). I do not gloss this combined marker as separate morphemes, since in some cases the vowel lengthening is not easy to detect and its realisation depends on the vowel quality of the subject prefix. Some exceptions are when the subject marker is the 1st person singular N-, which has a phonological requirement for it to always follow a CV- or V- shape tense/aspect prefix but never precedes it. In this case, tense is marked via a separate marker *a-* preceding the nasal prefix, as shown in (140); when the subject marker is the null 2nd person singular, this tense marker *a-* just stands by its own as in (141).

- (140) Me a-n-bólik-i Gilbert kírí.
 1SG.PRO PST-1SG.SM-break-PST 1.Gilbert 5.chair
 ‘I broke Gilbert’s chair.’

- (141) We me bu-ní á-swool-i?
 2SG.PRO 1SG.PRO I4-which PST-find-PST
 ‘How did you find me?’

The class 7 subject marker *ki-* serves as the default subject marking strategy in this language. It is used when there is no overt preverbal subject for the verb to agree with, and this default use is most notably seen on a copula, as in (142). An additional example is given in (143), in which the class 7 *ki-* functions as a default relative marking strategy, as there is no candidate in the sentence for the relative pronoun to agree with.

- (142) Ka-kí-li ya kí-lóko ki-báá-sí ni.
 NEG-7SM-COP with 7-thing 7REL-2SM.PST-do.PST NEG
 ‘There was nothing that was done.’

- (143) Ki-n-dzií me ki-dzá wína ŋama.
 7REL-1SG.SM-like 1SG.PRO INF-eat only 1.meat
 ‘What I like to eat is only meat.’

2.4.1.3 TAM markers and other verbal prefixes

Next I introduce the verbal prefixes that occur between the subject marker and the verb root, which include aspect prefixes in Meeussen’s Formative slot and some other verbal prefixes that may correspond to the Limitative slot. When any of these aspect prefixes are present on the verb, the final vowel of the verb always has the default form (see below on the final vowel).

Prefixes *kâ-*, *maá-*, and *yi-*

The prefix *kâ-* is the most commonly attested aspect prefix, which often expresses habitual meaning that specifies an action as taking place habitually or regularly, or it is used in a statement which is believed to be universally true. I analyze this *kâ-* marker with the overarching aspectual meaning of “imperfective” as opposed to *maá-* in the following section. Some examples are given below in (144)-(147). The *kâ-* in (144) specifies an action that occurs habitually every year; (145) is a statement on the speakers’ daily life; (146) describes a fact on the woman’s general way of walking. Example (147) is a Kukuya proverb on the spread of Christianity,

saying that though a leopard can always eat a messenger, but the messages s/he carries can ultimately remain en route. The tone pattern on the verb stem and the following prefix is the same as in the infinitive verb phrases, i.e. the metatonic H tone appears on the FV of the verb and the following nominal prefix.

(144) Ná mvúla me kâ-m-fúúm-á bí-ko.
 every 1.year 1SG.PRO IMPF-1SG.SM-buy-FV 8-clothes
 'Every year I buy clothes.'

(145) Bhií ba-lí-kâ-sál-á lí-kâ-bák-á mí-pará.
 1PL.PRO 2REL-1PL.SM-IMPF-work-FV 1PL.SM-IMPF-get-FV 4-money
 'We who work earn money.'

(146) Mu-kái wu-kíí-pí-í miilí mbaá kâ-dzalak-á
 1-woman 1REL-7SM.PST-catch-PST 4.leg 9.fire 1SM.IMPF-walk-FV
 yă mpúru.
 with 9.anger
 'The woman whose legs were burnt by fire walks with difficulties.'

(147) Ngo kâ-dzá ntúmá, mi-kiené
 1.leopard 1SM.IMPF-eat 1.messenger 4-message
 míí-siil-i mu mfúúlá.
 4SM.PST-remain-PST 18.LOC 9.road
 'The leopard eats the messenger, messages remained on the way.'

While expressing the career or common activities in one's daily life, the object is often seen fronted to the preverbal position as in (148), with the use of the habitual marker *kâ-*.

(148) (*answer to ‘what do you do in your life?’*)

Me báa-ntabá **kâ**-n-ték-e.
 1SG.PRO 2-goat IMPF-1SG.SM-sell-FV
 ‘I sell goats.’ (I am a goat-seller)

The prefix *kâ-* can also have the progressive reading, representing a situation in progress at and around reference time. The source of progressive reading may be an expansion from the general present that *kâ-* expresses. Some examples are given in (149)-(152). The similar *kâ-* prefix that is used to express habitual and progressive is also attested in other Teke varieties such as Teke-Eboo (SIL Congo, 2016, 2020).

(149) Mu-lúmi áá-yení mbhii, mu-kálí **kâ**-sál-á
 1-husband 1SM.PST-go.PST 9.hunting 1-wife 1SM.IMPF-work-FV
 nséke.
 9.field
 ‘The husband went for hunting, and the wife is cultivating in the field.’

(150) Me **kâ**-n-kín-a ña-kí-yím-á mú-kálí aa
 1SG.PRO IMPF-1SG.SM-dance-FV 16REL-7SM-sing-FV 1-wife 1.CONN
 me.
 1SG.PRO
 ‘I am dancing while my wife is singing.’

(151) ñama wǎ-ból-í **kâ**-wá mfúúlá nsúli.
 1.meat 1REL-decompose-PST 1SM.IMPF-give 9.road 9.smell
 ‘The rotten meat is making the street smelly.’

(152) Yáal-a nzó, tál-a mwáana ki-kálá **kâ**-ya.
 open-IMP 9.house see-IMP 1.child INF-stay 1SM.IMPF-come
 ‘Open the door and see if the child is coming.’

To encode an event that happens habitually or is happening at the moment, the tense marker *kâ-* is used. Sometime it can be difficult to determine which interpretation is meant in specific.

The aspect prefix *maá-* encodes perfective aspect, specifying an action as already finished, which can be translated as 'have (done)'. This perfective prefix may have its lexical origin in the verb *ki-maná* 'to accomplish' and has been grammaticalised. Some examples of this prefix are given in (153)-(156).

- (153) Mu-sálá wúa á-**maá**-fâ.
 3-work 3.DEM.II 3SM-PERF-end
 'The work has (been) finished.'
- (154) Mbhielé wu-kí-**maá**-pfiil-a maamá
 3.knife 3REL-7SM.PST-PERF-abandon-FV 1.mother
 báá-tól-i.
 2SM.PST-take-PST
 'The knife that mother has abandoned was picked up.'
- (155) Ma-tála bu-lii-báan-i bhiáwe, me
 AUX-see 14REL-1PL.SM.PST-begin-PST 1PL.PRO 1SM.PRO
 ka-**maá**-n-sak-á kí-lóko ku we ni.
 NEG-PERF-1SG.SM-search-FV 7-thing 17.LOC 2SG.PRO NEG
 'See that since we started, I have not asked for anything from you.'
- (156) Mbuurú wŭ-yi-n-tsuom-ó me pirí ya
 1.person 1REL-IMPF-1SG.SM-think-FV 1SG.PRO 1SG.COMP with
 má-lúa á-**maá**-kwâ.
 6-disease 1SM-PERF-die
 'The person who I was thinking that was ill has died.'

The prefix *yi-* is used to describe events that last over a period of time, or situations in which neither the beginning nor the end is mentioned, most

commonly in the past. I gloss it as “imperfective” but it also looks like a relative tense marker, such as situative or consecutive. This prefix may have its lexical origin in the word *ki-ye* ‘to go’. Some examples are shown in (157)-(159). In (157), the prefix *yi-* is used to express past progressive, specifying a “singing” event that lasted for a long duration; in (158) *yi-* marks an unaccomplished action; in (159) the event “coming to do harm to the chief” may also have lasted for some time before the soldier detected and killed the suspect. I leave more on the functions of this prefix for future research.

- (157) Ma-tsíka mu nkunkólo mu ma-táli mǎ wurí ma,
 6-yesterday 18.LOC 9.evening 18.LOC 6-sunshine 6REL like 6.DEM.I
 mu-kái á-yi-yím-á kí-pfupfumi kúko.
 1-woman 1SM-IMPF-sing-FV 7-song loudly
 ‘Yesterday evening at the same time, the woman was singing a song loudly.’

- (158) Li-ká-yi-som-ó ndé, mu-káli aa ndé
 5REL-1SM-IMPF-enter-FV 1.PRO 1-wife 1.CONN 1.PRO
 á-maá-yéem-e.
 1SM-PERF-sleep-FV
 ‘When he returned, his wife had already fallen asleep.’

- (159) Mu-mboró áá-dwî wũ-yi-yá kí-sá mpfúúmú
 1-soldier 1SM.PST-kill.PST 1REL-IMPF-come INF-do 1.chief
 bu-bí.
 14-badness
 ‘The soldier killed the one who came to do harm to the chief.’

Negative *bú-*

In the introduction on the Pre-initial slot, I mentioned that the negative prefix and the relative prefix cannot co-occur on the verb. In a relative clause, negation is usually expressed via the prefix *bú-* (see other negation strategies of relative clauses in section 2.5.3) which has its origin in the

verb *ki-bía* ‘to refuse’, and I suppose that it has at least partially been grammaticalised as a negative auxiliary, since in most cases it does not express an intentional refuse but just inverts the polarity and it can be attached directly to the verb without the use of an infinitive prefix (as in (161)). Some examples are shown in (160)-(163). This auxiliary can also be used in matrix clauses to mark negation, as in example (163).

- (160) Me n-dzií baarí ba-bá-bí-í kí-sá
 1SG.PRO 1SG.SM-like 2.people 2REL-2SM-refuse-PST INF-do
 má-ndáka maa-bi.
 6-word 6-bad
 ‘I like the people who do not complain.’

- (161) Mbuurú wǎ-bí-fur-á líi-mpó áá-dzáam-i.
 1.person 1REL-NEG-pay-FV 5-tax 1SM.PST-hide-PST
 ‘The person who did not pay the tax hid himself.’

- (162) Nkaaká bu-ka ná ndé káá-bí-í kí-wâ?
 1.grandmother 14-cassava 1.who 1.PRO 1SM.PST-refuse-PST INF-give
 ‘To whom didn’t the grandma give the cassava?’

- (163) Me se m-bí m-búon-o nzaami, me
 1SG.PRO even.if 1SG.SM-refuse 1SG.SM-believe-FV 1.God 1SG.PRO
 ka-kâ-n-dzwá bí-búólí ni.
 NEG-IMPF-1SG.SM-kill 8-organism NEG
 ‘Though I don’t believe in God, I don’t kill animals.’

Counterfactual *ku-*

The prefix *ku-* expresses a counterfactual meaning, which can usually be translated as ‘could/should have (done)’. The origin of this morpheme is not clear to me. Two examples are given in (164) and (165).

- (164) Ndé báá-ku-búk-í (ndé), yii-m-bí ndé á-maá-kwâ.
 1.PRO 2SM.PST-CF-cure-PST 1.PRO 9REL-9-bad 1.PRO 1SM-PERF-die
 ‘He could have been cured, unfortunately he has died.’
- (165) Me ali a-n-ku n-yá ŋa mu mvúla
 1SG.PRO RPST PST-1SG.SM-CF 1SG.SM-come 16.PRO 16.LOC 1.year
 wũ-fíŋa wu mu ki-ná kéki.
 1REL-pass 1.DEM.I 18.LOC 7-moment same
 ‘I should have been here at the same time last year’

2.4.1.4 Final vowel

In Kukuya, there are no verbal suffixes apart from the final vowel (FV). The FV appears as *-a* as its default form in infinitives and non-past tenses when the stem vowel is *a*, *i* or *u*, as a vowel copy suffix when the stem vowel is *e* or *o*, as shown in (166). As seen in many examples above, the FV occurs as *-i* in all past tenses regardless of the quality of the stem vowel, for presentation I show one more example in (167).

- (166) *ki-bhúima* ‘to empty’
ki-kúula ‘to be older’
ki-sóo ‘to search’
ki-téke ‘to sell’
ki-khéé ‘to try’
ki-lheeme ‘to become clear’
ki-bolo ‘to rot’
- (167) Me n-tsíúul-í mú-fiemé mu ki-wá ndzulí
 1SG.PRO 1SG.SM-make.fall-PST 3-bottle 18.LOC INF-give 1.cat
 buokó.
 14.fear
 ‘I made the bottle fall to scare the cat.’

Therefore, it seems that the FV alternation in Kukuya is used to encode

the past/non-past tense distinction, functioning together with the vowel lengthening tense prefix and grammatical tones like a circumfix. This is also what Nurse (2008) documented for some Bantu languages including Zone B which exploit the combination of two different positions such as TA and FV to carry one tense reference, where he notes that it is often “impossible to split the functions” of the two positions (Nurse 2008:81). However, in some grammar sketches of other Teke varieties, the FV is usually glossed as an aspect marker whereby -a and the vowel copy prefix encode imperfective and -i encodes perfective (Mouandza 2001 on Iyaa B73; Etsio 1999 and Raharimanantsoa 2012, 2017, 2020 on Eboo B74), or glossed as ambiguous between tense and aspect (Tsoue 2017 and Linton 2016 on Tege B71a; Calloc’h 1911 and Makouta-Mboukou 1976 on Fumu B77b). For Kukuya, in Paulian (1997: 213, 2001) she labelled the FV -a as MA ‘*marque d’aspect*’ and -i as MTA ‘*marque de temps et d’aspect*’.

An apparent counterargument for the FV in Kukuya to be an aspect marker is that when an aspect *prefix* occurs, whether the prefix is the perfective/resultative *maá-* or the imperfective/durative *yi-*, the FV can only take the default form -a or the vowel copy suffix but can never be -i, which shows that the FV alternation does not function to distinguish aspects. However, this co-occurrence constraint on the aspect markers and the FV may be explained as that diachronically the aspect prefixes such as *maá-* and *yi-* were very likely to have been grammaticalised from lexical verbs *ki-mana* ‘to finish’ and *ki-ye* ‘to go’, then they were phonologically reduced and fused together with a following *infinitive* verb stem, which may have an impact on the synchronic realisation of the FV on the verb complex that it can only take the unmarked -a as in infinitives.

In addition to these inconsistent analyses on whether the function of FV in Teke is more towards tense or aspect marking, it should also be noted that the FV alternation is also attested in conditional and subjunctive clauses (also see next subsection). One example is given in (168), in which the verb in the subjunctive clause takes the FV -i which is not related to past tense marking, so I suppose that the FV can also be used to encode the subjunctive mood.

- (168) Me ki-n-yáab-a pírí me bú-kíá
 1SG.PRO DEP-1SG.SM-know-FV 1SG.COMP 1SG.PRO 1.tomorrow
 â-ŋ-kwá, me n-ték-i bi-lóko bvíí me
 FUT-1SG.SM-die 1SG.PRO 1SG.SM-sell-SBJV 8-thing 8.CONN 1SG.PRO
 bvheí lo.
 8.all today

'If I know that I will die tomorrow, I would sell all my things today.'

Based on the facts introduced above, I propose that the FV in Kukuya, at least for its modified form *-i*, should not be treated as a single suffix with one unique function, but may have multiple homophonic counterparts which may have originated from different historical suffixes, whose functions are not confined to tense (and possibly aspect) but also mood marking. The suffix *-i* used in the past tenses may originally have represented perfective or anterior, and since the semantic shift from anterior to past is commonly attested cross-linguistically, it can result in the ambiguity on whether the FV encodes tense or aspect or both (Nurse 2008: 95). For simplicity, and also due to the fact that in Kukuya the alternation of FV is overwhelmingly used to encode past/non-past distinction, I treat the FV as a tense suffix when aspect and mood are not marked in a sentence.

2.4.2 TAM conjugations

The Kukuya language can distinguish as least three degrees of future tenses and two degrees of past tenses; the habitual, perfective and imperfective aspects; as well as the indicative/subjunctive/imperative moods. The distinctions on tenses are mainly expressed by tonal strategies on the subject-tense prefix and the final vowel, with the help of some independent tense auxiliaries. In this subsection I mainly focus on tense expressions. Since I have discussed aspect expressions in the description of aspect prefixes in the previous subsection, aspect is not discussed again here.

2.4.2.1 Near future tense

The near future tense in Kukuya usually encodes events which take place several hours later on the same day, or on the next day. The tone pattern on the verb of this tense is summarised in Table 2.27, and two examples are given in (169) and (170), one with a 2nd person subject, the other with the class 1 subject. We see that the H tone on the final vowel has also spread onto the following nominal prefix, it is not clear whether this is just a metatony effect or it should be characterised as the grammatical tone that marks future tense.

In (169) the subject-tense prefix of 2SG singular is realised as a L tone prefix, which consists of the L tone subject marker and the L grammatical tone of future tense. While in (170), the tone on the STM is realised as HL, which is combined by the H tone subject agreement marker for a 3rd person subject and the grammatical L tone for future tense. We also see that the grammatical L tone does not delete the subject agreement H tone.

Verbal tone		
SM	tense tone	Final vowel
1st and 2nd person: L 3rd person: H	L	H

Table 2.28: The verbal tone pattern in the near future tense

(169) We a-fúúm-á má-ko?
 2SG.PRO 2SG.FUT-buy-FV 6-banana
 ‘You will buy some bananas?’

(170) Ndé ló â-fúúm-á má-ko.
 1.PRO today 1SM.FUT-go-FV 6-banana
 ‘S/He will buy some bananas today.’

A note is added here on the cases such as (171) in which there is a focused

element occurring in the immediate preverbal position. We see that here the 1st person plural subject-tense marker in the near future tense is realised as HL instead of L. I will show in chapter 4 that this insertion of an additional H tone is related to relative marking. In this subsection, the tone pattern that I display for each tense are restricted to canonical word order.

- (171) Ngiemé ndíri a wu-mó ndíri ndúku, ka ki-bhiimá
 1.bat 1.COMP PREP 1-other 1.COMP 1.friend EMP 7-corpse
 kíí maamá [mu píriba] lí-dziik-á?
 7.CONN 1.mother 18.LOC 9.night 1PL.FUT-bury-FV
 ‘The bat says to the sunshine, ‘are we going to bury the corpse of
 mother AT NIGHT?’

2.4.2.2 Immediate future tense

The immediate future tense is used to express an intention which will be realised in very short time, and is encoded via the SOV word order. Some examples are given in (172)-(174) with different kinds of subjects. The tone pattern on the verb in the immediate future tense is similar to that in the near future with 3rd person subjects, which have a HL subject-tense marker. For 1st and 2nd person subjects, the tone pattern on the STM is also realised as HL, which is due to the insertion of an additional H tone in SOV that I mentioned above.

- (172) Bó má-ko báa-fúum-a.
 2.PRO 6-banana 2SM.FUT-buy-FV
 ‘They’ll buy some bananas.’
- (173) Ndé má-ko kâ-fúum-a.
 1.PRO 6-banana 1SM.FUT-buy-FV
 ‘S/He’ll get some bananas.’

- (174) Me nzó â-n-yé.
1SG.PRO 9.house FUT-1SG.SM-go
'I'll go home.'

In (173) when the subject is in class 1, we see that the subject marker shifts from *â-* to *kâ-*, which I will return to discuss in chapter 4 and 5. In causal speech, when the subject is 1SG/2SG as in (174), the STM *â-* can be optional and the speakers say that this vowel prefix can be just "swallowed".

2.4.2.3 Far future tense

The far future tense in Kukuya usually encodes events which will take place at least some days later or in the far future, it is not compatible with the time expressions such as "today" or "tomorrow". To mark this tense, an invariant auxiliary *lí* is placed before the verb, for some Kukuya speakers this auxiliary is pronounced as *líá*. The tone pattern of the far future tense on the verb is the same as that in the near future, which consists of the grammatical L tone on the STM and the H tone on the FV of the verb and the following prefix. See the examples in (175) and (176).

- (175) Bhií lí líi-tsú-á máa-nzó.
1PL.PRO FUT 1PL.SM-build-FV 6-house
'We will build some houses.'
- (176) Bó lí báa-fúúm-á bí-ko nziimi.
2.PRO FUT 2SM.FUT-buy-FV 8-clothes much
'They will buy many clothes.'

2.4.2.4 Recent past tense

The recent past tense in Kukuya usually encodes events which just took place several hours ago on the same day, or on the previous day. The tone

pattern of the recent past tense on the verb phrase is shown in Table 2.28. Some examples are given in (177)-(179).

Subject-tense prefix		FV+Prefix
SM	tense tone	grammatical tone
1st and 2nd person: L 3rd person: H	H	L

Table 2.29: The tone pattern of the recent past tense

(177) Bhií líí-fúum-i ma-ko.
 1PL.PRO 1PL.SM.PST-buy-PST 6-banana
 ‘We bought some bananas.’

(178) Me ló á-n-ká-i ba-ntsúú.
 1SG.PRO today PST-1SG.SM-grill-PST 2-chicken
 ‘Today I just grilled some chicken.’

(179) Bó báá-fúum-i ma-ko.
 2.PRO 2SM.PST-buy-PST 6-banana
 ‘They bought some bananas.’

In the above examples we find that the the grammatical H tone of recent past tense on the STM may have deleted and replaced the L tone of the 1st/2nd person subject markers, so the STM in (177) and (178) are realised as H rather than LH. In (179) the class 2 STM of recent past tense consists of a H subject marker and a grammatical H tone on the lengthened vowel, and is realised as a long prefix with H tone. In all these examples, the FV on the verb and the following nominal prefix surface as L, which may be due to a floating grammatical L tone that marks the past tense.

In the recent past tense, an optional auxiliary *âli* can be used and is often realised as *li* when contracted with the preceding subject, as shown

in (180). The presence of this auxiliary does not affect the tone pattern on the verb phrase. The speakers consider that there is no difference on the degree of past time reference whether there is the auxiliary or not, while the presence of the auxiliary may add some emphasis on the past tense. I leave the function of this auxiliary for further research.

- (180) Bóó li báá-fúum-i ma-keé.
 2.PRO PST 2SM.PST-buy-PST 6-tobacco
 ‘They just bought some tobacco.’

2.4.2.5 Remote past tense

The remote past tense encodes the events that happened as least some days ago or very far in the past. This tense requires the presence of an auxiliary *âli*, and the grammatical tone on the STM is L, which is different from the grammatical H tone in the recent past tense. The tone pattern of remote past tense is schematised in Table 2.29, and some examples are given in (181) and (182).

In (181) the STM prefix of 1PL is realised as *lii-* which consists of the L subject marker and the grammatical L tone; while in (182) the STM for a class 2 subject consists of the H subject marker and the grammatical L tone, which is realised as HL. In both examples the FV on the verb and the following nominal prefix surface as L.

Subject-tense prefix		FV+Prefix
SM	grammatical tone	grammatical tone
1st and 2nd person: L 3rd person: H	L	L

Table 2.30: The tone pattern of the remote past tense

(181) Bhií âli lii-fúum-i ma-ko.
 1PL.PRO PST 1PL.PST-buy-PST 6-banana
 ‘We had bought some bananas.’

(182) Bó âli báa-fúum-i ma-keé
 2.PRO PST 2SM.PST-buy-PST 6-banana
 ‘They had bought some tobacco.’

2.4.2.6 Sequence of actions

To express a sequence of actions, a dependent tense marker is attested on the dependent verb, for 1st and 2nd SG subjects, this dependent marker is *ki-*, for other subjects this dependent marker appears as *kâ-*. Two examples are given in (183) and (184).

(183) Me ki-n-ténem-e, n-dá-a bi-ko.
 1SG.PRO DEP-1SG.SM-stand.up-FV 1SG.SM-put.on-FV 8-clothes
 ‘Having stood up, I put on the clothes.’

(184) Ndé ká-ya nzó, ndé télek-i bvi-kídzá.
 1.PRO 1SM.DEP-come 9.house 1.PRO prepare- 8-food
 ‘Having returned home, s/he prepared the food.’

2.4.2.7 Imperative and subjunctive

An imperative sentence in Kukuya is usually marked by using a bare verb form without subject marking, and the verb does not show metatony effect, which may indicate that there is a verb-final grammatical L tone marking the imperative mood, as shown in (185).

- (185) a. Kúr-a ki-dzulibi!
 close-IMP 7-door
 ‘Close the door!’
 b. Fúum-a ma-li!
 buy-IMP 6-wine
 ‘Buy the wine!’

In hortative imperative sentences which express a command to a 3rd person, as shown in (186), we see that for a singular subject it takes the subject marker *ká-* on the verb, while for a plural subject the class 2 SM *bá-* is used. In both cases there is no overt subject, and there is a verb-final grammatical L tone blocking the metatony.

- (186) a. Ká-kúr-a ki-dzulibi!
 1SM.IMP-close-IMP 7-door
 ‘(Let) him close the door!’
 b. Bá-fúum-a ma-li!
 2SM.IMP-buy-IMP 6-wine
 ‘(Let) them buy the wine!’

The subjunctive mood is encoded by modifying the FV of the verb from the default *-a* to *-i*, which is formally identical to a past tense marking. Two examples are shown in (187) and (188).

- (187) Me n-dzií ndé yení Mpúru.
 1SG.PRO 1SG.SM-like 1.PRO go.SBJV 9.Europe
 ‘I want that he goes to Europe.’
- (188) Me ka-n-kukí mu ki-yé ŋa-kítí we kí-na
 1SG.PRO NEG-1SG.SM-can 18.LOC INF-go 16-COMP 2SG.PRO 7-period
 ka-á-kaal-í ndáka aa me ni.
 NEG-2SG.SM-answer-SBJV 9.question 9.CONN 1SG.PRO NEG
 ‘I cannot go if you don’t answer my question now.’

To summarise, in this section I have shown how the verb complex is structured in Kukuya and what sorts of verbal affixes can occur on it. I also presented how different tenses, aspects and mood are expressed. Next I discuss some syntactic issues in this language.

2.5 Clause structure

In the final section of the grammar sketch, I present some syntactic issues in Kukuya. Since the canonical word order and deviation from it, which is influenced by information structure, are discussed in detail in chapter 3, and the syntax of relative constructions is analysed in chapter 5, I don't use much space to discuss them in this section but just give some brief description. I first introduce the forms and functions of different types of complementisers in Kukuya. Then I provide a description on the functions of various types of relative constructions and adverbial clauses. At the end of the section I shed some light on the non-verbal predication.

2.5.1 Complementisers

2.5.1.1 Say-complementiser

In this subsection, I introduce the *say*-complementiser that heads an embedded declarative complement clause. In a large number of Bantu languages, there is a type of complementiser which is diachronically derived from Proto-Bantu **tì*, which in many languages has evolved into an element meaning “say” as well as a complementiser (Meeussen, 1967; Güldemann, 2002, 2008; Gluckman 2023). In some sources that document Bantu complementisers, this *say*-complementiser is mentioned as also related to “evidentiality” (Botne 1995, 2020; Devos and Bostoen 2012).

In Kukuya, an embedded complement clause that is selected by the perceptive verbs such as “think”, “see”, “say” and “know” is commonly headed by a complementiser-like element, which at first glance seems to agree with the matrix subject in person and number. The forms of this agreeing complementiser are summarised in Table 2.30.

From the table we see that the agreement pattern of this complementiser is relatively exceptional, since it does not show obvious correlations with the subject agreement or adnominal concord prefixes. The forms of these

Person	singular	plural
1st	píri	píri
2nd	wurí	wurí
3rd	ndíri	bóri

Table 2.31: The forms of the complementiser *-ri*

complementisers seem to be derived from the contraction of person pronouns and an element *-rí*. for example *ndíri* is derived from “*ndé+rí*” and *wurí* from “*we+rí*”, which still needs further research.

Some straightforward examples are shown in (189)-(192). In these examples, the complementiser agrees with the matrix subject in person and number.

- (189) Ndé kâ-tsuom-ó **ndíri** mbuká taará á-fúum-i
 1.PRO 1SM.IMPf-think-FV 1.COMP 9.bed 1.father 1SM.PST-buy-PST
 ku dzáandu.
 17.LOC 5.market
 ‘S/He thinks that the bed was bought by father in the market.’

- (190) Mu-kiliŋomo ka-bir-á, baarí ba-bá-li ku
 3-drum 3SM.DEP-sound 2.people 2REL-2SM-COP 17.LOC
 ntáli bá-yuk-á mu-kiliŋomo báá-yáab-i
 9.other.place 2SM-hear-FV 3-drum 2SM.PST-know-PST
bóri mpfúúmú wŷ-yi-lúa á-maá-kwâ.
 2.COMP 1.chief 1REL-IMPf-sick 1SM-PERF-die
 ‘The drum sounds, people who are in far places hear it and know that the chief who was ill has died.’

- (191) Nzaamí â-lak-í **ndíri** li-bák-á kí-lóko,
 1.God 1SM.RPST-say-PST 1.COMP 1PL.SM-get-FV 7-thing
 li-káb-í baa-mó.
 1PL.SM-share-SBJV 2-other
 ‘God said (if) we get something we share with others.’

- (192) Me n-dzwî li-bakí píri téme ku
 1SG.PRO 1SG.SM-kill.PST 5-connection 1SG.COMP 5.hoe 17.LOC
 kíí taará á-m-fu-í.
 7.CONN 1.father PST-1SG.SM-put-PST
 ‘I forgot that I put the hoe in father’s place.’

This complementiser can never show class agreement. In example (193) where the matrix subject is in class 4, the complementiser takes the form *bóri*, which shows that it only reflect the number and person feature of the matrix subject.

- (193) Mi-fémé mí-kâ-tsuom-ó **bóri** me njé
 4-pig 4SM-IMPF-think-FV 2.COMP 1SG.PRO 4.PRO
 â-n-siiba.
 FUT-1SG.SM-attack
 ‘The pigs think that I will attack them.’

When I asked the speakers about the meaning of this complementiser, they clearly said that it means “to say”, and in many cases it can just replace the verb *ki-lák* ‘to say’, as shown in (194). However, its inflectional pattern is very different from a common verb; it can also co-occur with the verb ‘to say’, as shown in (191) above; but it cannot be modified by a manner adverb as in (195). Therefore I suppose that it has been grammaticalised and functions as a complementiser. Prosodically, there is often a clear intonation break after this complementiser.

- (194) Me píri a ndé.
 1SG.PRO 1SG.COMP PREP 1.PRO
 ‘I said to him.’

- (195) Ndé ndíri tswáatswáa.
 1.PRO 1.COMP quickly
 ‘S/He said ‘be quickly’!’ (*S/He spoke quickly.)

Interestingly, in many contexts this complementiser does not necessarily agree with the matrix subject, but it always agrees with the perspectival source of the embedded event, namely the logophoric center. In (196) and (197) the complementisers apparently do not agree with the matrix subject. The matrix verb in (196) is ‘to believe’ and the source of information is from the addressee, so the complementiser takes the 2SG form, which may literally mean ‘I don’t believe that (you say)’; while in (197) the verb is ‘to hear’ and the source of the news is from some other people, so the complementiser takes the 3PL form, which has the reading of ‘I heard that (they say)’. I suppose that the complementiser here agrees with a covert operator that is linked to the perspectival source of the embedded event. I leave the more elaborated analysis on the agreement patterns of these *say*-complementisers and selectional properties of the matrix verb for future research.

- (196) Me ka-kâ-n-kiik-a wu-rí ba-lóbi ntsúí
 1SG.PRO NEG-IMPF-1SG.SM-believe-FV 2SG.COMP 2-fisherman 1.fish
 wúna ngúumo báá-lóob-i ni.
 only 1.one 2SM.PST-fish-PST NEG
 ‘I don’t believe that the fishermen caught only one fish.’

- (197) We á-yúk-i bóri mbuurú mu báa kuní
 2SG.PRO PST-hear-PST 2.COMP 1.person 18.LOC 9.fire 17-which
 káá-kwî?
 1SM.PST-die.PST
 ‘Where (the place of hearing) did you hear that the man died of fire?’

2.5.1.2 Embedded questions

When the embedded clause is a yes-no question, it is introduced by a complementiser *káli* ‘if, whether’, as shown in (198) and (199). Interestingly, in (199) we see that the complementiser *ndíri* and *káli* can co-occur, which is reminiscent of the similar ‘if-that’ co-occurrence in some Germanic languages, despite that the order of the two complementisers is just the reverse of that in Kukuya here. The shape of this complementiser *káli* is also identical to the conditional marker (see section 2.5.2.3 below).

- (198) Maamá áa-fúul-i me **káli** me n-dzíí ki-yé
 1.mother 1SM.PST-ask-PST 1SG.PRO if 1SG.PRO 1SG.SM-like INF-go
 tála bi-dziladzíli yă ndé.
 look 8-film with 1.PRO
 ‘Mother asked me if I want to go to watch the film with her.’

- (199) Mpfúúmú wu-kâ-n-sál-á me kúkí ndé
 1.chief 1REL-IMPV-1SG.SM-work-FV 1SG.PRO PREP 1.PRO
 áá-fúul-i me **ndíri káli** me maá-m-bák-á
 1SM.PST-ask-PST 1SG.PRO 1.COMP if 1SG.PRO PERF-1SG.SM-get-FV
 mú-pára.
 3-money
 ‘My patron whom I work for asked me if I have received the money.’

When the embedded clause is a *wh*-question, the interrogative word can either be placed in the immediate-before-verb position, as shown in (200) and (201); or what is embedded can be a relative clause that is headed by an interrogative element, as in example (202).

- (200) Me ka-n-yáab-i bó ka **munkí** báa-tó ni.
 1SG.PRO NEG-1SG-know-PST 2.PRO EMP when 2SM.FUT-arrive NEG
 ‘I did not know when they would arrive.’

- (201) Ndé ká-tsuom-ó ndíri kǐ-má kǐí-wéén-i ku
 1.PRO 1SM.IMPf-think-FV 1.COMP 7-what 7SM.PST-lack-PST 17.LOC
 ndé.
 1.PRO
 ‘He thinks about what he lacked.’

- (202) Me ka-n-yáabi ku-ní ku-m-báka me
 1SG.PRO NEG-1SG-know-PST 17-which 17REL-1SG.SM-get-FV 1SG.PRO
 ma-dzá ni.
 6-water NEG
 ‘I did not know (the place) where I could get water.’

2.5.2 Relative constructions

In this subsection I provide a brief description on the forms and functions of relative constructions in Kukuya. I first introduce the structure of subject and non-subject relatives, then I present diverse functions of relative constructions for conditional, temporal and reason expressions. Since the agreement pattern of relatives will be analysed from a syntactic perspective in chapter 5 (section 5.3), here I don’t discuss it in detail.

2.5.2.1 Subject and non-subject relatives

The verb in Kukuya relative constructions consists of a relative marker which agrees with the relativised NP in noun class, and a subject marker which can also agree with the subject, but depending on the relative position of the subject to the verb. The agreeing relative marker always takes the same shape as the proximal demonstratives (see section 2.3.2 and Table 2.25) and bears a L tone, while the subject marker always has a H tone. Some examples on the subject and non-subject relative clauses (in brackets) are given in (203)-(206).

- (203) Amenda yǎ [ko lí ndé li-lí-maá-bol-ó].
 1.Amenda with 5.banana 5.CONN 1.PRO 5REL-5SM-PERF-rot-FV
 ‘Amenda has her banana that has rotten.’
- (204) [Mu-kái wǔ-kwî mu ntsá ma-dzá] nkélé aa
 1-woman 1REL-die.PST 18.LOC inside 6-water 1.sister 1.CONN
 me.
 1SG.PRO
 ‘The woman who died in the river is my sister.’
- (205) [Mi-féme mi-kíí-fúúm-í mú-kái] míí-bár-i.
 4-pig 4REL-7SM.PST-buy-PST 1-woman 4SM.PST-escape-PST
 ‘The pigs that the woman bought escaped.’
- (206) Mbuká yi-kíí-sweek-í mú-kái ntséke njíibi
 9.place 9REL-7SM.PST-hide-PST 1.woman 9.product 1.thief
 áá-swoolí (yó).
 1SM.PST-find-PST 9.PRO
 ‘The place where the woman hid food was found by the thief.’

From these examples, we see that in the subject relative (203), the relativised subject NP is preverbal, being indexed both by a relative marker and a subject marker. When the relativised subject is in class 1, as in (204), the class 1 SM is usually suppressed thus the relative marker bearing a rising tone. In the object relative (205) and adjunct relative (206), the relativised NPs agree with the relative pronouns, while the subjects are placed in a postverbal position and do not show canonical subject agreement, but a class 7 SM *kí-* appears on the verb. I introduce the structure of Kukuya non-subject relatives a bit more here.

In a Kukuya non-subject relative, the subject is always placed postverbally but can never be preverbal. The postverbal subject does not always show agreement on the verb, but the shape of SM varies according to the type of the subject. When the postverbal subject is the 1st and 2nd person

pronouns, the SM appears just as the agreeing subject markers (see Table 2.26 above on the shape of SMS), as shown in (207) and (208).

- (207) Yǎ nkú má ngámo yi-má-n-tá **me.**
 with 9.story 9.one 9REL-AUX-1SG.SM-tell 1SG.PRO
 ‘There is a story that I will tell.’

- (208) Baa-mvá ba-líí-yíí **bé** ná nkunkólo
 2-dog 2REL-2PL.SM.PST-bring.PST 2PL.PRO every 9.evening
 bá-kâ-bík-a.
 2SM-IMPF-bark-FV
 ‘The dogs that you(pl.) brought bark every evening.’

When the postverbal subject is a lexical DP or class pronoun, the SM slot can only be filled by a class 7 default subject marker *-kí-* but cannot show noun class agreement with the subject, as shown in (209a) and (209b).

- (209) a. Ki-pfúo ki-kíí/*báá-télék-í bá-káli báa-nzulí
 7-bread 7REL-7/*2SM.PST-prepare-PST 2-wife 2-cat
 báá-dzí.
 2SM.PST-eat.PST
 ‘The bread that wives prepared was eaten by the cats.’
 b. Bviila bvi-kíí/*míí-dz-í ɲé
 8.food 8REL-7/*4SM.PST-eat-PST 4.PRO
 ‘the food that they (the pigs) ate’

An exceptional case is when the postverbal subject is the class 1/2 pronouns *ndé* or *bó*, which can trigger either the default SM *kí-* or real subject agreement, as illustrated in (210). I will discuss these differential SM strategies in non-subject relatives further in chapter 5.

- (210) a. Ntaba wu-**káá/kíi-lélék-í ndé** mu mbára mú-tí
 1.goat 1REL-1/7SM-tie-PST 1.PRO 18.LOC against 3-tree
 á-maá-kol-o.
 1SM-PERF-be.weak-FV
 ‘The goat that s/he tied against the tree has become weak.’
- b. bi-ko bi-**báa/kíi-lí** báa/kíi-ték-í bó
 8-clothes 8REL-2/7SM.RPST-COP 2/7SM.RPST-sell-PST 2.PRO
 ‘the clothes which they had sold’

Another co-referential relative pronoun may also follow the postverbal subject, but it is not obligatory, as the class 9 REL *yi* shown in (211). This clause-final use of the relativiser is also reported in Nsong (B85d, Koni Muluwa and Bostoen 2019).

- (211) Mfúúlá yi-kâ-n-yé me yi báa-sinwá
 9.road 9REL-IMPF-1SG.SM-go 1SG.PRO 9.REL 2-Chinese
 báá-sî (yó).
 2SM.PST-do.PST 9.PRO
 ‘The road on which I am walking was built by the Chinese.’

A summarising table on the shape of SMS in Kukuya non-subject relatives is given in Table 2.32 below.

SM	1st	2nd	class 1/2 pronouns	others
SG	-N-	-Ø-	-ka-/-ki-	-ki-
PL	-li-	-li-	-ba-/-ki-	

Table 2.32: Subject marking in Kukuya non-subject relatives

In some cases the thematic relation between the extracted element and the subject can be rather complicated. In examples (212) and (213), the relativised NPs ‘girl’ and ‘person’ do not seem to be arguments of the verb ‘to die’ and ‘to go away’, but rather the experiencer of the whole event. In (214) the

head NP ‘reason’ is not an argument of the verb ‘to cry’, but some implicit thematic relations may be involved.

- (212) [Mwáana wu-mu-kái wu-kíi-kwí ngúku]
 1.child IREL-1AGR-female IREL-7SM.PST-die.PST 1.mother
 á-yika kí-lílá kâ-líl-a.
 1SM-IMPF INF-cry 1SM-IMPF-cry-FV
 ‘The girl whose mother died is crying.’

- (213) [Mbuurú wu-kíi-yení mu-káli yă baraka
 1.person IREL-7SM.PST-go.away.PST 1-wife with 1.man
 mu-kimá] mí-kíi kâ-ló ná tsúku.
 1-other 4-curse 1SM-IMPF-curse every 5.day
 ‘The person whose wife went away with another man curses every-day.’

- (214) [Ki-sáli ki-káá-líl-í ndé] ka-kí-li tsítsie ni.
 7-reason 7REL-1SM.PST-cry-PST 1.PRO NEG-7SM-be clear NEG
 ‘The reason why s/he cried is not clear.’

In a non-subject relative that involves more than two arguments, the subject is commonly placed right-adjacent to the verb, as shown in (215) and (216).

- (215) Mwáana wu-mu-balaka wu-kíi-pfur-í **báa-ndúku**
 1.child IREL-1AGR-male IREL-7SM.PST-cheat-PST 2-friend
 taará á-dzií ki-yé bíta.
 1.father 1SM-like INF-go revenge
 ‘The boy whose father was cheated by friends wants to revenge.’

- (216) Mu-kái wu-kíi-pí-í **miilí** mbaá kâ-dzalak-á
 1-woman IREL-7SM.PST-catch-PST 4.leg 9.fire 1SM-IMPF-walk-FV
 yă mpúru.
 with 9.force
 ‘The woman whose legs were burnt by fire walks with difficulties.’

The relative marker is not always attached to an inflected verb, but can also be followed by a noun phrase to modify the head NP. In (217) the relative marker is followed by a noun phrase *míaka mi-lilá* ‘long arms’ to express ‘long sleeved clothes’. The same fact is observed in (218), in which the intended meaning for *wǔ ngambú* may be ‘the one who is the elder twin’ and a copula seems to be elided between the relative marker and the noun phrase. The relative part in (219) has a non-restrictive reading, and we also see that the relative marker is followed directly by a noun phrase.

- (217) Me n-lá-i ki-kó ki-kí-níni **ki-kí**
 ISG.PRO ISG.SM-wear-PST 7-clothes 7REL-7AGR-big 7REL-7AGR
míaka mi-lilá bu-kítí bi-mbú bí-niak-a ki-tsá
 4.hand 4AGR-long 14-COMP 8-mosquito 8SM-abandon-FV INF-bite
 me.
 ISG.PRO
 ‘I wore a big coat with long sleeves in order that the mosquitoes do not bite me.’

- (218) Yǎ mpiiní ku me mu ki-yáaba pirí **bǎ**
 with 9.force 17.LOC ISG.PRO 18.LOC INF-know ISG.COMP 2REL
báa-mbú wǔ ngambú wǔ ngampíka.
 2-twin 1REL 1.elder.twin 1REL 1.younger.twin
 ‘It is difficult for me to distinguish which is the elder twin and which is the younger.’

- (219) ndúku aa me, wǔ mwáana aa mvá
 1.friend 1.CONN ISG.PRO 1REL 1.child 1.CONN 1.dog
 ‘my friend, the puppy’

2.5.2.2 Complementiser *-kítí*

Relative constructions can also be expressed via a set of complementisers with the root *-kítí* that always agree with the head NP. The origin of this complementiser is not clear to me. We may notice that the root *-kítí* has

a HLH tone pattern that does not fit into the five tone melodies of the language, which may suggest that its *-kí-* part was possibly a prefix and *-tĩ* was the stem in some earlier stage.

This complementiser is used in both subject and non-subject relatives, as shown in (220)-(222). We see that in these examples the complementiser *-kítĩ* always agrees with the relativised NP in noun class. Since this relative marker is used independently and is not attached to the verb complex, in the non-subject relatives (221) and (222) the use of *-kítĩ* does not require inversion of the subject.

- (220) Mwáana **wu-kítĩ** ba-búri ka-bá-dzií ndé ni
 1.child 1-COMP 2-parent NEG-2SM-like 1.PRO NEG
 áá-mar-í ɲa nzó.
 1SM.PST-leave-PST 16.LOC 9.house
 ‘The boy who is not liked by the parents left home.’

- (221) Me n-sweék-í lóoro ɲa mbúka yi-kítĩ
 1SG.PRO 1SG.SM-hide-PST 5.gold 16.LOC 9.place 9-COMP
 wũ-yáab-a yó wúna me-nkúlu.
 1REL-know-FV 9.PRO only 1SG-oneself
 ‘I hid the gold in a place that the one who knew it is only I myself.’

- (222) Ya bí-na **bi-kítĩ** me ka-n-li ya
 with 8-period 8-COMP 1SG.PRO NEG-1SG.SM-be with
 baá-n-kab-a bó ki-sáábí kíí me ni.
 2REL-1SG.SM-share-FV 2.PRO 7-happiness 7.CONN 1SG.PRO NEG
 ‘There are some moments when I do not have the ones with whom I share my happiness.’

The *-kítĩ* is commonly seen in the negation of relative clauses. Since the negative prefix *ka-* can never appear on a relative verb, when expressing negation in a relative clause, the complementiser *-kítĩ* is used, which is followed by a non-relative verb form that allows the use of the negative prefix *ka-*, as shown in examples (223)-(226).

- (223) Mbuurú **wu-kítĩ** ka-ká-fur-í mpákí ni
 1.person 1-COMP NEG-ISM.PST-pay-PST 9.tax NEG
 áá-dzáam-i.
 ISM.PST-hide-PST
 ‘The person who did not pay the tax hid himself.’
- (224) Bhií ka-líí-wol-í baarí **ba-kítĩ**
 IPL.PRO NEG-IPL,SM-take-PST 2.people 2-COMP
 ka-bá-yáab-i ki-líra ni.
 NEG-2SM-know-SBJV INF-read NEG
 ‘We do not employ people who cannot read.’
- (225) **Ki-kítĩ** we ka-á-dzií ki-múna ni kí-ma?
 7-COMP 2SG.PRO NEG-PST-like INF-see NEG 7-what
 ‘What didn’t you see?’
- (226) Ma-dzá **ma-kítĩ** ka-má-li má-bvé ni
 6-water 6-COMP NEG-6SM-be 6AGR-good NEG
 báá-yiir-i (mó).
 2SM.PST-overthrow-PST 6.PRO
 ‘The water which is not clean was overthrown.’

2.5.2.3 Adverbial clauses

Subordinate adverbial clauses in Kukuya are generally relative clauses and frequently involve the use of class 16 and 17 locative markers. Temporal clauses are usually expressed by using a class 16 relative prefix which does not agree with an overt NP. Two examples are shown in (227) and (228).

- (227) *ɲa-kíí-kín-á* *bá-kái, me*
 16REL-7SM.PST-dance-FV 2-woman 1SG.PRO
wu-kâ-n-yím-a.
 1REL-IMPF-1SG.SM-sing-FV
 ‘While the women are dancing, I am singing.’

- (228) *ɲa-n-som-í* *me nzó, maamá kí-ni*
 16REL-1SG.SM-enter-PST 1SG.PRO 9.house 1.mother 7-period
ka-ká-yi-bvúruk-a ni.
 NEG-1SM-IMPF-return-FV NEG
 ‘When I entered home, mother did not come back yet.’

To express ‘until’, the preposition *tíí* is used, which is followed by a relative construction with the class 17 relative prefix *ku-*, as shown in examples (229) and (230).

- (229) *Me n-kalá ɲa tíí ku-m-bak-á me*
 1SG.PRO 1SG.SM-stay 16.PRO until 17REL-1SG.SM-get-FV 1SG.PRO
wũ-fur-á me tiki.
 1REL-pay-FV 1SG.PRO 5.ticket
 ‘I stay here until I have someone pay the ticket for me.’

- (230) *Mwáana wúna mu kâ-líla tíí ku-kí-yi-bvúruk-a*
 1.child only 18.LOC IMPF-cry until 17REL-7SM-IMPF-return-FV
ngúku.
 1.mother
 ‘The child remained crying until mother returned.’

The conditional meaning can be expressed by using the complementiser *ɲa-kítí* which also takes the class 16 prefix, as illustrated in (231) and (232).

- (231) **ŋa-kítĩ** me ka-man-í bi-lokó bvi-kí-sá ló
 16-COMP 1SG.PRO NEG-1SG.finish-SBJV 8-thing 8REL-7SM-do today
 bu ni, bu-kía mpfúúmú aa me
 14.PRO NEG 14-tomorrow 1.chief 1.CONN 1SG.PRO
 â-béer-e me.
 1SM.FUT-beat-FV 1SG.PRO
 ‘If I do not finish the work today, my patron will punish me tomorrow.’
- (232) Me ka-kâ-n-soom-ó nzó **ŋa-kítĩ** mu-káli
 1SG.PRO NEG-IMPf-1SG.SM-enter-FV 9.house 16-COMP 1-woman
 aa me ka-ká-yíí mu ki-búon-o ni.
 1.CONN 1SG.PRO NEG-1SM-come 18.LOC INF-apologize NEG
 ‘I do not return home if my wife does not come to apologize.’

A conditional clause can also be expressed by using a particle *káli* which is formally identical to the complementiser for embedded yes-no questions, as shown in (233a). Another strategy is to use a L tone prefix *ki-* on the verb, whose primary use may be to mark clause dependency, see (233b). The two strategies expressing conditional meaning cannot be applied at the same time.

- (233) a. **Káli** me n-yáaba pirí me bú-kía
 COND 1SG.PRO 1SG.SM-know-FV 1SG.COMP 1SG.PRO 14-tomorrow
 á-ŋ-kwá, me n-ték-i bi-lóko bvíí
 FUT-1SG.SM-die 1SG.PRO 1SG.SM-sell-SBJV 8-thing 8.CONN
 me bvheí ló.
 1SG.PRO 8.all today
 ‘If I know that I will die tomorrow, I would spend all my money today.’

- b. Me ki-n-yáaba pirí me bú-kíá
 ISG.PRO DEP-1SG.SM-know-FV ISG.COMP ISG.PRO 14-tomorrow
 â-ŋ-kwá, me n-ték-i bi-lóko bvíí
 FUT-1SG.SM-die ISG.PRO ISG.SM-sell-SBJV 8-thing 8.CONN
 me bvheí ló.
 ISG.PRO 8.all today
 ‘If I know that I will die tomorrow, I would spend all my money
 today.’

To express intention or purpose in Kukuya, the complementiser *-kítĩ* is used to introduce the aim of the action. It can take the class 14 or class 17 prefixes, as shown in (234) and (235).

- (234) Me mpólo kâ-n-sak-á **bu-kítĩ** me
 ISG.PRO 9.chance IMPF-1SG.SM-search-FV 14-COMP ISG.PRO
 m-pál-a.
 ISG.SM-go.out-FV
 ‘I’m looking for a chance for me to go out.’

- (235) Mu-kái áá-lúok-i kúkó **ku-kítĩ** baari bhoí
 1-woman 1SM.PST-shout-PST loudly 17-COMP 2.people 2.all
 bá-yúk-á.
 2SM-hear=FV
 ‘The woman shouted loudly in order that all the people hear.’

The expression of reason is usually introduced by the class 18 preposition *mu*. It can be followed simply by a nominal phrase as the causer (see section 3.1.1 on the use of locative markers), or by a relative clause explaining the reason, which takes the class 14 relative prefix or the complementiser *bu-kítĩ*, as shown in (236)-(238).

- (236) Maamá bu ká-tsiim-á nziimí mu bu-kítĩ ndé
 1.mother 14.PRO 1SM.IMPf-regret-FV much 18.LOC 14-COMP 1.PRO
 li á-wĩ ngálíbakí ñamá wu-mu-bi.
 PST 1SM-give.PST 1.baby 1.meat 1REL-1AGR-bad
 ‘The mother regret much that she gave the baby bad meat.’
- (237) Mu-kokó mú-káli áá-tsilik-í mu-tswé mu
 1-king 1-wife 1SM.PST-cut-PST 3-head 18.LOC
 bu-kí-yi-sak-á ndé ki-woló mu-kái mu-kíma.
 14REL-7SM-IMPf-search-FV 1.PRO INF-take 1-woman 1-other
 ‘The king was cut the head by his wife because he looked for and
 married another woman.’
- (238) Me yă ngúbalaka wu-kítĩ kâ-wá me mu-para
 1SG.PRO with 1.uncle 1-COMP 1SM.IMPf-give 1SG.PRO 3-money
 mu bu-m-wéén-á mú-káli.
 18.LOC 14REL-1SG.SM-lack-FV 1-wife
 ‘I have an uncle who often gives me money because I do not have a
 wife.’

The concessive meaning is expressed by the particle *se* which means ‘though’ or ‘regardless of’. In (239) the particle is followed by a relative clause that takes the class 14 relative prefix. In (240) the meaning ‘regardless of/however’ is encoded by a fixed expression *se bálakí bóri* which literally means ‘although they say that...’. In (241) ‘unless’ is expressed by a paraphrased phrase which literally means ‘(if) badness comes to him’.

- (239) Se bu-ká-kwí ndé, ki-tóli kii ndé
 though 14REL-1SM.PST-die.PST 1.PRO 7-spirit 7.CONN 1.PRO
 kii-kal-á kí-na.
 7SM.FUT-stay-FV 7-time
 ‘Although he is dead, his spirit will last forever.’

- (240) Se bá-lak-í bóri bu-kí-tsuom-ó maamá, me
 though 2SM-say-PST 2-COMP 14REL-7SM-think-FV 1.mother 1SG.PRO
 ka-n-kukí mu ki-sóó mí-tsíomi míí me ni.
 NEG-1SG.SM-can 18.LOC INF-change 4-idea 4.CONN 1SG.PRO NEG
 ‘However (they say that) mother thinks, I cannot change my ideas.’
- (241) Ndé lí â-yá, [mu ki-bí kí-yá ndé] ndé
 1.PRO FUT 1SM.FUT-come 18.LOC 7-badness 7SM-come 1.PRO 1.PRO
 kâ-kalá ya má-lúa.
 1SM.IMPf-stay with 6-illness
 ‘He will come, unless he will still be sick.’

2.5.3 Copula and non-verbal predication

The most common strategy for non-verbal predication in Kukuya is to use the copula which generally has the form *-li* and takes the subject agreement prefixes. The copula is usually used to express identification, which includes equation and qualitative specification. Two examples are given in (242) and (243).

- (242) Kímá ké kí-li mu ntsá nkíé?
 7-what 7.PRO 7-COP 18.LOC inside 9.pot
 ‘What is it in the pot?’
- (243) Me ka-n-li Albert ni.
 1SG.PRO NEG-1SG.SM-COP Albert NEG
 ‘I am not Albert.’

Adjectives can also function as predicative and are introduced by the copula. In (244) the adjective *mú-bvé* agrees with the class 1 subject and serves as predicative introduced by the negative copula *ka-ká-li* that also agrees with the subject. While in (245) *mpémbe* has the reading of ‘white’ but does not take class 6 agreement with the subject, and only the clause-final copula

takes the agreeing prefix. This may be due to the fact that *mpémbe* is not a true adjective thus does not show any agreement.

- (244) Mwáana wu-balaka ka-ká-li mú-bvé ni.
 1.child 1-male NEG-ISM-COP IAGR-good NEG
 ‘The boy is not good-looking.’

- (245) mjíini máá mwáana wu-mu-kái mpémbe má-li.
 6.teeth 6.CONN 1.child 1-LAGR-female white 6-COP
 ‘The teeth of the girl are white.’

In many cases of non-verbal predication, especially in affirmative sentences, the copula is just omitted. In (246) and (247) the predicative expression is realised by juxtaposing two NPs in which the prefix of the second NP bears a H tone. Example (248) expresses possessive meaning and the possessive phrase *máá biábe* ‘of us’ is predicative. The pseudo-cleft in (249) is formed by a free relative plus a predicative NP with a H tone prefix, and the copula is omitted. See more examples of cleft constructions in chapter 3 section 3.4.

- (246) Ndé mú-tsúli.
 1.PRO 1-goldsmith
 ‘He is a goldsmith.’

- (247) Ki-báka kí-báka, bu-bila.nkele múu-nkwáará.
 7-obtain 7-obtain 14-question 3-keeping
 ‘To obtain is to obtain, the question is (how) to keep.’
 (Paulian 1975: 194, glossing added by the author)

- (248) Ma-sáani ma-báa-niak-i máá biábe.
 6-plate 6REL-2SM.PST-abandon-PST 6.CONN 1PL.PRO
 ‘The plates that are abandoned are ours.’

- (249) Kĩ-n-dzii me ki-nywâ má-dzá maa-mfé.
 7REL-1SG.SM-please 1SG.PRO INF-drink 6-water 6-cold
 ‘What I like to drink is cold water.’

In Kukuya, ‘to have’ is usually expressed by using the preposition *yă* in a non-verbal predication, which literally means ‘be with’, as shown in (250) and (251) whereby the copula is omitted. In (251), ‘be happy’ is expressed by *ya kí-sáabi* which is literally translated as ‘with happiness’.

- (250) Mu-tsúli ma-biele ya mú-dzí wu-lilá wu-ŋwa.
 1-forge 6-iron with 3-rope 3-long 3-red
 ‘The blacksmith has a long red rope.’
- (251) Mwáana wu-bá-wî bu-ká ka ná ndé ya
 1.child 1REL-2SM-give.PST 14-cassava EMP 1.who 1.PRO with
 kí-sáabi?
 7-happiness
 ‘The child that who gave cassava to is happy?’

I suppose that that in the examples above the copula is underlyingly present in the structure but is just unpronounced. As in examples (252) and (253), the copula *-li* can appear in negation and relative constructions.

- (252) William ka-ká-li yă ntá nna ni, ndé ntá
 William NEG-1SM-COP with 10.melon four NEG 1.PRO 10.melon
 tíri ká-li yă yó.
 three 1SM-COP with 10.PRO
 ‘William does not have four melons, he has three melons.’
- (253) Ndé ba-kái ba-ká-yááb-í ndé wúna ba-bá-li
 1.PRO 2-woman 2REL-1SM-know-PST 1.PRO only 2REL-2SM-COP
 ŋa ki-yínga.
 16.LOC 7-festival
 ‘He only knew the women who were at the festival.’

CHAPTER 3

Word order and topic/focus expressions

This chapter provides a description on the expression of information structure in Kukuya. In Kukuya there is crucial morphosyntactic variation that cannot be accounted for by the traditional point of view on the grammar but must be explained with reference to information structure. In this chapter I show that the word order in Kukuya is to a large extent determined by information structure more so than by grammatical relations. All kinds of topical elements tend to occur in the preverbal domain, while focused elements are usually placed in a dedicated immediate-before-verb (IBV) focus position which is rarely attested in eastern and southern Bantu languages but seems to be an areal feature shared by most if not all West-Coastal Bantu languages (Grégoire 1993, Hadermann 1996, Bostoen and Mundeke 2011, 2012; De Kind 2014; Koni Muluwa and Bostoen 2014; Bostoen and Koni Muluwa 2021). I show that in Kukuya, this IBV position is productively exploited, the element that is placed in the IBV position must be itself focal or part of a larger focal constituent.

The chapter is organised as follows: section 3.1 introduces the canon-

ical word order SVO and various types of focus that this word order can express; section 3.2 is dedicated to illustrating different functions of the IBV focus position and interpretations associated with it, as well as some morphological and tonal variation related to this position; section 3.3 introduces the expressions of (multiple) topical elements in the preverbal domain, and functional passive constructions that combine the use of IBV focus position and topic fronting; and section 3.4 turns to different types of cleft constructions.

3.1 Canonical word order

In this section I present the canonical SVO word order in Kukuya and show that SVO can be used to express various types of focus such as argument and adjunct focus, VP focus and truth focus. The SVO order is also the most common way of expressing athetic sentence.

3.1.1 SVO as canonical word order

When talking about word order, it can be sometimes problematic to generalise what the “canonical” word order is in a language, depending on different criteria and discourse types. Synchronic variation within the language can also provide different clues on its canonical word order. Here I follow the criterion that the canonical word order of a language is commonly reflected in a “topic-comment articulation” where the subject of the sentence has a discourse function of topic representing presupposed or given information, and the rest of the sentence expresses new information (Lambrecht 1994; Andrews 2007). Thus the canonical word order is expected to show up in the answer to a question such as “What did s/he do” which places focus on the predicate, namely the VP. In (1) we see that to answer such a question, the SVO order is usually attested, which has been reported to be the canonical word order of most other Bantu languages (Bearth 2003; van der Wal 2015; Downing and Marten 2019 among others).

- (1) a. Mu-kái kí-má ká-sí?
 1-woman 7-what 1SM.PST-do.PST
 ‘What did the woman do?’
 b. Ndé á-búnum-i baa-ntsúú.
 1.PRO 1SM.PST-feed-PST 2-chicken
 ‘She fed the chicken.’

However, as we will see shortly, the felicitous answer to a VP question is not restricted to SVO, but can also be SOV as shown in (2), although the occurrence of SVO for VP focus largely surpasses that of SOV in my corpus and is always the first intuition of the speakers. Based on these facts, I assume that the canonical word order of Kukuya, if there is one, should be SVO, which is also the most common and frequent word order attested in a this language. We will also see later in this chapter that any deviation of the SVO order, to a larger or smaller extent, involves some discourse-related manipulations. The SOV expressing VP focus is possibly used to mark contrast on the VP (see section 5.2.2), or it is in the process of being further grammaticalised from a more marked focus construction to a pragmatically neutral word order secondary to the canonical SVO.

- (2) (*visual stimulus: what are the two women doing?*)
 Bó ntáli bá-kâ-yílik-a.
 2.PRO 9.bed 2SM-IMPF-clear.up-FV
 ‘They are making the bed.’

The canonical position of different kinds of adjuncts is usually postverbal and after the object(s) in a transitive construction, as illustrated in (3). Here I refer to adjuncts as adverbial phrases that add extra information (temporal, locative, manner) to the sentence, which is distinguished from adverbs which modify the verb. From (3) we can also see that in a ditransitive construction in Kukuya, the recipient object always precedes the theme. Example (4) shows that it is ungrammatical to place the locative phrase between the verb and the object, even if this adjunct is in focus. This also indicates that in Kukuya there is no IAV focus position which is well-known in many other Bantu languages. I thus conclude that the canonical constituent order in Kukuya is Subject-Verb-Object-Adjuncts.

- (3) Nkaaká á-wî baa-ndzulí bvi-kídzá mu
 1.grandmother 1SM.PST-give.PST 2-cat 8-food 18.LOC
 nkunkólo yi.
 9.evening 9.DEM.I
 ‘The grandmother gave the cats food this evening.’
- (4) (*answer to “where did you see Gilbert?”*)
- a. *Me á-mún-i ku dzáandu Gilbert.
 1SG.PRO PST-1SG.SM.see-PST 17.LOC 5.market Gilbert
Int: ‘I saw Gilbert at the market.’
- b. Me á-mún-i Gilbert ku dzáandu.
 1SG.PRO PST-1SG.SM.see-PST Gilbert 17.LOC 5.market
 ‘I saw Gilbert at the market.’
- c. Me ku dzáandu á-mún-i Gilbert.
 1SG.PRO 17.LOC 5.market PST-1SG.SM.see-PST Gilbert
 ‘I saw Gilbert at the market.’

3.1.2 Focus expressions in SVO

As said above, the canonical SVO word order is usually captured when the whole VP is in focus. In this subsection I will show that SVO can also be used to express term focus (focus on an argument or adjunct or a subpart of these), and in fact all types of arguments and adjuncts can be focused in their canonical linear position. In addition, SVO can also express different types of predicate-centered focus (PCF) such as verb focus and truth value focus. I will discuss them in turn.

A *wh*-element, which is usually considered to be inherently focused, as well as its corresponding answer, are commonly seen as reliable diagnostics for focus expressions (Rooth 1992; Lambrecht 1994; Krifka 2007; van der Wal 2016). In Kukuya, an object can be focused in its canonical postverbal position. In (5a) we see that the answer to an object *wh*-object question can be SVO with the object being focused in its canonical position. We find in (5b) that this question can also be answered in an SOV order with the

- (8) *(The speaker thinks that it is impossible for the person to have seen Alain.)*

Ndé á-mún-i Alain **ku-ní**?
 1.PRO 1SM.PST-see-PST Alain I7-which
 ‘He saw Alain, (but) where?’

An element modified by “only” is always associated with an exhaustive focus reading. In (9a,b) we see that to place an object DP modified by “only” in its canonical postverbal position and in the IBV position are both grammatical. This shows that even exhaustive focus on the object can be expressed in the canonical SVO order.

- (9) a. Mu-loí á-wí báana **wúna** maa-nkúru.
 1-teacher 1SM.PST-give.PST 2.children only 6-pen
 ‘The teacher gave the children only pens.’
 b. Nkaaká **wúna** mvá ká-wí bú-ká.
 1.grandmother only 1.dog 1SM.PST-give.PST 14-cassava
 ‘The grandmother gave only the dog cassava.’

A subject can also be focused preverbally in SVO, as shown in the question-answer pair in (10) and the preverbal subject modified by “only” in (11). The availability of preverbal subject focus in Kukuya is somehow exceptional considering the rigid constraint against the preverbal subject to be focal in many other Bantu languages (Morimoto 2000; Zerbian 2006; van der Wal 2009, 2015; Downing and Marten 2019). It should be noted here that I am not claiming that the preverbal subject is structurally focused *in situ*, but in a position which is structurally different from the canonical subject position, though under both circumstances the linear word order is SVO. In section 5.2.1.2, I will distinguish the preverbal topical subject position from the focal subject in IBV position. Here I say that a topical and a focal subject, which structurally occupy different positions, overlap in their linear position in the preverbal domain.

- (10) a. Kí-má kíí-súruk-i?
 7-what 7SM.PST-fall-PST
 ‘What fell down?’
 b. Mpúku á-súruk-i.
 1.rat 1SM.PST-fall-PST
 ‘A/The rat fell down.’
- (11) Wúna baa-ntsúú báá-ból-i.
 only 2-chicken 2SM.PST-decompose-PST
 ‘Only the chickens got bad.’

SVO is also compatible with various types of predicate-centered focus (PCF) (Güldemann 2003, 2009) including verb focus (state-of-affairs focus) and truth focus (also known as verum focus). In (12b) SVO as the answer to the question in (12a) expresses focus on the verb and SVO is used. In (13b) SVO is used to correct the truth value of (13a), thus we see that SVO is also felicitous to express truth focus.

- (12) a. Taará mi-féme kí-má ké ká-sí?
 1.father 4-pig 7-what 7.PRO 1SM.PST-do.PST
 ‘What did father do to the pigs?’
 b. Ndé á-dzwí mi-féme.
 1.PRO 1SM.PST-kill.PST 4-pig
 ‘He killed the pigs.’
- (13) a. Gilbert ka-ká-bvúúr-í we mi-pará ni?
 Gilbert NEG-1SM.PST-return-PST 2SG.PRO 4-money NEG
 ‘Gilbert did not return you the money?’
 b. Ndé á-bvúur-i me mi-pará.
 1.PRO 1SM.PST-return-PST 1SG.PRO 4-money
 ‘He did return me the money’

In the above, I have shown that SVO can have different uses in terms of information structure. According to my intuition, SVO is most commonly used

as a “topic-comment structure” where the subject functions as the topic and the whole VP is focused, and it can also express term focus and different types of PCF. There is no constraint against preverbal subject focus.

3.1.3 Thetic sentences

In this subsection, I will show that SVO can also be used as a thetic sentence. A thetic sentence is used to present all the information that the sentence carries in one piece, as opposed to a “categorical” sentence in which the topic and comment can be further divided (Kuroda 1972; Sasse 1987, 1996). The thetic sentence is also referred to as “all-new” or “all focus” utterance (van der Wal 2021). The answer to a question such as “what happens” can thus be used to investigate the formation of a thetic sentence, as this type of question often does not presuppose a topical referent and requires information on the whole event. In Kukuya, a thetic sentence usually surfaces in SVO. As shown in (14), to answer the question “what happened outside”, only (14a) with SVO is felicitous, while any deviation of this word order cannot be an appropriate answer. The answer in (14b) is only felicitous when *mwáana* “child” has been already mentioned and is what the speakers are talking about. In (15) the preverbal subject is indefinite and non-specific, which are not characteristics of topic. From the context we see that it expresses a thetic meaning as there is no old information presupposed. Here we see that the distinction on definiteness of the preverbal subject may help discern a thetic SVO sentence from a categorical one.

(14) (*What happened outside?*)

- a. Mvá á-bvî ku ntsá dzuná.
 1.dog 1SM.PST-fall.PST 17.LOC inside 5.hole
 ‘A dog fell into the hole.’
- b. #Mwáana taará á-béer-i.
 1.child 1.father 1SM.PST-beat-PST
 ‘The child was beaten by father.’

- (15) (*You returned home and found some footprints on the floor, you said to your roommate:*)

Mbuurú á-yení.

1.person 1SM.PST-come.PST

'Someone came.'

We have already seen above that the preverbal subject in SVO can be topical or focal. The availability of SVO to express thethetic meaning shows that the preverbal subject can also be non-topical (and non-focal), since there is no distinction on topic and comment in athetic sentence. However, an answer to a question “what happened” may also contain a topic expression due to the tendency to “accommodate information” in a sentence (Lewis 1979; Stalnaker 2002; von Fintel 2008; van der Wal 2016). Even if there is no identifiable topical referent in the common ground before the discourse starts, the interlocutors tend to accept the referent that occurs at the beginning of the dialogue, for example the “dog” in (14a) above, to ensure a coherent communication. In this regard, SVO may never be really “thetic”, but the preverbal subject can function as an “immediate topic” that can always rescue the discourse from not having a topic.

A question that asks about the reason may also have athetic answer, as the reason may not contain any presupposed information known by the addressee. In the examples (16) and (17) below, a subject relative construction is used to answer this kind of *why*-questions. At first glance, the subject relative clause looks like a dedicated strategy to express athetic meaning, just as in French a cleft can be used in athetic sentence in which the subject is detopicalised by relativisation. However, since the *wh*-word for “why” is formed as *mu kima* which literally means “for what”, the relativisation of the subject is more likely to nominalise the whole sentence to congruently answer the question, as the *why*-question is actually a *what*-question which targets at a nominal. For example in (16) the question literally means “for what are the children afraid” which may target a certain object that causes the fear rather than a whole event, therefore the answer is interpreted as “for the crocodile that is walking in the yard”.

- (16) a. Báana mu **ki-má** bá-li ya buokó?
 2.children 18.LOC 7-what 2SM-COP with 14.fear
 ‘Why are the children afraid?’
 b. Mu-ɲaaní **wu-kâ-dzîe** ɲa kalá mbalí.
 1.crocodile 1REL-IMPF-walk 16.LOC inside 9.yard
 ‘A crocodile that is walking in the yard.’
- (17) a. Mu-kái mu **ki-má** ké ká-mal-í ɲa
 1-woman 18.LOC 7-what 7.PRO 1SM.PST-leave-PST 16.LOC
 nzó?
 9.house
 ‘Why did the woman leave home?’
 b. Mwáana aa ndé **wũ-dzínim-i**.
 1.child 1.CONN 1.PRO 1REL-disappear-PST
 ‘Her son who disappeared.’

In (18) we see that the *mu* always requires a nominal or nominalised element following it. To answer the question in the context, only (18a) is grammatical as the reason is nominalised thus can be selected by *mu*, while (18b) the sentence is ungrammatical due to the fact that the construction after *mu* is still clausal rather than nominal. From (18a) we also see that the relativisation of the subject can nominalise the clause after *mu* as a whole, since the child becomes happy not because of the “father” himself, but of the *fact* that “father bought him a small goat”. In this sense, the relativisation strategy above in (16) and (17) can also be considered as nominalising the whole clause for the sake of expressing the information as *one piece* (Sasse 1987).

- (18) (*Why is the child happy?*)
 a. Mwáana li yă **kí-sáábí** mu taará **wũ-fúum-i**
 1.child COP with 7-happiness 18.LOC 1.father 1REL-buy-PST
 ndé ntaba.
 1.PRO 1.goat
 ‘The child is happy that father bought him a goat.’

- b. *Mwáana li yǎ kí-sáábí mu taará
1.child COP with 7-happiness 18.LOC 1.father
á-fúum-i ndé ntaba.
1SM.PST-buy-PST 1.PRO 1.goat
Int: 'The child is happy that father bought him a goat.'

In summary, a thetic meaning is commonly expressed by the canonical SVO word order in Kukuya. A subject relativisation strategy may also be used to express thetic meaning, the motivation of which seems to be nominalising the whole information in the sentence as one chunk. However, the use of relativisation may also be due to the fact that some interrogative words tend to require a nominal answer.

In the previous sections, I have shown that the canonical word order SVO is compatible with different information structural constructions. It can be used to express VP focus, term focus, predicated-centered focus and as well as theticity. In the next section I will introduce how a deviation of this canonical word order, namely the use of the dedicated IBV focus position, is associated with information structure.

TOP	TOP	FOC	Verb
S _{TOP}			V
		S _{FOC}	V
O _{TOP}	(O _{TOP})	S _{FOC}	V
S _{TOP}	(O _{TOP})	O _{FOC}	V

Table 3.1: Linear slots of the preverbal domain in Kukuya

3.2 Dedicated IBV focus position

This section gives an overview on the availability and interpretation of the dedicated immediate-before verb (IBV) focus position in Kukuya. I show that the IBV position is available for arguments including subject and object, for adjuncts and even for the infinitive predicate to get focused. Compared to the *in situ* focus strategy introduced above, the element placed in IBV often has an identificational focus reading in which a referent is identified in an existential presupposition. The interrogative words and contrastively focused elements more strictly occur in the IBV position than other focal elements. The focal interpretation can project from the IBV position to the whole VP.

An immediate question here is how to define the “IBV” position in this language. Throughout the chapter, the notion “IBV position” refers to a particular *structural* position, whether occupied or empty, that is adjacent to the verb and no other constituent can intervene in between. Linearly, a topical or non-focal grammatical subject in SVO can also occur immediately left-adjacent to the verb, but it is not structurally placed in the IBV position, because other constituents can still be inserted between it and the verb (see section 3.3.1). When the IBV focus position is not filled, a topical or non-focal subject just linearly overlaps the IBV position. If we consider all possible elements in the preverbal domain to occur in different slots which correspond to different structural positions (but not necessarily linear position), which is illustrated in Table 3.1, we see that the IBV focus position can be clearly discerned. In this table, each line represents a particular construction that

will be introduced in the rest of the chapter.

3.2.1 Argument and adjunct focus in IBV

3.2.1.1 Object and adjunct focus in IBV

In Kukuya, a *wh*-element is usually placed in the IBV position. As shown in (19a,b), the *wh*-objects *kímá* “what” and *ná* “who” are both placed in the IBV position. In (19c) we see that the *wh*-word must be strictly adjacent to the verb and the intervention of another element turns the sentence ungrammatical. Example (19d) shows that to place the *wh*-object *kímá* “what” in its canonical postverbal position is also ungrammatical.

- (19) a. Mvá **kí-má** ká-siib-i?
 1.dog 7-what 1SM.PST-catch-PST
 ‘What did the dog catch?’
- b. Taará **ná** ká-mún-í ku mu-súru?
 1.father 1.who 1SM.PST-see-PST 17.LOC 3-forest
 ‘Who did father see in the forest?’
- c. *Taará **ná** ku mu-súru ká-mún-í?
 1.father 1.who 17.LOC 3-forest 1SM.PST-see-PST
Int: ‘Who did father see in the forest?’
- d. *Mvá á-siib-i **ki-ma**?
 1.dog 1SM.PST-catch-PST 7-what
Int: ‘What did the dog catch?’

In ditransitive constructions (20a,b), the recipient and the patient objects are questioned in the IBV position respectively, while the other non-focal objects also tend to occur in the preverbal domain preceding the IBV position. We see from these examples that a *wh*-object is strictly placed in the IBV position. We can also see in (20) that there is an agreeing pronoun following the *wh*-word. I leave the function of the pronoun to be discussed later

and assume the *wh*-word and the pronoun to form one inseparable constituent in the IBV position.

- (20) a. Nkaaká ma-désu ná ndé ká-wî?
 1-granny 6-bean 1.who 1.PRO 1SM.PST-give.PST
 ‘To whom did grandmother give the beans?’
 b. Nkaaká mvá kí-má ké ká-wî?
 1-granny 1.dog 7-what 7.PRO 1SM.PST-give.PST
 ‘What did grandmother give to the dog?’

The answer to an object *wh*-question also tends to occur in the IBV position, though it is not restricted to this position. As introduced in section 2 and also as in (21), we see that the answer to a *wh*-object question can be SVO and SOV, with the focal answer being either in IBV or its canonical postverbal position. Here we see that both preverbal and postverbal focus strategies are available in Kukuya. In the elicitation of question-answer pairs, I had a strong impression that when I put emphatic intonation on the focal answer in French, the speakers were more likely to use the preverbal focus strategy in the corresponding translation. I will discuss the interpretational differences of the IBV and *in situ* object focus strategies in section 3.2.4.

- (21) (*What did mother buy yesterday?*)
 a. Ndé á-fúum-i mu-ngwa.
 1.PRO 1SM.PST-buy-PST 3-salt
 ‘She bought some SALT.’
 b. Ndé mú-ngwa ká-fúum-i.
 1.PRO 3-salt 1SM.PST-buy-PST
 ‘She bought some SALT.’

In an alternative question that asks for a choice or preference, as well as in its corresponding answer, the IBV focus strategy is always used as shown in (22) and (23). SVO is viewed as infelicitous as in (23c). Here we see that when some (at least one) alternative is explicitly mentioned in the context, SOV must be used for exclusion and identification.

- (22) a. Maamá lóoso ká-télek-i wó bú-ka?
 1.mother 5.rice 1SM.PST-prepare-PST or 14-cassava
 ‘Did mother cook the rice or the cassava?’
 b. Ndé bú-ka ká-télek-i.
 1.PRO 14-cassava 1SM.PST-prepare-PST
 ‘She prepared the cassava.’
- (23) a. We báa-ntsúú kâ-dzií kí-dzá wó kí-wáli?
 2SG.PRO 2-chicken 2SG.IMPF-like INF-eat or 7-duck
 ‘Do you like to eat chicken or duck?’
 b. Me kí-wál-í kâ-n-dzií kí-dzá.
 1SG.PRO 7-duck IMPF-1SG.SM-like INF-eat
 ‘I like to eat duck.’
 c. #Me kâ-n-dzií kí-dzá kí-wáli.
 1SG.PRO IMPF-1SG.SM-like INF-eat 7-duck
 ‘I like to eat duck.’

The *wh*-adjuncts such as *munkí* ‘when’, *kuní* ‘where’, *buní* ‘how’ and *mu kimá* ‘why’, as inherently focal, are also most commonly placed in the IBV position, as shown in (24)-(26). As for the answer to a *wh*-adjunct question, the focused adjunct in the answer can be either in IBV or its base position, as illustrated in (24).

- (24) a. Mwáana **munkí** ká-dzí ntsúí?
 1.child when 1SM.PST-eat.PST 1.fish
 ‘When did the child eat the fish?’
 b. Ndé ntsúí **mu ngwaalí** ká-dzí.
 1.PRO 1.fish 18.LOC 9.morning 1SM.PST-eat.PST
 ‘S/He ate the fish in the MORNING.’
 c. Ndé á-dzí ntsúí **mu ngwaalí**.
 1.PRO 1SM.PST-eat.PST 1.fish 18.LOC 9.morning
 ‘S/He ate the fish in the MORNING.’

- (25) Li-dzwá nyama wúa, biáwe ndé **ku-ní**
 1PL.SM-kill 1.animal 1.DEM.II 1PL.PRO 1.PRO 17-which
 líi-kab-a?
 1PL.FUT-share-FV
 ‘(As) we kill that animal, where will we share it?’
- (26) a. Mwáana ki-yinga **bu-ní** kíi-wir-i?
 1.child 7-festival 14-which 7SM.PST-pass-PST
 ‘How did the child pass the festival?’
 b. Ndé ki-yinga **kí-bvé** kíi-wir-i.
 1.PRO 7-festival 7-good 7SM.PST-pass-PST
 ‘S/He passed the festival WELL/HAPPILY.’

Some *wh*-adjuncts, in particular the manner and reason interrogatives *buní* ‘how’ and *mu kimá* ‘why’, are also attested to occur in their canonical postverbal position as in (27), in free variation with their occurrence in the IBV position, without triggering interpretational differences. Some *wh*-adjuncts occur in the canonical postverbal position in the context of a rhetorical question, see example (8) above.

- (27) Ngo ká-kwí **mu** **ki-ma**?
 1.leopard 1SM.PST-die.PST 18.LOC 7-what
 ‘Why did the leopard die?’

An element modified by ‘only’ is always associated with an exhaustive focus reading. In (9) above we have already seen that to place an object DP modified by ‘only’ in the IBV position and the canonical postverbal position are both grammatical. This shows that exclusive focus is not necessarily expressed via the IBV position. Example (28) shows that when excluding some alternatives in an explicitly mentioned set, the exclusively focused phrase can either occur in IBV or its canonical position. In spontaneous speech, I also found both postverbal and preverbal distribution of the *only*-phrases, although the occurrence in the IBV position is more often attested.

- (28) a. Mu-kái á-fíuum-i ntaba yǎ má-sáani?
 1-woman 1SM.PST-buy-PST 1.goat with 6-plate
 ‘Did the woman buy a goat and some plates?’
- b. Ndé wúna ma-sáani ká-fíuum-i.
 1.PRO only 6-plate 1SM.PST-buy-PST
 ‘She only bought some plates.’
- c. Ndé á-fíuum-i wúna ma-sáani.
 1.PRO 1SM.PST-buy-PST only 6-plate
 ‘She only bought some plates.’

Interestingly, where I do find restrictions on the position of phrases with exclusive (exhaustive) focus is in yes-no questions. As (29) shows, in a yes-no question the “*only*”-phrase can only occur in the IBV position but is judged to be ungrammatical in the postverbal position. This restriction does not hold when there is no “*only*”-phrase in the sentence; both SOV and SVO are felicitous to form a yes-no question in that case.

- (29) a. Taará **wúna** ma-sáani ká-swaak-í?
 1.father only 6-plate 1SM.PST-wash-PST
 ‘Did father only wash the plates?’
- b. *Taará á-swaak-í **wúna** ma-sáani?
 1.father 1SM.PST-wash-PST only 6-plate
Int: ‘Did father only wash the plates?’

Contrastively focused objects and adjuncts also commonly occur in the IBV position, with rare exceptions. The postverbal locative phrase and the object in statement (30a) are corrected respectively in (30b) and (30c) in the IBV position, while correcting them postverbally is degraded. In (31) the instrumental phrase is also corrected in the IBV position. In (30b) and (31b) we also notice that the focal element in IBV can be preceded by multiple non-focal elements which can be subject, object and adjunct. I will return to discuss this in section 3.3.1.

- (30) a. Ngajwâ maamá ká-wéek-i mu-nkáání ku
 9.truth 1.mother 1SM.PST-send-PST 3-letter 17.LOC
 Djambala?
 Djambala
 ‘Was it true that mother sent the letter to Djambala?’
- b. Ambú, ndé mu-nkáání mfaí ká-wéek-i.
 no 1.PRO 3-letter 3.capital 1SM.PST-send-PST
 ‘No, she sent the letter to BRAZZAVILLE.’
- c. Ambú, ndé kí-dzídzilá ká-wéek-i.
 no 1.PRO 7-parcel 1SM.PST-send-PST
 ‘No, she sent a PARCEL.’
- (31) a. Ki-yélé kii-nyánim-i kii-mbúli mu míaka.
 7-hare 7SM.PST-save-PST 7-lion 18.LOC 4.hand
 ‘The hare saved the lion by hand.’
- b. Ambú, ndé kii-mbúli mu mu-siá káá-nyánim-i.
 no 1.PRO 7-lion 18.LOC 3-rope 1SM.PST-save-PST
 ‘No, he saved the lion with a ROPE.’

Some additional examples on contrastive focus with clear context are illustrated in (32) and (33). Example (32) is felicitous in the context when you did not feed the chicken and went out with your wife, when you returned home, you found that the chickens were full and there were beans on the ground. Your wife did not notice the beans and asked “did someone feed the chickens with rice?” and you corrected her with this sentence. The speaker also suggested some possible context for (32) and (33) as shown in the brackets. From these examples we see that when displacing an element in the IBV position from its canonical position, a set of alternatives is at least implicitly available from the context.

- (32) (*There were bags of beans and rice, you found that the chickens were full and only the beans were reduced.*)

Mbuurú baa-ntsúú **má-désu** ká-búnum-i.

1.person 2-chicken 6-bean 1SM.PST-feed-PST

'The person/Someone fed the chickens the BEANS.'

- (33) (*You see that the child is sitting on the ground and crying, your friend asks from some distance away "did something hurt the child's legs?", and you correct her/him.*)

Ki-lóko mwáana **mú-tswé** kíí-búl-i.

7-thing 1.child 3-head 7SM.PST-hurt-PST

'The (particular) thing hurt the child's HEAD.'

An interim generalisation here is that a focal object or adjunct can be either placed in IBV position or its canonical postverbal position, while some types of foci such as *wh*-words and contrastively focused elements particularly favour the IBV position. This is in line with the idea that specific types of focus and different "degrees" of contrast can be syntactically identified (Cruschina 2021, a.o.). The IBV position, as the more marked focus position than the canonical position in terms of word order, is reserved for higher degree of contrast while the canonical postverbal position may encode less or no contrast. I will show in section 3.2.4 that the IBV position usually expresses identificational focus, while assertively focused elements tend to stay in their canonical positions.

3.2.1.2 Subject focus in IBV

Subject focus in Kukuya can be expressed in three ways, namely in the canonical SVO order, by an OSV order or by using a pseudo-cleft construction. We will see shortly that the former two means should be considered as different realisations of the same IBV subject focus strategy.

First, to question a subject, the pseudo-cleft construction seems to be the most widely used in my corpus, and an example is shown (34a-c). In

these constructions, the subordinated clause is a relative clause with a covert head, and the predicative focused subject DP occur sentence-finally. A copula linking the relative part and the predicative DP is only visible in negative context as in (34c), in which the subject marking on the copula is by default the class 7 subject marker *kí-*.

- (34) a. *Ki-kí-túm-í mbaá ki-namá kí-ma?*
 7REL-7SM-cause-PST 9.fire INF-burn 7-what
 ‘What caused the fire?’
- b. *Wũ-fúum-i ma-li taará.*
 1REL-buy-PST 6-wine 1.father
 ‘(The one) who bought the wine is father.’
- c. *Wũ-dzí baa-ntsúú ka-kí-li mvá ni.*
 1REL-eat.PST 2-chicken NEG-7SM-COP 1.dog NEG
 ‘(The one) who ate the chicken was not the dog.’

An alternative strategy to focus the subject is to place the subject in the IBV position, as the *wh*-word in (35), the subject modified by “only” in (36) and the answer to a subject question in (37). It is noteworthy that the subject is focused in the IBV position which is structurally different from its canonical preverbal position. When a focused subject appears preverbally, no other element can occur between this subject and the verb as shown in (35b), which is not characteristic of the topical subject, therefore the focused subject must be placed in a different structural position, which is the IBV.

- (35) a. *Ná á-ték-i mu-ngwa?*
 1.who 1SM.PST-sell-PST 3-salt
 ‘Who sold the salt?’
- b. **Ná mú-ngwa ká-ték-i?*
 1.who 3-salt 1SM.PST-sell-PST
Int: ‘Who sold the salt?’

- (36) *Wúna baa-ntsúú báá-ból-i.*
 only 2-chicken 2SM.PST-rot-PST
 ‘Only the chicken rotted.’

- (37) (“Who gave the child the oranges?”)
Bí-búru **bíí-wí** **mwáana** **ma-láara**.
 8-parent 8SM.PST-give.PST 1.child 6-orange
 ‘The PARENTS gave the child the oranges.’

Intriguingly, an exception to the legitimacy of a preverbal focal subject is the *which*-phrase. According to many speakers, a *which*-phrase cannot be placed in IBV position in the same way as other *wh*-phrases, but can only occur in a reverse pseudo-cleft sentence. We see in (38) and (39) that the *whose*-phrase and *which*-phrase respectively are not compatible with canonical subject marking, which indicates that they cannot function as the grammatical subject of the sentence, but can only occur in a pseudo-cleft construction.

- (38) a. ***Mu-káli** **wuu ná á-níak-i** **mwáana?**
 1-wife 1.CONN 1.who 1SM.PST-abandon-PST 1.child
Int.: ‘Whose wife abandoned the child?’
 b. **Mu-káli** **wuu ná wǔ-níak-i** **mwáana?**
 1-wife 1.CONN 1.who 1REL-abandon-PST 1.child
 ‘Whose wife abandoned the child?’
- (39) a. ***Mwáana** **wu-ní á-mún-i** **Zacharie?**
 1.child 1-which 1SM.PST-see-PST Zacharie
Int.: ‘Which child saw Zacharie?’
 b. **Mwáana** **wu-ní wǔ-mún-i** **Zacharie?**
 1.child 1-which 1REL-see-PST Zacharie
lit.: ‘Which child is the one who saw Zacharie?’

A *which*-phrase is usually considered to be discourse-linked and presupposes an antecedent in the given discourse, thus does not necessarily trigger discourse-new information (Şener 2010). On the opposite, a non-discourse linked *wh*-phrase does not presuppose an antecedent and always functions as a focal phrase. As we will see throughout this chapter that the preverbal domain is available for elements of various information structural status,

it thus seems unexpected that the D-linked *which-* and *whose-* phrases are not compatible with preverbal focus in a mono-clausal construction. One possible motivation for the D-linked interrogatives to prefer a cleft construction may be that the presupposed existence makes the question as selective, which patterns with the pseudo-cleft construction. Here again it shows that different types and “degrees” of contrast and focus may be encoded through different grammatical strategies.

There are some interpretational differences between the pseudo-cleft construction and subject focus in IBV. In (40a) the subject of the embedded clause is questioned in a pseudo-cleft, and in (40b) it is questioned in the IBV position. According to the speakers, (40a) is used in the context where there is a presupposed set of candidates who killed the king, which means the speaker has already a group of suspects; while in (40b) there is no candidate invoked in the speaker’s mind. In this sense the pseudo-cleft construction is more discourse-linked than the IBV subject focus strategy.

- (40) a. Ndé kâ-tsuomó ndíri [wũ-dzwí mu-kóko na].
 1.PRO 1SM.IMPf-think 1.COMP 1REL-kill.PST 1-king 1.who
 ‘S/He is thinking about who killed the king.’
- b. Ndé kâ-tsuomó ndíri [ná á-dzwí mu-kóko].
 1.PRO 1SM.IMPf-think 1.COMP 1.who 1SM.PST-kill.PST 1-king
 ‘S/He is thinking about who killed the king.’

Subject focus in the IBV position is very commonly accompanied by the fronting of topical object(s) to the preverbal domain. For example the answer to a subject question in (37) above can be alternatively expressed as in (41) in which the subject is focused in the IBV position and the given objects are all preposed to the preverbal domain, surfacing an OOSV order. Similarly in (42), both in the question and the answer, the focused subject is placed in the IBV position with some object being fronted to the preverbal domain. The OSV order is in fact very commonly attested in Kukuya for expressing subject focus, and we will also see in section 4.2 that OSV can function as an equivalent of the passive construction. I assume that this OSV construction is not a third strategy for subject focus, but it is just a different realisation of the IBV focus strategy introduced above, and here in the OSV construction

some element must be marked salient as topic of the sentence and fronted to the initial position. The preposing of other preverbal constituents can also help to identify the subject as being placed in the IBV focus position.

- (41) (“Who gave the child the oranges?”)
 Mwáana ma-láara **bí-búru** bíí-wí.
 1.child 6-orange 8-parent 8SM.PST-give.PST
 ‘The child was given the oranges by the PARENTS.’
- (42) a. Taará téme **ná** á-sonom-i?
 1.father 5.hoe 1.who 1SM.PST-lend-PST
 ‘Who lent father the hoe?’
 b. Téme **nkaaká** á-sonom-í taará.
 5.hoe 1.grandmother 1SM.PST-lend-PST 1.father
 ‘Grandmother lent father the hoe.’

I have shown in the previous subsection that focal objects and adjuncts can occur either in the IBV position or their canonical postverbal position. Here we see that the subject can also be focused in the IBV position, and we can generalise that all arguments and adjuncts in Kukuya can be focused in this verb-adjacent IBV position. Next I will investigate whether the IBV position can be used to express focus on an element which is structurally smaller than the argument/adjunct, namely a modifier, or on a larger constituent such as the VP.

3.2.1.3 Sub-NP focus

In this subsection I show that sub-NP focus can also be expressed by placing the NP in the IBV position, as it can express focus on a modifier. In example (43) we see that the interrogative quantifier *kwê* “how many” occurs in the IBV position, following the NP that it modifies. From this example it is not clear whether it is the whole DP including the head noun and the quantifier that is placed in the IBV position, or solely the quantifier occupies the IBV position.

- (43) Ba-nziá ma-tsúku kwê bâ-sá ɲa ntsá
 2-foreigner 6-day how.many 2SM.FUT-stay 16.LOC inside
 bu-lá ba?
 14-village 14.DEM.I
 ‘How long will the foreigners stay in this village?’

The sentences in (44) were elicited with a picture in which a woman is holding three knives in her hand. In (44a) the numeral quantifier is correctively focused and takes the H tone prefix, while the nominal prefix of the head NP keeps the L tone; in (44b) it is only the head NP that is focal and takes the H tone prefix, with the quantifier following it; (44c) conveys focus on the whole quantified NP, and in this case the H tone prefix only appears on the head NP.

- (44) (*visual stimuli: the woman is holding three knives in her hand.*)
- a. (*Is the woman holding TWO knives?*)
 Ndé maa-mbhiélé **má**-tíri kâ-kwaal-a.
 1.PRO 6-knife 6-three 1SM.IMPF-hold-FV
 ‘She is holding THREE knives.’
- b. (*Is the woman holding three SPOONS?*)
 Ndé **máa**-mbhiélé **ma**-tíri kâ-kwaal-a.
 1.PRO 6-knife 6-three 1SM.IMPF-hold-FV
 ‘She is holding three KNIVES.’
- c. (*Is the woman holding TWO SPOONS?/What is the woman holding?*)
 Ndé **máa**-mbhiélé ma-tíri kâ-kwaal-a.
 1.PRO 6-knife 6-three 1SM.IMPF-hold-FV
 ‘She is holding THREE KNIVES.’

From these examples we see that only one H tone prefix can occur on the preverbal elements, either on an NP or its modifier, and it is only when the modifier itself is focal that it can take the H tone prefix. One question that arises from (44) is whether the H tone prefix always aligns with focus and with the IBV position. Here I suppose that the head NP and the modifier are

separated in (44a), the head NP functions as a dislocated topic and only the modifier is focused in IBV, since the head NP in this case can also be elided; in (44b) and (44c) the head NP and the modifier are one constituent and only the prefix on the head NP can take the H tone which maps onto the focus reading and the IBV position.

So far in this section I have shown that the IBV focus position is available for argument focus including subject and object, adjunct focus as well as sub-NP focus on a modifier. In the next section I will look into how the IBV position is exploited beyond term focus, namely in predicate-centered focus.

3.2.2 Predicate(-centered) focus and IBV

3.2.2.1 VP focus and verb focus

In this section I refer to predicate focus as focus on the whole verb phrase. Predicate-centered focus (PCF), as defined in Gildemann (2003), indicates the focus on *part* of the predicate and can be further divided into state-of-affairs focus which is also referred to as verb focus, tense/aspect/mood (TAM) focus, and truth value focus (also known as verum focus); the latter two are also referred to as operator focus.

As introduced in section 2, VP focus in Kukuya is most commonly expressed via SVO. Example (45) is extracted from a written translation task done by the speakers in which they gave most answers to the VP questions in SVO. However, the IBV focus position can also be employed when answering a VP question, as shown (46c). In elicitation the speakers usually cannot explain the interpretational differences between SVO and SOV when expressing VP focus, while there are in fact more pragmatic restrictions for the SOV order to occur, as I will introduce shortly.

- (45) a. Huguette bu-kía kí-má kâ-sá?
 Huguette 14-tomorrow 7-what 1SM.FUT-do
 ‘What will Huguette do tomorrow?’
 b. Bu-kía, ndé â-yé kîe baa-ndúku.
 14-tomorrow 1.PRO 1SM.FUT-go visit 2-friend
 ‘Tomorrow she will go to visit friends.’
- (46) a. Taará kí-má ké ká-sí ŋa ngwaalí?
 1.father 7-what 7.PRO 1SM.PST-do.PST 16.LOC 9.morning
 ‘What did father do in the morning?’
 b. Ndé á-dzwí mi-féme.
 1.PRO 1SM.PST-kill.PST 4-pig
 ‘He killed some pigs.’
 c. Ndé mí-féme ká-dzwí.
 1.PRO 4-pig 1SM.PST-kill.PST
 ‘He killed some pigs.’

The availability of SOV to express VP focus could be viewed as a counter-argument for the IBV focus position, as in this case the focus is not only on the element immediate before the verb, but is on the whole predicate that contains the IBV element. I propose that we do not need to reject the hypothesis of IBV as a focus position but can revise the hypothesis to say that the IBV element should at least be part of the focus, and the IBV element as the *nucleus* of the focus set can project up to the whole verb phrase, which depends on the discourse context (Selkirk 1995: 555; Reinhart 2006; van der Wal 2009: 241).

An alternative question on the VP and its congruent answer can both be expressed by SOV as shown in (47). The yes-no question in (48a) focuses the whole VP and also uses SOV, here we see that SVO in (48b) and SOV in (48c) can both be felicitous additive responses to this question, which again shows that both SVO and SOV can signal focus on the whole VP.

- (47) a. We má-sáání á-swaakí wó bi-báa-wî
 2SG.PRO 6-plate 2SG.PST-wash-PST or 8REL-2SM.PST-give.PST
 we á-sî?
 2SG.PRO 2SG.PST-do.PST
 ‘Did you wash the plates or do your homework?’
- b. Me má-sáání á-n-swaak-í.
 1SG.PRO 6-plate PST-1SG.SM-wash-PST
 ‘I washed the plates.’
- (48) a. Ndé wúna bi-ko ká-swaak-í?
 1.PRO only 8-clothes 1SM.PST-wash-PST
 ‘Did he only wash the clothes?’
- b. Ndé á-búnum-i bii-ndomó hé.
 1.PRO 1SM.PST-feed-PST 8-goat also
 ‘He also fed the goats.’
- c. Ndé bíi-ndomó hé ká-búnum-i (hé).
 1.PRO 8-goat also 1SM.PST-feed-PST also
 ‘He also fed the goats.’

Verb focus, also known as state-of-affairs focus which locates focus on the lexical value of the verb, can be expressed in different ways in Kukuya. As mentioned in section 3.1.2, SVO can be used to signal verb focus. The answer to a question like “what did X do to Y?” can be used to diagnose verb focus expressions, in which the subject and the object are both topical since they are already given in the background and the focus is on the verb itself. Interestingly, we see in (49) and (50) that SVO and SOV can both signal verb focus, while OSV with the subject in IBV cannot be used as a felicitous answer.

- (49) a. Ngolo Marie kí-má ká-sî?
 Ngolo Marie 7-what 1SM.PST-do.PST
 ‘What did Ngolo do to Marie?’

- b. Ngolo á-pfur-í Marie.
Ngolo 1SM.PST-cheat-PST Marie
'Ngolo betrayed Marie.'
- c. Ngolo Marie ká-pfur-í.
Ngolo Marie 1SM.PST-cheat-PST
'Ngolo betrayed Marie.'
- d. #Marie Ngolo á-pfur-í.
Marie Ngolo 1SM.PST-cheat-PST
Int: 'Ngolo betrayed Marie.'

- (50)
- a. Ngúku baa-ntaba kí-má ká-sí?
1.mother 2-goat 7-what 1SM.PST-do.PST
'What did mother do to the goats?'
 - b. Ngúku á-dzwí baa-ntaba.
1.mother 1SM.PST-kill.PST 2-goat
'Mother killed the goats.'
 - c. Ngúku báa-ntaba ká-dzwí.
1.mother 2-goat 1SM.PST-kill.PST
'Mother killed the goats.'
 - d. #Baa-ntaba ngúku á-dzwí.
2-goat 1.mother 1SM.PST-kill.PST
Int: 'Mother killed the goats.'

Here it is somehow problematic to explain why SOV is applicable to express verb focus, if we hypothesise the IBV to be a dedicated focus position from which the focus can project up to the whole VP, we still cannot account for why the focus on the IBV can be "transferred" to the verb. According to the focus projection hypothesis above, the object being placed in the IBV position is consistent with the whole VP being in focus, since the object is counted as within the scope of the VP focus. Here we may wonder whether the preposed objects in (49c) and (50c) indeed occupy the IBV position or they are just fronted as some topical elements. The most obvious evidence that they are placed in the IBV position rather than some higher positions lies in the H tone on the nominal prefix in (50c). Since this H tone marking

only occurs when the preposed element is in IBV, here we can confirm that SOV is indeed felicitous to express verb focus with the IBV position being occupied. In (49d) and (50d) we see that the answers become infelicitous when the subject is placed in IBV, this may be accounted for by an economy principle. Since verb focus here must involve something to be placed in the IBV position and both the subject and object in this question are topical, it may be easier to just place the object to the IBV position, rather than place the subject in the IBV while also topicalising the object.

Let us consider some more examples of verb focus. In the answers to the question in (51a), some other actions taken on the object “pig” are introduced in addition to just “washing” it, so the verb “to kill” in (51b-d) itself is focused. We see that (51b) is felicitous with SVO, while (51c) with SOV is infelicitous here with the additive particle *hé* and it only implies that the grandmother must have killed other animals beforehand, thus the focus can only be on the object rather than the verb. SOV in (51d) without the additive particle is felicitous to correctively focus the truth value of the verb, and in (51e) SOV with the additive particle is felicitous when the whole VP is focused as an additive action that is not related to the pig.

- (51) a. Nkaaká á-swaak-í mu-féme?
 1.grandmother 1SM.PST-wash-PST 4-pig
 ‘Did grandmother wash the pig?’
- b. Ndé á-dzwí hé mú-féme.
 1.PRO 1SM.PST-kill.PST also 4-pig
 ‘She also KILLED the pig.’
- c. #Ndé mú-fémé (hé) ká-dzwí (hé).
 1.PRO 4-pig also 1SM.PST-kill.PST also
Int: ‘She also KILLED the pig.’
 ‘She also killed the PIG.’
- d. Ndé mú-fémé ká-dzwí.
 1.PRO 4-pig 1SM.PST-kill.PST
 ‘She KILLED the pig.’

- e. Ndé báa-ntsúú hé ká-ká-i.
 1.PRO 2-chicken also 1SM.PST-grill-PST
 ‘She also GRILLED THE CHICKEN.’

The infelicity in (51c) above is unexpected given that we have already seen examples above in which the SOV word order can be used to express verb focus. I suggest that this infelicity is due to the presence of the additive particle *hé* which is always associated with the focal element in a sentence. This particle may disambiguate the nucleus of the focus from the domain to which it may project up, thus in (51c) it is more intuitive for the speakers to interpret the focus on the object only. In the absence of this additive particle, the SOV word order becomes a possible way of expressing verb focus as in (51d). We also see that in (51e) the presence of *hé* does not prevent the SOV word order from expressing VP focus, this may be explained by the fact that the whole VP in (51e) is new, so the VP focus reading can be rescued from the intervention of the additive particle, thus it can be an appropriate answer to (51a).

From above I have shown that when the IBV position is occupied by an object DP, it can be used to express VP focus and verb focus. Next I present another strategy for expressing predicate-centered focus, which also involves the use of the IBV focus position, namely the predicate doubling construction.

3.2.2.2 Predicate doubling

Predicate doubling is first documented by Meeussen (1967: 121) as the “advance verb construction” that can express truth focus, intensity and concession. In many other Bantu languages, predicate doubling is a common strategy to express state-of-affairs focus and truth focus on the verb, and is reported to be situated in different stages in the grammaticalisation path to the progressive and future tense (Güldemann et al. 2010, 2014; Morimoto 2016). In some neighbouring languages of Teke, such as in the Kikongo group of Zone H and other Zone B languages, some of which also favour the IBV focus position, the predicate doubling construction is also

well attested expressing verb focus and truth focus, as well as progressive and future tense (Hadermann 1996; De Kind 2014; De Kind et al. 2015; Güldemann and Fiedler 2022). Some examples from these languages illustrate the phenomenon in (52)-(55) below.

- (52) Ku-tá:nga ndyeká-tá:nga.
 INF-read 1SG:FUT-read
 ‘I will READ.’ [Suundi H31b] (Hadermann 1996: 161) [verb focus]
- (53) Mona mbwene N-kenda za zula ki-ame
 INF.see 1SG.see.PERF 10-affliction 10.GEN 7.people 7-1SG.POSS
 kina.
 7.DEM
 ‘I have surely seen the affliction of that people of mine there.’
 [Ndibu H16] (De Kind et al. 2015: 12) [truth focus]
- (54) Ba-ka:só bá-ná; vádó bâ:vádə pénda.
 2-woman 2-DEM INF.cultivate 2-cultivate groundnut
 ‘These women, they are cultivating groundnuts.’
 [Nzebi B52] (Hadermann 1996: 162) [progressive]
- (55) Vuumbuka yi-vuumbuka.
 INF-dress 1SG.SM-dress
 ‘I’ll dress myself.’ [Yaka H33] (De Kind et al. 2015: 36) [future]

In this subsection I introduce the predicate doubling construction in Kukuya. In Kukuya, the predicate doubling construction is mainly attested as IBV doubling, while topic doubling is judged to be infelicitous and cleft doubling to be quite marginal. In (56), to exclusively focus the lexical value of a verb while excluding some alternatives, we see that SOV can be appropriately used in (56b) with the exclusive focus particle *wína* preceding the preposed object, while to place the particle immediately in front of the verb in either SVO or SOV is judged to be ungrammatical as shown in (56c, d). It seems that the exclusive particle can only modify nominal elements or a VP

but not a bare verb. In (56e) we see that verb focus can also be expressed by placing an infinitive form of the verb immediately before the inflected verb.

- (56) a. Ngúku á-télek-i bu-ká á-dzí?
 1.mother 1SM.PST-prepare-PST 14-cassava 1SM.PST-eat.PST
 ‘Did mother prepare and eat the cassava?’
- b. Ndé wúna bu-ká ká-télek-i.
 1.PRO only 14-cassava 1SM.PST-prepare-PST
 ‘She only PREPARED the cassava.’
- c. ??Ndé wúna á-télek-i bu-ká.
 1.PRO only 1SM.PST-prepare-PST 14-cassava
Int: ‘She only PREPARED the cassava.’
- d. *Ndé bu-ká wúna ká-télek-i.
 1.PRO 14-cassava only 1SM.PST-prepare-PST
Int: ‘She only PREPARED the cassava.’
- e. Ndé bu-ká wúna ki-téléké ká-télek-i.
 1.PRO 14-cassava only INF-prepare 1SM.PST-prepare-PST
 ‘She only PREPARED the cassava.’

One additional example of predicate doubling expressing verb focus in Kukuya is given in (57). There is an important interpretational difference between the use of SOV and predicate doubling in expressing verb focus: while (56b, e) and (57b) all express exclusive focus on the verb, (56b) with SOV indicates that the event is completed and the mother only prepared the cassava but does not need to go on making it, while (56e) and (57b) imply that the event is still continuing and there must be other things that need to be done with the cassava and the goats.

- (57) a. Maamá á-dzwî baa-ntabá á-ték-i.
 1.mother 1SM.PST-kill.PST 2-goat 1SM.PST-sell-PST
 ‘Mother killed the goats (and) sold (them).’

- b. Ambú, ndé bó wúna ki-téké káa-ték-i.
 no 1.PRO 2.PRO only INF-sell LSM.PST-sell-PST
 'No, she only SOLD them.'

Truth focus on the verb can be expressed neither by predicate doubling nor by SOV in Kukuya. In (58) and (59) we see that to correct a negative truth value on the verb, there is no other construction than the canonical SVO, and the speakers tend to put some intonational emphasis on the verb to express the truth focus. SOV in (58c) and (59c) is infelicitous here, while it can actually express focus on the object or the VP or the lexical value of the verb. The predicate doubling in (58d) and (59d) is also infelicitous and implies that there are other actions that need to be done with the oranges, expressing verb focus. We see that both the SOV order and predicate doubling can trigger alternatives either on the object or on the verb and imply a contrast with other actions or tasks that remain to be done.

- (58) a. Taará ka-ká-kí ma-láala ni?
 1.father NEG-ISM.PST-pick.PST 6-orange NEG
 'Did father not pick the oranges?'
 b. Ndé á-kí ma-láala.
 1.PRO 1SM.PST-pick.PST 6-orange
 'He DID pick the oranges.'
 c. #Ndé má-láálá ká-kí.
 1.PRO 6-orange 1SM.PST-pick.PST
Int: 'He DID pick the oranges.'
 d. #Ndé ma-láala ká-ká ká-kí.
 1.PRO 6-orange INF-pick 1SM.PST-pick.PST
Int: 'He DID pick the oranges.'
- (59) a. Ndé ka-ká-bvúúr-í we mi-pará ni?
 1.PRO NEG-ISM.PST-return-PST 2SG.PRO 4-money NEG
 'Didn't s/he return you the money?'

- b. Ndé á-bvúur-i me mi-pará.
 1.PRO LSM.PST-return-PST 1SG.PRO 4-money
 ‘S/He DID return me the money’
- c. #Ndé me mí-para ká-bvúur-i.
 1.PRO 1SG.PRO 4-money LSM.PST-return-PST
Int: ‘He DID return me the money.’
- d. #Ndé me mi-pará kí-bvúúrá ká-bvúur-i.
 1.PRO 1SG.PRO 4-money INF-return LSM.PST-return-PST
Int: ‘He DID return me the money.’

The predicate doubling construction that expresses verb focus in Kukuya looks quite like the IBV focus construction that encodes narrow focus on the preposed DP, and predicate doubling in this case is just a particular realisation of the IBV focus, in which the predicate is doubled as an infinitive form and is focused in the IBV position. In this sense the predicate doubling and the SOV/OSV orders are actually the same structure that places focus in the IBV position, which is also consistent with the fact that infinitives are also DPs in Kukuya and most other Bantu languages. If this is true, we may expect the fronted infinitive and a preverbal *wh*-word to be in complementary distribution as they should compete for the unique IBV position, and this is borne out as the ungrammaticality in (60) and (61), which also shows that there is only one preverbal focus site in this language. In these examples the predicate doubling is intended to be used for expressing a *progressive* meaning, which I will introduce shortly.

- (60) *Ndé kí-má kí-dzá kâ-dzá?
 1.PRO 7-what INF-eat LSM.IMPF-eat
Int: ‘What is he/she eating?’

- (61) *Ná kí-tsúka kâ-tsúka?
 1.who INF-speak LSM.IMPF-speak
Int: ‘Who is talking?’

We have seen above that the IBV position is associated with argument and adjunct focus, as well as VP focus and verb focus. It is not clear here whether

the infinitive in the predicate doubling construction should be viewed as an argument of the verb, if so, the predicate doubling is analogous to term focus on an argument DP. In fact, predicate doubling and term focus in the IBV position have some important interpretational similarities: predicate doubling usually implies the potential occurrence of other actions, while term focus in IBV also hints that some alternatives should be available for the proposition. I will discuss more on these interpretational properties in section 3.2.4.

Similar to many other Bantu languages, predicate doubling in Kukuya can express progressive aspect. In examples (62) and (63) the fronted infinitive expresses a neutral progressive meaning without focusing on the verb itself. Verb focus and progressive reading are often said to have a close semantic and pragmatic relation and the progressive is considered to be an inherently focused verb category in which the “ongoing nature of the event described by the verb” constitutes the focus domain of the sentence (Hyman and Watters 1984; Güldemann 2003; De Kind 2014; De Kind et al. 2015). The predicate doubling with progressive reading is sometimes ambiguous and can only be distinguished from PCF focus through the pragmatic context. Example (62) can be a felicitous corrective response to focus on the progressive aspect expressing TAM focus, while predicate doubling in (63) is used outside the PCF focus context. In Kukuya there is a dedicated aspect marker *-ká-* that can mark habitual as well as progressive aspect without the fronting of an infinitive verb, so the predicate doubling is not the only way of expressing progressive in Kukuya.

(62) *(Have they already eaten?)*

Bó kí-dzá bá-kâ-dzá.
 2.PRO INF-eat 2SM-PROG-eat
 ‘They are eating.’

(63) Mwáana wu-kái wu-kíi-kwî ngúku á-yiká kí-líla
 1.child 1-female 1REL-7SM.PST-die.PST 1.mother 1SM-IMPF INF-cry
 kâ-líl-a.
 1SM.PROG-cry-FV
 ‘The girl whose mother died is crying.’

In Kukuya grammar as well as in many other Teke languages, the expression of immediate future tense also involves the SOV order, as in (64). The predicate doubling construction in Kukuya can also have the immediate future reading, as shown in (65).

- (64) Bó má-ko báa-fíuum-a.
 2.PRO 6-banana 2SM.FUT-buy-FV
 ‘They’ll buy some bananas.’
- (65) a. We ka-á-bvúúr-í ndé mi-pará ní?
 2SG.PRO NEG-2SG.PST-return-PST 1.PRO 4-money NEG
 ‘You did not return her the money?’
 b. Me mi-pará kí-bvúúrá kâ-n-bvúur-a.
 1SG.PRO 4-money INF-return IMPF-1SG.SM-return-FV
 ‘I am (surely) going to return the money.’

The response in (65b) has a truth focus reading, meaning that the speaker will definitely return the money and does not imply that there are other things to be done with the object “money”, which differs from the interpretation in example (59d) above. The contrast between (59d) and (65b) is that, in (59d) the alternative could be “borrow again” the money in addition to just returning it, while in (65b) the alternative is “not to return” the money as opposed to returning it. We see that when predicate doubling has an immediate future reading, it can express truth value focus, which may be due to the SOV order being grammaticalised to express certain tense, thus becoming pragmatically equal to the canonical word order; for the predicate doubling in other tenses, it cannot be used to express truth focus.

From the above presentation on the expressions of different types of predicate(-centered) focus, we see that in Kukuya VP focus (predicate focus) can be expressed by SVO as well as SOV. The use of the IBV focus position to express VP focus can be explained by the focus projection account. Verb focus (state-of-affairs focus) can also be realised via SVO and SOV, while OSV cannot express verb focus. It remains to be investigated why the IBV focus position is also involved in expressing verb focus. The

predicate doubling construction is used mostly to express verb focus and usually triggers alternatives to the verb, while truth focus is commonly expressed by the canonical SVO order.

3.2.3 IBV as a dedicated focus position

So far we have encountered and discussed many examples which suggest that the IBV position is always associated to some type of focus. In this subsection I will investigate some intrinsic properties of the IBV position. First I will provide more tests on whether the IBV position is really a dedicated focus position. Then I will discuss the interpretational differences between the IBV focus strategy and *in situ* focus that was introduced in the previous section, showing that the IBV position is reserved for expressing identificational focus, while the *in situ* focus strategy seems to be more frequently used to express assertive focus.

Many Bantu languages have been reported to have a dedicated focus position, most of which are the so-called immediate-after-verb (IAV) position that are commonly attested in languages such as Aghem (Watters 1979, Hyman and Polinsky 2010), Bemba (Costa and Kula 2008), Matengo (Yoneda 2011), Makhuwa (van der Wal 2009), and Zulu (Buell 2009). The immediate-before-verb (IBV) focus position is much more rarely attested only in some West-Coastal Bantu languages (WCB), which has already been described in detail for Mbuun (B87, Bostoen and Mundeke 2011, 2012), Nsong (B85d, Koni Muluwa and Bostoen 2019) and in the Kikongo cluster (Hadermann 1996; De Kind 2014; De Kind et al. 2015). Here I provide some more evidence to show that the IBV in Kukuya is indeed a dedicated focus position.

At this moment, we consider first non-subject elements (the IBV and the canonical subject position will be disentangled later). We see that if some element is placed in the IBV position, it must be focal or at least within the scope of focus, while elements in other positions cannot be focal at the same time. This is illustrated in (66) and (67). In (66b) we see that a *wh*-element cannot co-occur with another element being placed in

the IBV position, the ungrammaticality can only be explained by the focal status of the adverb in IBV and the generalisation that multiple foci are not allowed; (66c) is not a felicitous answer to (66a), as what is placed in the IBV position is a manner adverb but not the object which is the target of the question; (66c) can only be an appropriate answer to the question “HOW did the person eat the cassava”, which indicates that the adverb in the IBV must be focal. Similarly in (67), only (67b) can be a felicitous answer to (67a) while (67c) can only be the answer to the question that asks for the location. From these examples we see that if there are multiple preverbal elements (in an affirmative sentence), the IBV slot must be occupied by a focal element, while other elements in the sentence are prohibited to be focused.

- (66) a. Mbuurú **kí-má** ká-dzí tswáatswáa?
 1.person 7-what 1SM.PST-eat.PST fast
 ‘What did the person eat quickly?’
- b. *Mbuurú kí-má tswáatswáa ká-dzí?
 1.person 7-what fast 1SM.PST-eat.PST
Int: ‘What did the person eat quickly?’
- c. #Mbuurú bu-ka tswáatswáa ká-dzí.
 1.person 14-cassava fast 1SM.PST-eat.PST
 ‘The person ate the cassava QUICKLY.’
- (67) a. We ná á-mún-i ku dzáandu?
 2SG.PRO who 2SG.PST-see-PST 17.LOC 5.market
 ‘Who did you see at the market?’
- b. Me Gilbert á-mún-i ku dzáandu.
 1SG.PRO Gilbert 1SG.PST-see-PST 17.LOC 5.market
 ‘I saw Gilbert at the market.’
- c. #Me Gilbert ku dzáandu á-mún-i.
 1SG.PRO Gilbert 17.LOC 5.market 1SG.PST-see-PST
 ‘I saw Gilbert AT THE MARKET.’

Neither SOV nor OSV can be used to answer the question such as “What happened” as in (68a) which requires athetic answer in which the whole

utterance provides the new information thus no topic or focus is subdivided in the sentence (Kuroda 1972; Sasse 1987, 1996; also see section 2.3). Here we see that only SVO in (68b) can be felicitous. The answers in (68c) and (68d) are both inappropriate here, as there must be some focal reading triggered by the IBV position being occupied, namely the object “child” in (68c) and the subject “father” in (68d), thus they are both incompatible with thethetic requirement.

- (68) a. Me a-n-yúk-i nkelé ku mbali, kí-má
 1SG.PRO PST-1SG.SM-hear-FV 9.noise 17.LOC 9.outside 7-what
 kí-sî?
 7SM.PST-do.PST
 ‘I heard some noise outside, what happened?’
- b. Taará á-béer-i mwáana.
 1.father 1SM.PST-beat-PST 1.child
 ‘Father beat the child.’
- c. #Taará mwáána ká-béer-i.
 1.father 1.child 1SM.PST-beat-PST
 ‘Father beat the CHILD.’
- d. #Mwáána taará á-béer-i.
 1.child 1.father 1SM.PST-beat-PST
 ‘FATHER beat the child.’

Idiom tests can also help to justify that the IBV position is indeed associated with focus function (van der Wal 2016, 2021). In idiom sentences, the idiomatic reading arises as a whole chunk thus is considered to be non-compositional. We would predict that any part of an idiom sentence cannot be focused, since no expressions in an idiom refers to something that is accessible in the reality, thus no alternatives can be triggered for focus. Examples (69)-(71) illustrate several idioms in Kukuya, and all these idiom sentences surface as SVO, which is further evidence for SVO as the canonical order. Crucially, we find that when the word order is shifted to SOV, the sentence is still grammatical but the idiomatic reading is not retained, and the sentence can only have the literal meaning. These idiom tests show that the formation of SOV must have involved some discourse-related operations,

namely the IBV element must be focal, as the translations indicate.

- (69) a. Ndé á-tín-i ko li-búi.
 1.PRO 1SM.PST-pick-PST 5.banana 5-immature
 ‘S/He had a sexual relation with a child.’
lit: ‘S/He picked the unripe banana.’
- b. Ndé ko li-búi ká-tín-i.
 1.PRO 5.banana 5-immature 1SM.PST-pick-PST
 *‘S/He had a sexual relation with a child.’
 ‘S/He picked the unripe banana.’
- (70) a. Me a-n-dzwî ntaalí mu kíí.
 1SG.PRO PST-1SG.SM-pick.PST 1.snake 18.LOC 7.pipe
 ‘I have lost all.’
lit: ‘I killed a snake with a pipe.’
- b. Me ntaalí mu kíí n-dzwî.
 1SG.PRO 1.snake 18.LOC 7.pipe PST-1SG.SM-pick.PST
 *‘I have lost all.’
 ‘I killed a snake with a PIPE.’
- (71) a. Maa-nkala máá-dzí mbúlu.
 6-charcoal 6SM.PST-eat.PST 9.blanket
 ‘The problem becomes burning (rather than coldness when you use too much charcoal).’
lit: ‘The coal is eating the blanket.’
- b. Maa-nkala mbúlú máá-dzí.
 6-charcoal 9.blanket 6SM.PST-eat.PST
 *‘The problem becomes burning (rather than coldness when you use too much charcoal).’
 ‘The coal is eating the BLANKET.’

So far we have seen that any non-subject constituent that is in IBV position must be interpreted as focused; now I provide examples to illustrate that it

is also a dedicated focus position for a subject. Example (72) is partially repeated from (35) above, in which we see that the interrogative subject in (72a) seemingly occupies the same near position as the grammatical subject in canonical word order; however from (72b) and (72c) we see that the focal and topical subjects are subject to different constraints on their linear position: the focal subject can only occur in the IBV position but cannot be followed by other elements in the preverbal domain as in (72b), while the topical subject can be followed by other DPs, such as by focused object in the IBV position in (72c). In other words, the focal subject has an IBV requirement while the topical subject does not, therefore they must stay in different structural positions. Similarly in (73), we see that the answer to a subject question must be adjacent to the verb as in (73a,c) and another DP cannot intervene as in (73b).

- (72) a. **Ná** á-ték-i mu-ngwa?
 1.who 1SM.PST-sell-PST 3-salt
 ‘Who sold the salt?’
- b. ***Ná** mú-ngwa ká-ték-i?
 1.who 3-salt 1SM.PST-sell-PST
Int: ‘Who sold the salt?’
- c. (*What did the grandmother sell?*)
Nkaaká mú-ngwa ká-ték-i?
 1.grandmother 3-salt 1SM.PST-sell-PST
 ‘The grandmother sold some SALT.’

- (73) (*Who brought the dog?*)
- a. **Taará** á-yi-í mvá.
 1.father 1SM.PST-bring-PST 1.dog
 ‘FATHER brought the dog.’
- b. #**Taará** mvá ká-yi-í.
 1.father 1.dog 1SM.PST-bring-PST
Int: ‘FATHER brought the dog.’

- c. Mvá taará á-yi-í.
 1.dog 1.father 1SM.PST-bring-PST
 ‘FATHER brought the dog.’

In (74a) the subject precedes the negative marker on the verb, and the focus is on the polarity of the sentence rather than on the subject; while (74b) expresses constituent negation and the subject is somehow “inserted” between the negative marker and the verb, providing evidence that it must be situated in a different position than the subject in (74a). The interpretation in (74b) is that it is not “father” but someone else that killed the leopard, so the focus is apparently on the subject. From the minimal pair in (74) the canonical subject position and the IBV can be distinguished: in (74a) the subject appears in the canonical subject position, while in (74b) the subject is placed in the IBV position. The position of the negative marker here may also support that the IBV position is indeed “immediate” before the verb, since when the IBV slot is empty as in (74a), the negative morpheme is always prefixed to the verb and prosodically phrased together with it.

- (74) a. Ngo taará ka-ká-dzwí ni.
 1.leopard 1.father NEG-1SM.PST-kill.PST NEG
 ‘The leopard, father did NOT kill (it).’
 b. Ngo ka taará á-dzwí ni.
 1.leopard NEG 1.father 1SM.PST-kill.PST NEG
 ‘The leopard was not killed by father (but by someone else).’

The analysis above has provided strong evidence on the presence of a dedicated focus position in Kukuya, i.e. everything that is in this position is focused, and this position is located immediately left-adjacent to the verb. I have also shown that the subject is focused in the IBV position which is distinct from its canonical preverbal position. When this IBV position is filled, the sentence must have undergone some discourse-related operations for information packaging, in most cases some argument or adjunct gets focused. Recall that in section 3.1.2 we have seen that an element such as a postverbal object can also be focused in its canonical postverbal position,

next I will discuss the distinction between IBV focus and non-IBV *in situ* focus with regard to their interpretation.

3.2.4 Interpretational properties of IBV focus

In the introduction on the expressions of term focus, we have already noticed that in some examples the IBV focus strategy is preferred over the *in situ* focus strategy: in the answer to an alternative question; in a contrastive focus expression; in the predicate doubling construction; and in most of the SOV sentences with a clear context in which some overt alternatives are available for the focused element. In this subsection I show more details on the interpretational distinction between IBV focus and the non-IBV focus strategies, arguing that the IBV position is usually, if not in all cases, used to express identificational focus. Here I refer to the identificational focus as a focus type that identifies a referent in an existential presupposition. For example in English sentence “What I like is sunshine”, where the presupposition is that there is something that I like and this something is identified as sunshine. The concept of identificational focus is also used as a hypernym of contrastive and exhaustive foci.

I begin with comparing a minimal sentence pair that only differs in the order of the constituents. Both the sentences in (75a,b) can be felicitous answers to the *wh*-question “what did father eat?”, while they differ in interpretation as shown in the contexts. Example (75a) with SOV word order is used to identify exactly what father ate; while (75b) with SVO word order is used just to provide some new information. For example (75a) the speakers clearly told me that there must be some alternatives invoked in mind and you want to identify what exactly the correct answer is.

- (75) a. (*There were many dishes and in fact father ate only some fish, and you may suspect him to have eaten something else.*)

Taará baa-ntsúí ká-dzí.
1.father 2-fish 1SM.PST-eat.PST

‘The father ate some FISH.’

- b. (*There were some fish and the father ate them all, and you just wanted to know what father ate.*)

Taará á-dzí **baa-ntsúi.**

1.father 1SM.PST-eat.PST 2-fish

'The father ate some FISH.'

The same distinction is attested in (76), which is a sentence extracted from a written task done by two speakers. One speaker was asked to write a letter in Kukuya to another speaker, and at the beginning of the letter after some greetings, the speaker asked the other if he saw me, using the SVO sentence in (76a). When I asked them if this sentence can be replaced by SOV as (76b), both of them judged it as infelicitous, saying that (76b) is used only when the speaker thought the other had seen someone and wanted to know who exactly he saw. From this minimal pair we see again that SOV is used for identification and SVO simply provides new information.

- (76) (*At the beginning of a letter: 'How are you? Did you see Zhen yesterday?'*...)

a. We á-mún-i Zhen?

2SG.PRO 2SG.PST-see-PST Zhen

'Did you see Zhen?'

b. #We Zhen á-mún-i?

2SG.PRO Zhen 2SG.PST-see-PST

'Did you see ZHEN?'

Example (77a) is used when someone is asking about your profession. Here an identificational reading can also be deduced, since a person's career is usually the regular and unique activity that s/he is involved, and (77a) implies that the speaker lives by only selling goats but not other animals; (77b) is used in case where the speaker has a farm and s/he is just telling the others what s/he sells, in which the goats are not necessarily the only animal that the speaker sells.

- (77) a. Me **báa-ntabá** **kâ-n-téke**.
 ISG.PRO 2-goat IMPF-1SG.SM-sell
 ‘I sell goats.’
- b. Me **kâ-n-téké** **báa-ntabá**.
 ISG.PRO IMPF-1SG.SM-sell 2-goat
 ‘I sell goats.’

The sentences in (78) intend to express focus on the subject and are both felicitous as answer to a subject question. While (78a) is used when “you see that child crying and you want to know whether the father or the mother beat the child”, (78b) according to the speakers can also mean “it is father but not someone else that beat the child”, here it seems that both subjects may have been placed in the same IBV position.

- (78) a. (*You see that child crying and you want to know whether the father or the mother beat the child.*)
Mwáana taará á-béer-i.
 1.child 1.father 1SM.PST-beat-PST
 ‘The child is beaten by FATHER.’
- b. (*It is father but not someone else that beat the child.*)
Taará á-béer-i mwáana.
 1.father 1SM.PST-beat-PST 1.child
 ‘Father beat the child.’

The sentences in (79) are examples of the construction that functions as an equivalent of the passive in Kukuya, which I will introduce in the next section. (79a) is used in the context where you discovered the theft and were worrying about your things to have been all stolen, after checking you found that only the necklace was missing; while (79b) is used when simply telling a truth that the thieves had come and a necklace was stolen.

- (79) a. **Mú-dzirá báá-túr-i**.
 3-necklace 2SM.PST-steal-PST
 ‘The necklace was stolen.’

- b. B^áa-t^úr-i mu-dzirá.
 2SM.PST-steal-PST 3-necklace
 ‘They stole the necklace. (The necklace was stolen)’

Some more evidence comes from the interpretation on the word *mbuurú* ‘person’ in different positions, which is inspired by the same test used for diagnosing exclusive focus in van der Wal (2016). In Kukuya, the expression *mbuurú* can have the reading ‘person’ or ‘someone/anyone’, which depends on the context. In (80a) when *mbuurú* is placed in the IBV position, it can only have a generic reading as ‘human-being’ that contrasts with an animal; while in (80b) *mbuurú* can have either the reading ‘someone’ or ‘person’. The generic reading in (80a) is consistent with the hypothesis we make here on the IBV position being an identificational focus position (van der Wal 2016, 2020). The reading of ‘someone’ is indefinite so is never identifiable, while the reading ‘person’ can only be identified when contrasted with ‘non-person’, namely the animals.

- (80) a. Ngo mbuurú ká-dzí.
 1.leopard 1.person 1SM.PST-eat.PST
 ‘The leopard ate a PERSON (not an animal).’
 b. Ngo á-dzí mbuurú.
 1.leopard 1SM.PST-eat.PST 1.person
 ‘The leopard ate someone/the person/a person.’

The generic reading is also attested in OSV as in (81) that expresses focus on the subject. Example (81) is used when you saw a dead leopard and you were wondering how the leopard died until you found an arrow on its body which indicated that it was killed by a human. Notice here that the ‘person’ reading, though it can be definite referring to a given person, can only show the contrast between ‘this person’ and ‘that person’ when demonstrative modifiers are present, thus in (80a) above and (81) the *mbuurú* in the IBV position cannot express contrast between different ‘persons’ but can only have the generic reading.

- (81) (*You saw a dead leopard and you were wondering how the leopard died until you found an arrow on its body.*)

Ngo **mbuurú** á-dzwî.
 1.leopard 1.person 1SM.PST-kill.PST

'The leopard was killed by a PERSON.'

Another crucial piece of evidence supporting the IBV position as an identificational focus site lies in the negation strategy on the focal elements. In examples (82) and (83) we see that to negate the element in the IBV position, the often omitted copula can somehow "show up" with the negative marker and precede the IBV item. Example (82) means that the gecko was not eaten by the dog but by some other animals, and the negation targets only the subject (dog) and does not negate the action/sentence. Example (83) means that father bought some other things instead of the bed. Given that the copula has an identifying function, its being placed immediately before the IBV focused element suggests that the IBV element is identificationally focused. The possible presence of the copula can also provide evidence on the origin and the nature of the IBV position, namely its connection with the cleft construction that is dedicated for identification and specification, which will be discussed in the next chapter.

- (82) Ngwangúlu ka-kí-li **mvá** á-dzí ni.
 1.gecko NEG-7SM-COP 1.dog 1SM.PST-eat.PST NEG

'The gecko was not eaten by the DOG.'

- (83) Taará ka(-kí-li) **ntáli** ká-sí me ni.
 1.father NEG-7SM-COP 9.bed 1SM.PST-make.PST 1SG.PRO NEG

'Father did not make a BED for me.'

Although there is much evidence on the identificational nature of the IBV position, this position is not necessarily a dedicated exclusive focus position, from which we may see the difference between identificational and exclusive focus in this language. If the IBV position is used to express exclusive focus, we expect that an element modified by the strong quantifiers

“every/each” and “all” should be incompatible with the IBV position, since a DP modified by these quantifiers is not exclusive (É. Kiss 1998; van der Wal 2009, 2011, 2016). However, an *every*-phrase can occur in the IBV position as shown in (84a), and to specify a set of alternatives such as “every chicken” to contrast with “every fish” in this example is possible but is judged to be unnecessary according to the speakers. The context of (84a) can be either “there were several species of fish and you tasted each” or “there were many dishes and you only tasted each fish but not other meat”. In (84b) we see that a DP modified by the universal quantifier “all” is also compatible with the IBV position, and here again to explicitly specify the alternatives such as “all the cakes” to show a contrast is possible but not necessary.

- (84) a. Me ná ntsúí á-n-dziin-i.
 ISG.PRO every 1.fish PST-1SG.SM-taste-PST
 ‘I tasted each fish.’
- b. Me báa-ntsúí bhoî á-n-dziin-i.
 ISG.PRO 2-fish 2.all PST-1SG.SM-taste-PST
 ‘I tasted all the fish.’

In (85) we see that a DP modified by a scalar additive particle “even” that does not exclude the alternatives can occur at IBV. In (86) the reply to an incomplete question with the additive particle “also” can surface in the SOV order, which again indicates that the IBV position is not necessarily an exclusive focus position.

- (85) (*There is a lazy boy who never did any housework but today he has washed many things, the clothes, the curtains, the plates, and...*)
 Ndé ntswê ki-tséké kíí me ká-swaak-í.
 1.PRO even 7-hat 7.CONN ISG.PRO ISM.PST-wash-PST
 ‘He even washed my hat.’

- (86) (*Did Gilbert wash the clothes?*)
 Ndé **bí-ko** ká-swaak-í, ndé hé **má-saani**
 1.PRO 8-clothes 1SM.PST-wash-PST 1.PRO also 6-plate
 ká-swaak-í.
 1SM.PST-wash-PST
 ‘He washed the clothes, and he also washed the plates.’

There are also some counterarguments against the IBV position as being an identificational focus site. The first puzzle that remains to be explained is what we have already seen above: since the *wh*-words show the strongest tendency to be placed in the IBV position and if this preference is related to the identificational nature for most *wh*-questions, it is unexpected that SOV and SVO are both acceptable as the answer, if only the IBV position is employed for identificational focus.

Moreover, if the IBV position is identificational in nature which must have a presupposition of existence, a question with the IBV *wh*-phrase cannot have an empty set answer, since the existence of a possible candidate is contained in the presupposition. In (87) we see that the *wh*-question can be answered by “nobody”, which indicates that there is no presupposition in the question thus it is not necessarily identificational focus.

- (87) a. We **ná** á-mún-i ku mu-súru?
 2SG.PRO 1.who 2SG.PST-see-PST 17.LOC 3-forest
 ‘Who did you see the the forest?’
 b. Mbuurú ni.
 1.person NEG
 ‘Nobody.’

For these counterarguments against the IBV to be an identificational focus position, I will leave them open for now. The assumption is that, at an earlier stage the IBV position was indeed innovated for the sake of expressing identificational focus, which can be deduced from its possible origin from a cleft construction that I will discuss in the next chapter, but synchronically not all the uses of IBV position in all contexts are necessarily identificational,

and in fact the IBV position has been observed to be in a further grammaticalisation process to become pragmatically neutral.

* * *

In this section I have introduced some syntactic and interpretational properties of the IBV focus position in Kukuya. I argue that the IBV position is a dedicated focus position which is structurally different from the canonical subject in the SVO order. I have shown that the IBV focus position is available for argument focus including subject and object, adjunct focus, sub-NP focus on a modifier as well as various types of predicate-centered focus such as VP focus and verb focus. The element that is placed in the IBV position can be an argument NP, an adjunct PP, or an infinitive verb in the predicate doubling construction. While focus can also be expressed postverbally for non-subject constituents, IBV focus tend to have an identificational reading but in some contexts it becomes pragmatically neutral. After investigating focus expressions in this language, in the next section I introduce topic expressions.

3.3 Topical elements in the preverbal domain

In Bantu languages and in general cross-linguistically, topical elements show the general tendency to occur in the left periphery or the preverbal domain of the sentence (Gundel 1988; Henderson 2006; van der Wal 2009; 2015; Kerr et al. 2023). Likewise, the topical elements in Kukuya also tend to occur in the preverbal domain. In this section I first introduce that in Kukuya there are multiple types of topical elements and they all tend to occur in the preverbal domain. As illustrated above, there is a dedicated IBV focus position in this language, and in fact this IBV position can also interact with topical expressions. We will see in this section that in many sentences in which the IBV focus slot is occupied, all other non-focal elements tend to occur in the left periphery preceding the IBV slot, leaving the verb to the right boundary of the clause. Then I present two specific constructions that can function as the equivalent of a passive, namely the OSV and the impersonal *ba-* constructions, which can functionally compensate the absence of morphological passive marking in this language.

3.3.1 Multiple topics in the preverbal domain

I start from classifying different types of topical elements in this language. According to different syntactic and interpretational properties, at least four types of topical elements can be distinguished, which are the topical subject, the topical object, the scene-setting topic, and the secondary topic, which all precede the IBV position. Though in this section I will not investigate in detail the structural positions that these preverbal elements may occupy, their syntactic properties that are relevant to the discussion will be mentioned. Next I present these elements one by one.

3.3.1.1 Topical subject

First I investigate how the topical subject in Kukuya behaves in terms of both syntactic and information structural status. I show that a topical

subject in Kukuya can occur in various preverbal positions, while its interpretational characteristics can differ. It should be noted that the available positions for a topical subject that I mention are not necessarily different structural positions but different linear positions *relative to* other preverbal elements.

The first possible position that a topical subject can occur in is the initial position of a sentence, for example the topical subject in the SVO or SOV word order. Some examples are given in (88): in (88a) the subject is topical while the focus is on the VP, and the pronominal subject in the congruent answer (88b) is also topical since it is given and is what the predicate is about.

- (88) a. Nkaaká kǐ-má ká-sǐ?
 1.grandmother 7-what 1SM.PST-do.PST
 ‘What did grandmother do?’
- b. Ndé á-tól-i ma-buokó ma-kí-ték-e.
 1.PRO 1SM.PST-collect-PST 6-mushroom 6REL-7SM-sell-FV
 ‘She collected mushrooms to sell.’

A topical subject that occurs in the sentence-initial position can often be followed by multiple other topical elements, in which case the IBV focus position cannot be empty but is always filled by a focal element. The other in-between topical elements in the preverbal domain are usually the objects of the verb. I will introduce the latter type in detail as “secondary topic” in section 3.3.1.3. In example (89) we see that the subject *nkaaká* “grandmother”, which controls the class 1 subject marking on the verb, is followed by the object of the verb *buká* “cassava” and the interrogative word in the IBV position. In (90) and (91) are illustrated two answers in which the object and the adjunct are focused in IBV, and again the subject occurs in the sentence-initial position with another topical object sandwiched between the subject and the IBV element.

- (89) Nkaaká bu-ka ná ndé ká-bí-í kí-wâ?
 1.grandmother 14-cassava who 1.PRO 1SM.PST-refuse-PST INF-give
 ‘To whom didn’t the grandma give the cassava?’
- (90) (*Did the grandma give the beans to the CATS?*)
 Ambú, ndé ma-désu báa-mvá ká-wí.
 no 1.PRO 6-bean 2-dog 1SM.PST-give.PST
 ‘No, she gave the beans to the DOGS.’
- (91) (*How did father go to Djambala?*)
 Ndé Dzambála mu miilí ká-yení.
 1.PRO Djambala 18.LOC 4.leg 1SM.PST-go.PST
 ‘He went to Djambala on foot.’

A difference between these two kinds of sentence-initial subjects with regard to whether they are followed by other topical elements is that, when the grammatical subject is the sole argument in the preverbal domain, it can be indefinite; when the subject is followed by other preverbal elements, namely in a SXXV construction, it cannot be indefinite. In (92) the indefinite and non-specific reading can be deduced from the given context; in (93) the subject is modified by a strong quantifier “every”. Since an indefinite non-specific element or a subject NP modified by strong quantifiers such as “all” and “every/each” cannot be dislocated nor a discourse topic (Rizzi 1986; Zerbán 2006; Zeller 2008; van der Wal 2009), these examples suggest that there is at least one non-dislocated subject position in the preverbal domain.

- (92) (*You returned home and found some footprints on the floor, you say to your roommate:*)
 Mbuurú (nguumó) á-yení.
 1.person 1.one 1SM.PST-come.PST
 ‘Someone came.’

- (93) (*Context: the headmaster came to the class and distributed the candies to each of the children.*)

Ná mwáana á-bák-i ba-bonbon.
 every 1.child ISM.PST-get-PST 2-candy
 ‘Every child got candies.’

When the initial subject is followed by other topical objects, the subject cannot be indefinite and non-specific. In examples (94) and (95), we find that while the NP *mbuurú* and *kilóko* can have both the indefinite and definite reading which depends on the context, they can have the definite reading only when followed by other topical elements. In other words, the initial subject must be topical if followed by other topical elements in the preverbal domain. According to most speakers it is infelicitous to place the modifier *nguumó* “one” with the initial subject in the presence of other preverbal topical elements, as shown in (95a); the sentence can only become appropriate if the subject is the only preverbal element as in (95b). However, there is some intra-speaker variation on the judgement of (95a), it can be felicitous according to some speakers in the context of contrast on the direct object.

- (94) Ki-lóko mwáana mú-tswê kíi-bólik-i.
 7-thing 1.child 3-head 7SM.PST-hurt-PST
 ‘The thing in question hurt the child’s HEAD.’

- (95) a. %*Mbuurú nguumó* baa-ntsúú má-désu ká-búnum-i.
 1.person 1.one 2-chicken 6-bean 1SM.PST-feed-PST
 ‘The person/Someone fed the chicken the BEANS.’
 b. *Mbuurú nguumó* á-búnum-i baa-ntsúú ma-désu.
 1.person 1.one 1SM.PST-feed-PST 2-chicken 6-bean
 ‘One person fed the chicken the BEANS.’

A topical subject can also occur in a non-initial position preceded by another topical element which also seems to be the subject of the sentence. In example (96) and (97) the sentence can be ambiguous whether it is actually about the “father” and the “child” or the “hoe” and the “lamp”, respectively.

The initial elements in both sentences are obviously the possessor or at least the user of the syntactic subjects that control subject marking on the verb, which looks like the “possessor-raising” construction as in (98). In a similar construction in (99), the initial element “child” is not necessarily the possessor of the syntactic subject “festival” but should be an “experiencer”, and the sentence is indeed about the “child” rather than the “festival” since the Q-A pair targets to the information on the “feeling” of the “child”.

- (96) Taará **téme** ku-ní líí-dzinim-i?
 1.father 5.hoe 17-which 5SM.PST-disappear-PST
 ‘Where did father lose the hoe? (*lit.*: As for father, where did (his) hoe disappear?)’
- (97) Mwáana **múnda** wu-kí-fúúm-í maamá ku dzáandu
 1.child 3.lamp 3REL-7SM-buy-PST 1.mother 17.LOC 5.market
 á-dzínim-i.
 3SM.PST-disappear-PST
 ‘The child lost the lamp that mother bought at the market.’ (*lit.*: As for the child, the lamp that mother bought at the market disappeared.)
- (98) Mu-kokó á-tsilik-í **ṅíibi mu-líeme.**
 1-king 1SM.PST-cut-PST 1.thief 3-finger
 ‘The king cut the thief the/his finger.’
- (99) a. Mwáana **ki-yinga** bu-ní kí-wir-i?
 1.child 7-festival 14-which 7SM.PST-pass-PST
 ‘The child, how did the festival pass (for him/her)?’
 b. Ndé **ki-yinga** kí-bvé kí-wir-i.
 1.PRO 7-festival 7-good 7SM.PST-pass-PST
 ‘(For) him/her, the festival passed well.’

These examples are reminiscent of the “double subject construction” as is attested in Chinese and Japanese, but from these examples we can observe

that the initial element is clearly not an argument of the verb so it cannot be the grammatical subject, but should be analysed as a “scene-setting” topic, which is the second type of preverbal topical element that I would like to introduce next.

3.3.1.2 Scene-setting topic

A scene-setting topic usually sets the “spatial, temporal or individual framework” of the rest of the sentence (Chafe 1976; Li and Thompson 1976). Some examples of scene-setting topics in Kukuya are given in (100)-(102). In (100) the sentence-initial topic is a DP which is co-referential to the pronominal subject in the pseudo-cleft construction that follows; in (101) and (102) the scene-setting topics are adverbial phrases. What distinguishes the scene-setting topics from topical subjects or objects besides their semantic function is that a scene-setting element never functions as an argument of the verb thus it is not originated from the rest of the sentence (Lambrecht 1994), it occurs only for the sake of limiting the frame of the proposition or semantically relate the event described by the core sentence to an “external topic”. In addition, there is a further division on the relation of the scene-setting elements and the rest of the sentence in these examples. In (99) above and (100), the initial element is what the sentence is “about” as the whole sentence is telling something about the “child” and the “woman”; while in (101) and (102) the locative and temporal phrases only set the background or the scene of the sentence and the aboutness topic expression is the 1st person pronoun.

(100) **Mu-kái** wu-ká-búr-í ndé mú-kái wó balaka?
 1-woman 1REL-1SM.PST-give.birth-PST 1.PRO 1-female or 1.male
 ‘The woman, whom she gave birth to was a girl or a boy?’

(101) **Mu mu-súru,** me á-mún-i ba-kái bá-kâ-tólo
 18.LOC 3-forest 1SG.PRO PST-1SG.see-PST 2-women 2SM-PROG-cut
 nkwî.
 9.firewood
 ‘In the forest, I saw women cutting the firewood.’

- (102) Mvúla wǎ-yá me â-m-fúúm-á báa-ntaba nkáma.
 3.year 3REL-come 1SG.PRO FUT-1SG.SM-buy-FV 2-goat hundred
 ‘Next year I will buy a hundred goats.’

3.3.1.3 Secondary topic

The third type of topical element is what I label as “secondary topic”. Cross-linguistically, an utterance can contain more than one topic under discussion simultaneously, which is often attested in the predicate-focus structure as shown in (103). In (103a) the question is on some relation between mother and the goats, and the answer in (103b) adds information on both the mother and the goats, thus here the two arguments should be both counted as topics. The question now is how to determine primary and secondary topichood.

- (103) a. Ngúku baa-ntaba kí-má ká-sí?
 1.mother 2-goat 7-what 1SM.PST-do.PST
 ‘What did mother do to the goats?’
 b. Ngúku á-dzwí báa-ntaba.
 1.mother 1SM.PST-kill.PST 2-goat
 ‘Mother killed the goats.’

In Nikolaeva (2001), a secondary topic is defined as “an entity such that the utterance is construed to be about the *relationship* between it and the primary topic”. The primary topic is considered to be more pragmatically salient and is closely associated with the subject function (Dalrymple and Nikolaeva 2011); as the secondary topic would often be realised as the object of the verb, which corresponds to some assumption that in historical terms, objects are grammaticalised secondary topics (Givón 1984, 1990, 2001). In Vallduví’s (1992) approach, the old information in the utterance can be further split into informationally more and less prominent material, namely the “link” and “tail” which correspond to the primary and secondary topic we discuss here. While the “link” is what the new information is anchored to, the “tail” entails the presence of the “link” and

implies that some update is to be carried out to complete the information on the relation between it and the “link”. In other words, the primary and secondary topics stand in a certain presupposed relation, the secondary topic presupposes the existence of the primary topic, and the proposition is to add new knowledge to some relation between the primary and the secondary topics (Dalrymple and Nikolaeva 2011). In the above example (103), the primary topic is the “mother” and the secondary topic is the “goats”, since the question is on what actions are done on the goats, thus the topic “goats” as the patient entails the presence of the “mother” as the agent.

In particular in the Kukuya language, I propose that the distinction on primary and secondary topics are grammatically encoded via word order: if there are more than one topical elements in the preverbal domain, only the primary topic can be placed sentence-initially (excluding the scene-setting topics), while the secondary topic should be non-initial. There are three informational types that involve secondary topic in Kukuya: the first type is as in example (103) in which the focus extends over the transitive predicate only, namely the verb focus expression where both arguments of the verb are given; the second type is the possessive secondary topic as in example (96) above, in which the two preverbal topics are in a possessive relation and the possessor functions as the primary topic, while the possessum is the secondary topic and the syntactic subject (Nikolaeva 2001), a similar example is given in (104).

- (104) Ngúku ndzulí ku-ní á-dzinim-i?
 1.mother 1.cat 17-which 1SM.PST-disappear-PST
 ‘Where did mother lose the cat? (*lit*: As for mother, where did (her) cat disappear?)

The third type of secondary topic is attested in adjunct or argument focus constructions. In example (105) and (106), when the locative interrogative word is focused in the IBV position, the object which is given in the context occurs between the initial subject and the IBV focused element, resulting in two topical elements in the preverbal domain. In (105b) the assertion updates the addressee’s knowledge on the relation between uncle and the rice by adding information that it was yesterday that uncle ate the rice, here the

“uncle” functions as the primary topic and “rice” as secondary topic. Similarly in (106), there are two preverbal secondary topics “falling” and “plates” which are the two objects of the ditransitive verb “to launch”. The word order pattern in (105) and (106) is very commonly attested in the formulation of question-answer pairs in Kukuya.

- (105) a. Mu-pfúru **lóoso** munkí ká-dzí?
 1-uncle 5.rice when 1SM.PST-eat.PST
 ‘When did uncle eat the rice?’
 b. Ndé **lóoso** má-tsíká ká-dzí.
 1.PRO 5.rice 6-yesterday 1SM.PST-eat.PST
 ‘He ate the rice YESTERDAY.’

- (106) Taará **ma-sáani** bví ku-ní ká-tí?
 1.father 6-plate 9.falling 17-which 1SM.PST-launch-PST
 ‘Where did father throw the plates?’

In (107) and (108) the division between primary and secondary topic is seen in the context of argument focus. In (107) the recipient object of the ditransitive verb is focused in IBV, and the topical theme object is placed in the preverbal domain as the secondary topic; in (108) it is the subject that gets focused in IBV, and both objects of the verb “to give” are placed preverbally, in this case the theme “oranges” is the secondary topic.

- (107) (*Did the grandfather give the food to the DOGS?*)
 Ambú, ndé **bvi-kídzá** báa-ndzulí ká-wî.
 no 1.PRO 8-food 2-cat 1SM.PST-give.PST
 ‘No, he gave the food to the CATS.’
- (108) (*“Who gave the child the oranges?”*)
 Mwáana ma-láara **bí-búru** bí-wî.
 1.child 6-orange 8-parent 8SM.PST-give.PST
 ‘The child was given the oranges by the PARENTS.’

In the above examples, it is interesting to see that the exploitation of IBV focus is usually accompanied by the “fronting” of other topical elements to the preverbal domain, while it is grammatical that the topical objects and adjuncts remain in their base positions, i.e. postverbally. The exact trigger of this topic fronting, whether syntactic or pragmatic, is left for further research. Here I propose that the Kukuya language can grammatically distinguish the primary and secondary topic by word order: the sentence-initial (excluding the scene-setting topics) topic is always primary while the non-initial one is secondary. Since the primary topic usually sets the most important framework and aboutness of the main predication, while a secondary topic is less important and continuous in terms of referential accessibility and thematic importance (Givón 1990; Nikolaeva 2001; Croft 1991; Tsao 1987; Shi 2000), it is necessary that the primary topic scopes over the secondary topic, so the former is placed in the initial position.

The secondary topic must have a definite reading. In (109) we see that it is infelicitous to have an indefinite object “someone/one person” occur in the preverbal domain and function as a secondary topic; in (110) the preposed object *kilóko* “thing” can only be interpreted as some particular thing that has been mentioned before but cannot be indefinite non-specific, as can be deduced from the context. Note here that *kilóko* “thing” is not in IBV since it precedes the negative prefix.

- (109) (*You are traveling in a very quiet small town and you did not see anyone on the street, your friend said she saw a person's figure on the way and you ask her:*)

??We **mbuurú nguumó** ku-ní á-mún-i?
2SG.PRO 1.person 1.one 17-which 2SG.PST-see-PST

Int: ‘Where did you see someone/one person?’

- (110) (*You have a precious gift in your home. One day you found a theft but fortunately the precious thing was not stolen.*)

múibi **ki-lóko** ka-ká-túr-i ni.
1.thief 7-thing NEG-1SM.PST-steal-PST NEG

‘The thief did not steal the thing/*anything.’

*‘The thief stole nothing.’

In example (111) some sentences with minimal difference on the position of the object *mbuurú* “person” are illustrated. For (111a) and (111b), as implied from the context, the proposed object can only be interpreted as definite: (111a) and (111b) have the same interpretation and can both be appropriately used in the context of (a) but neither can be used in the context of (b), which shows neither sentence can have the reading of “I saw nobody”; in (111c) the object is placed in the IBV position and can only have a generic reading as “human-being”; while in (111d) the object in its canonical post-verbal position can have the indefinite reading and functions as a negative polarity item (NPI), or a definite reading can also arise according to the context.

- (111) a. (*Your uncle asked you to call a certain person sitting under a tree nearby to come, you went but did not find the person, you returned and said to the uncle.*)

Mbuurú me ka-á-mún-i ni.
 1.person 1SG.PRO NEG-PST-1SG.see-PST NEG
 ‘I did not see the person/*anyone.’

- b. (*#Your mum and you are entering a dark hall, you are walking in front and your mum asked you from behind if you saw anyone in the hall.*)

Me **mbuurú** ka-á-mún-i ni.
 1SG.PRO 1.person NEG-PST-1SG.see-PST NEG
 ‘I did not see the person/*anyone.’

- c. (*You saw a “monster” in the forest; you did not know what animal it was, and after coming back someone asked you if you see anybody in the forest.*)

Me ka **mbuurú** á-mún-i ni.
 1SG.PRO NEG 1.person PST-1SG.see-PST NEG
 ‘I did not see a PERSON/*anyone.’

- d. (*felicitous in the context of both (a) and (b)*)

Me ka-á-mún-i **mbuurú** ni.
 1SG.PRO NEG-PST-1SG.see-PST 1.person NEG
 ‘I did not see anyone/the person.’

In this subsection, I have shown that there can be multiple preverbal topical elements, which can be further divided into primary topics which include sentence-initial topical subject (also object, see next subsection) and scene-setting topics, and secondary topics which are usually objects of the verb. The interpretation on these topics with regard to definiteness and specificity may depend to a large extent on their relative position in the preverbal domain. An initial subject, if it is not the sole preverbal argument, and a secondary topic must be definite. There are still some further questions that need to be investigated, such as potential restrictions on the order of the preverbal elements, and the connection between IBV focus and topic fronting.

As mentioned at the beginning of the section, there is a fourth type of topical element in Kukuya, which is the topicalised object that occurs in the initial position of the sentence. The sentence-initial object is usually attested in an OSV order in (112) and an impersonal *ba*-construction in (113), which can serve as functional *passives* in this language. In the next section I will introduce in detail these two functional passive constructions.

- (112) **Bii-ndomó** kii-mbúli kii-dzi.
 8-sheep 7-lion 7SM.PST-eat.PST
 ‘The sheep were eaten by the lion.’ (*lit*: The sheep, the lion ate them.)

- (113) **Mu-tí mu máa-ŋgúlu** áli báa-tsílik-i mbvúla
 3-tree 18.CONN 6-mango RPST 2SM.PST-cut.down-PST 3.year
 wǔ-fíŋ-a.
 3REL-pass-FV
 ‘The mango tree was cut down last year.’

3.3.2 Functional passives

In this subsection I introduce how Kukuya makes use of the IBV focus position and the topic fronting tendency to express the passive meaning. Two particular structures are presented, namely the OSV and the imper-

- (116) Bu-ká búú mwáana nzulí á-wool-i.
 14-cassava 14.CONN 1.child 1.cat 1SM-snatch-PST
 ‘The child’s cassava was snatched by the cat.’ (*lit*: The cassava of the child, the cat snatched it.)

In the ditransitive constructions (117) and (118), we see that both the theme and the patient object can be fronted. The passive reading can be verified in the elicitation: when I asked the speakers to translate the French passive into Teke and there is an explicit agent in the sentence, the OSV structure is always used.

- (117) a. Báana ngúku á-télek-i bvi-kídza.
 2.children 1.mother 1SM.PST-prepare-PST 8-food
 ‘The children were prepared the food by mother.’
 b. Bvi-kídza ngúku á-télek-i báana.
 8-food 1.mother 1SM.PST-prepare-PST 2.children
 ‘The food was prepared for the children by mother.’
- (118) a. Mu-safuká mú-kái á-kwá-i mu mbhiele.
 3-safou.tree 1-woman 1SM.PST-chop-PST 18.LOC 9.knife
 ‘The safou tree was chopped with a knife by the woman.’
 b. Mbhiele mú-kái á-kwá-i mu mu-safuká.
 9.knife 1-woman 1SM.PST-chop-PST 18.LOC 3-safou.tree
 ‘A knife was used to chop the safou tree by the woman.’

For the examples above I only gave the passive translation as a stimulus. However, the OSV construction itself does not show apparent grammatical means that are dedicated to passive expression, and here I want to decompose the OSV structure to see how the passive reading has emerged. Pragmatically, passiveness is often considered as a “foregrounding and backgrounding operation” (Keenan and Dryer 2007) in which the patient is foregrounded to the sentence-initial position while the agent is backgrounded or unspecified. In this sense, a passive construction is similar to the topicalisation operation in which the patient is fronted to the sentence-initial

position to become the topic of the sentence, while the agent can remain in the original position or be demoted to a less/non-topical position. In other words, a passivised element is usually made topical. The availability of OSV structure to express passive is thus consistent with the generalisation in the above subsection that in Kukuya topical elements tend to occur in the preverbal domain, so the topical object in OSV is expressed in the sentence-initial position. Nevertheless, in the OSV construction the agent subject is always explicitly expressed, which is not expected in a canonical passive construction. In addition, recall that the OSV structure is what I have introduced for subject focus (see section 3.2.1.2) and is always felicitous as an answer to a subject *wh*-question, as shown in (119) and (120). The focal status of the agent is pragmatically incompatible with a prototypical passive construction in which the agent is usually demoted or even deleted.

- (119) a. Mwáana láana ná á-wî?
 1.child 5.orange 1.who 1SM.PST-give.PST
 ‘Who gave the child the orange?’
 b. Mwáana láala taará á-wî.
 1.child 5.orange 1.father 1SM.PST-give.PST
 ‘FATHER gave the orange to the child.’
- (120) a. Wũ-fúum-i ma-li ná ndé?
 1REL-buy-PST 6-wine 1.who 1.PRO
 ‘Who bought the wine?’
 b. Ma-li taará á-fúum-i.
 6-wine 1.father 1SM.PST-buy-PST
 ‘The wine was bought by FATHER.’ (Li 2020: 15)

In (121a) we can see that the OSV functional passive construction cannot have a *wh*-adjunct, since the IBV focus position is occupied by the agent and there is usually only one focused element in a Kukuya sentence, the interrogative phrase becomes infelicitous even in its base position; the only possible rephrasing is (121b) in which the passive reading is lost.

- (121) a. *Mwáana taará á-béer-i munkí?
 1.child 1.father 1SM.PST-beat-PST when
Int: ‘When was the child beaten by father?’
- b. Taará mwáana munkí ká-béer-i?
 1.father 1.child when 1SM.PST-beat-PST
 ‘When did father beat the child?’

When I intend to elicit a sentence like “What was stolen by X?” in which the “passivised” object is an interrogative phrase, the speakers still use the OSV order as in (122). At first glance, we see that a *wh*-element can occur in the initial position of the OSV construction to express passive. However, as will be discussed in the next section, the sentence in (122a) is actually a cleft construction in which the class 1 subject marker shifts from *á-* to *ká-*; in (122b) we find that the initial *wh*-word is incompatible with the canonical subject marker. The OSV order in (122) is not the OSV functional passive construction that we are discussing here.

- (122) a. **Kí-má** njíbi ká-túr-i?
 7-what 1.thief 1SM.PST-steal-PST
 ‘What was stolen by the thief?’
- b. ***Kí-má** njíbi á-túr-i?
 7-what 1.thief 1SM.PST-steal-PST
Int: ‘What was stolen by the thief?’

Therefore, it shows that the OSV structure, though it can function as a translational equivalent of a canonical passive construction, is by no means dedicated to express passive and is at least pragmatically different from a true passive (see Bostoen and Mundeke 2011 for similar proposal for Mbuun). The primary function of the OSV construction is to clearly delimit the different discourse roles of the subject and object, in which the object is topicalised and fronted to the initial position; the subject is focused in the IBV position and the focus reading is somehow “strengthened” by fronting the topical object. Pragmatically, the OSV construction can function as the equivalent of passive but is used only when the agent serves as the new or contrasted information thus needs to be explicitly expressed.

3.3.2.2 Impersonal *ba-* construction

The second equivalent of passive in Kukuya is the so-called impersonal *ba-* construction. In this construction, the verb always takes the class 2 subject marker *ba-* which is not anaphoric to any lexical or pronominal subject in the sentence or the discourse. The patient object can occur either postverbally or preverbally, while the agent is in most cases deleted or unspecified, and this is why the construction is labeled as “impersonal”. Some examples are illustrated below. In (123) and (124), the patient object occurs preverbally, and the agent is unknown and suppressed; while in (125) and (126) there is no preverbal element and the patient object occurs after the *ba-* verb.

(123) (*visual stimuli: What about the food?*)

Bviilá báá-tél-i **bví** **ku** **mfúúlá.**
8.food 2SM.PST-throw-PST 9.falling 17.LOC 9.road

‘The food was thrown onto the road.’

(124) **Mu-ŋwâ** **wu-kí-som-í** **báa-mpúku báá-kí-i.**
3.hole 3REL-7SM-go.out-PST 2-rat 2SM.PST-fill-PST

‘The hole where the rats went out was filled.’

(125) (*in a story, a candle was extinguished due to some unclear reason...*)

Níjáa **báá-dzíib-i** **bu-dzí.**
suddenly 2SM.PST-extinguish-PST 14-candle

‘Suddenly the candle was extinguished.’ (Li 2020: 31)

(126) **Báá-tí** **ndé** **bví** **ku** **mbali.**
2SM.PST-throw.PST 1.PRO 9.falling 17.LOC 9.outside

‘It was thrown outside.’ (Saint Matthieu V: 13)

In example (127) we see from the context that the agent should be “I”, and the subject marking on the verb is still *ba-*, which shows the impersonal nature of the class 2 subject marker in this construction.

- (127) (*You cut some firewood in the morning and you gave it to your brother who could not work.*)

Nkwíi yi-m-baal-í me báá-wî
 9.firewood 9REL-1SG.SM-cut-PST 1SG.PRO 2SM.PST-give.PST
 ngândukú aa me.
 1.brother 1.CONN 1SG.PRO

'The firewood that I cut was given to my brother.'

Example (128) shows that both objects of the ditransitive verb can be preposed in the *ba-* construction. Interestingly, from the context we see that the preverbal theme object is topical in (128a), while in (128b) the preverbal recipient object is focal. The different discourse status of the preverbal object here is reminiscent of the information structure of the preverbal subject discussed in section 3.1.2, I propose that the preverbal objects in (128) occupy different structural positions, the preverbal object (128b) is in the IBV focus position. In this regard, the preverbal DP of the *ba-* construction behaves more like a preverbal subject which can be either topical or focal.

- (128) (*To whom did mother give the keys?*)

- a. **Ma-fungúla** báá-wî taará.
 6-key 2SM.PST-give.PST 1.father
 'The keys were given to father.'
- b. **Taará** báá-wî ma-fungúla.
 1.father 2SM.PST-give.PST 6-key
 'Father was given the keys.'

In example (129) we see that the *ba-* construction is used when the preverbal DP is contrastively focused and placed in the IBV position, and (130) shows that a preverbal interrogative word can occur in the *ba-* construction. In this sense, the *ba-* construction also shows deviance from the canonical passive construction in that the functionally "passivised" element is not always topical but can also be focal.

- (129) a. **Bi-ko** bvi-kí-dzílík-í mú-kái ku ngulu
 8-clothes 8REL-7SM-reserve-PST 1-woman 17.LOC 9.inside
 aa nzó **báá-túr-i.**
 9.CONN 9.house 2SM.PST-steal-PST
 ‘The clothes that the lady kept in the house were stolen.’
- b. Ambú, ndé **mí-pará báá-túr-i.**
 no 1.PRO 4-money 2SM-steal-PST
 ‘No, her MONEY was stolen.’

- (130) (*You found that the bananas on the table disappeared, and you asked father.*)
 Ma-ko ná **báá-wî?**
 6-banana 1.who 2SM.PST-give.PST
 ‘The bananas were given to whom?’/‘Who was given the bananas?’

This functional passive construction with class 2/3rd person plural subject marking is actually commonly attested in Bantu languages and beyond (Frajzyngier 1982; Keenan and Dryer 2007; Cobbinah and Lüpke 2009). A number of Bantu languages such as Bàsàá (Hamlaoui and Makasso 2013), Mbuun (Bostoen and Mundeke 2011), Bemba (Kula and Marten 2010), Lunda (Kawasha 2007) and Matengo (van der Wal 2015) have reported this construction as a functional passive. In all these languages the patient can either precede or follow the verb in this construction. As for the agent, in Bàsàá, Mbuun and Matengo, it is always unspecified and can not be present even via an oblique phrase, while in Bemba and Lunda an oblique agent is allowed and even preferred. In Kukuya, the agent is usually deleted but sometimes it can be introduced by an oblique phrase headed by a class 18 locative pronoun *mu*. However, two situations need to be distinguished.

There are some cases in which the DP introduced by the oblique phrase seems to be the agent of the verb, as shown in (131) and (132). Though these expressions are considered to be quite marginal and rare in use, the speakers often give the active constructions as equivalent translations to them. However, (131) and (132) can be used in various contexts in which the DP in the oblique phrase does not necessarily function as the agent but

rather a “causer” of the event. In (131) the context can be that someone else gave the child the orange due to father’s commission or network, while in (132) it was not necessarily your wife who caught you but perhaps your wife reported you to the police or you committed a crime due to your wife. Given that the class 18 preposition often introduces a reason, here the oblique phrases in these two examples should be interpreted as reason phrases rather than the demoted agents.

- (131) Mwáana **báá-wî** láala **mu** **taará**.
 1.child 2SM.PST-give.PST 5.orange 18.LOC 1.father
 ‘The child was given an orange because of father.’

- (132) Me **báá-siib-i** **mu** **mu-káli**.
 1SG.PRO 2SM.PST-catch-PST 18.LOC 1.wife
 ‘I was caught because of the wife.’

The *ba-* construction with an oblique phrase cannot be a felicitous answer to a subject *wh-* question. To answer the subject question in (133a), the OSV structure in (133b) is the an answer par excellence, while (133c) is infelicitous here because the oblique phrase can only be interpreted as a purpose or a reason. The question-answer congruence may also have some effect here, since the question in (133a) does not involve the *ba-* construction, (133c) is not expected to be a felicitous answer.

- (133) a. Nzó yĩ ya mú-tálikí ná ndé á-tsú-i?
 9.house 9REL with 3-height 1.who 1.PRO 1SM.PST-build-PST
 ‘The tall building was built by whom?’
 b. Yó **mii-ndéle** míi-tsú-i.
 9.PRO 4-foreigner 4SM.PST-build-PST
 ‘It was built by the foreigners.’
 c. #Yó báa-tsú-i **mu** **mii-ndéle**.
 9.PRO 2SM-build-PST 18.LOC 4-foreigner
Int: ‘It was built by the foreigners.’
 ‘It was built for/because of the foreigners.’ (Li 2020: 30)

Based on all these examples on the oblique phrase in the *ba-* construction, I would rather conclude that the DP introduced by *mu* is never a true agent but can only function as a reason, a purpose or a method, though sometimes it can be ambiguously interpreted as the agent. In this sense, it is more plausible to still label the *ba-* construction as impersonal. Compared to the OSV structure, the *ba-* construction is used when the agent is unspecified or there is no need to express it.

To summarise, I have presented two functional passive constructions in Kukuuya, namely the OSV and the impersonal *ba-* construction. Both of the constructions can serve as the translational equivalent of a prototypical passive structure. However, their syntactic and pragmatic properties differ from each other and also from the prototypical passive. The OSV construction is used when the utterance is about the patient and the agent needs to be explicitly expressed as new or contrasted information. The impersonal *ba-* construction looks more similar to the canonical passive as the agent is usually deleted, but the preverbal object can either be topical or focal, which differs from the canonical passivised element. Here these two constructions only partially overlap with some properties of the canonical passive construction and can only be treated as functional equivalents.

* * *

In this section I have shown that different types of topic expressions tend to occur in the preverbal domain in Kukuuya. The topic expressions can be divided into primary and secondary topics: a primary topic often includes the topical subject or the scene-setting topics, which occur sentence-initially; a secondary topic is non-initial and is often attested in the preverbal domain accompanied by the IBV focus position being occupied. Two functional passive constructions are presented, namely the OSV construction and the impersonal *ba-* construction, which are used in different pragmatic context and both make use of the topic fronting tendency and the IBV focus strategy to express passive. A scheme on the mapping of word order and information structure of Kukuuya is illustrated in (134). In the next section I introduce cleft constructions and their connection with the IBV focus strategy.

(134)

scene-setting TOP non-argument	subject TOP argument	secondary TOP argument
FOC argument/adjunct		VERB ...

3.4 Cleft constructions

Clefts are one of the well-known focus marking devices in Bantu languages (Demuth 1987; Sabel and Zeller 2006; Cheng and Downing 2013; Hamlaoui and Makasso 2015; Lafkioui et al. 2016). In this section I present different types of cleft constructions in Kukuya and their functions in information packaging. I first give a description on the formation and interpretation of the basic cleft and (reverse) pseudo-cleft constructions, then I introduce a special construction that I label as a “reduced” cleft. I also propose and show evidence that the IBV focus construction in this language is very likely to have its origin in the cleft construction, and different intermediate grammaticalisation stages can be identified.

As for a basic cleft, here I refer to a construction parallel to the English sentence “it was a pancake that we ate”, and it can also be labeled as the *it*-cleft. Syntactically, a basic cleft usually consists of two clauses: one contains a nominal predicate and one contains a free relative clause. The focus reading arises from the combination of the relative clause and the nominal predicate. The relative part of the cleft is presented as the maximal group of referents to which the predicate applies and is equated to the referent in the nominal predicate, and in this way an identificational and exclusive focus reading is rendered (van der Wal and Maniacky 2015).

In Kukuya, a basic cleft can be used to express focus on arguments and adjuncts. An example of a basic cleft in Kukuya that fulfills all the syntactic properties mentioned above is illustrated in (135). We see that the sentence contains an initial copula that takes a default class 7 subject marker, a nominal predicate that takes a H tone prefix and a relative clause with segmental relative marking. In fact, this kind of “complete” cleft construction is never uttered in natural speech. The copula is usually omitted in affirmative sentences, so a cleft construction in Kukuya is mostly formed just by a nominal predicate followed by a free relative clause. In (135) the focus is on the clefted object that occurs in the initial position, and the following relative clause is used to exclusively identify it. When using a cleft as in (135), the speaker intends to express that the person only bought a shelf but nothing else.

- (135) (Kí-li) **ki-taabí** ki-ká-fúúm-í ndé ku dzáandu.
 7SM-COP 7-shelf 7REL-1SM.PST-buy-PST 1.PRO 17.LOC 5.market
 ‘It was a SHELF that s/he bought at the market.’

In example (136a) we see that the object cleft sentence can only be a proper answer to an object question but not to a VP question, so apparently the focus reading cannot be extended to a larger constituent in a cleft. We also see that (136a) cannot be continued with an additive sentence such as “and also some sheep”, showing that the cleft sentence expresses exclusive focus. In (136b) we see that in the negative counterpart of the cleft sentence, the copula shows up and hosts the negative prefix. Here the scope of negation is not the whole sentence but only the focus. A subject cleft sentence is given in (137). In all these examples, the clefted arguments receive an exclusive focus reading.

- (136) a. Baa-ntaba ba-kíí-fúúm-í mú-kái.
 2-goat 2REL-7SM.PST-buy-PST 1-woman
 ‘It was some GOATS that the woman bought.’
 “What did the woman buy?” ✓
 “What did the woman do?” ✗
 “...and also some sheep” ✗
- b. Ka-kí-li báa-ntaba ba-kíí-fúúm-í mú-kái ni.
 NEG-7SM-COP 2-goat 2REL-7SM.PST-buy-PST 1-woman NEG
 ‘It was not some GOATS that the woman bought.’
- (137) Wúna mvá wu-á-wí baa-ntaba buókó.
 only 1.dog 1REL-1SM-give.PST 2-goat 14.fear
 ‘It was only the dog who scared the goats.’

A pseudo-cleft refers to a construction that equates the referent of a headless relative clause with a nominal predicate, for example the English sentence “what we want is pizza”, and is also known as *wh*-cleft. The pseudo-cleft construction seems to be more frequently attested in Kukuya than the basic cleft and is usually used to express subject focus (see section 3.1.2), as

shown in (138). In (139) the alternative question begins with a dislocated topic *mu-kái* “woman” and is followed by a pseudo-cleft construction sentence with the predicative focal object at the end.

- (138) a. Ki-kí-túm-í mbaá ki-namá kí-ma?
 7REL-7SM-cause-PST 9.fire INF-burn 7-what
 ‘What caused the fire?’
 b. Baá-fúum-i ma-li ba-na?
 2REL-buy-PST 6-wine 2-who
 ‘Who (*pl.*) bought the wine?’
 c. Wǔ-dzí baa-ntsúú ka-kí-li mvá ni.
 IREL-eat.PST 2-chicken NEG-7SM-COP 1.dog NEG
 ‘(The one) who ate the chicken was not the dog.’
- (139) a. Mu-kái wu-ká-búr-í ndé mú-kái wó
 1-woman IREL-ISM.PST-give.birth-PST 1.PRO 1-female or
balaka?
 1.male
 ‘The woman gave birth to a girl or a boy?’
lit: “The woman, to whom she gave birth was a girl or a boy?”
 b. Wu-ká-búr-í ndé **balaka.**
 IREL-ISM-give.birth-PST 1.PRO 1.male
 ‘The one she gave birth to was a boy.’

In (140) a reverse pseudo-cleft sentence is illustrated. Here again we see that the reverse pseudo-cleft cannot be continued by a sentence like “and also some sheep”, which shows that it express exclusive focus.

- (140) Báa-ntaba (bá-li) ba-kí-fúúm-í mú-kái.
 2-goat 2SM-COP 2REL-7SM.PST-buy-PST 1-woman
 ‘The GOATS were what the woman bought.’
 “...and also some sheep” ✗

There is also a commonly seen construction which surfaces in the OSV word order and in which the focus is placed on the initial element, as illustrated

in (141). I would analyse this construction as a somehow “reduced” version of a basic cleft rather than a monoclausal construction with initial focus, for the reasons that will become clear shortly. This cleft construction is reduced in the sense that there is no segmental relative marker on the verb, but there are clues of relative marking. In (141) we see that the class 1 subject marking on the verb takes the form *ka-* rather than the canonical form *a-*, which is an indicator of relative marking on the verb (also see chapter 4 section 4.2). This construction is a natural way of expressing exclusive focus on the initial element but never on the whole VP, which corresponds more to the cleft construction than the IBV focus construction. Prosodically, the initial focused element is always independently phrased from the rest of the sentence, which can also show evidence for the cleft nature of this construction (Cheng and Downing 2013). Therefore, I label this construction as a reduced cleft and will hypothesise that it can reflect an intermediate stage of the grammaticalisation process from the cleft to the IBV focus strategy. It is worth noting that this construction should be distinguished from the OSV construction presented in section 3.2 in which the focus is in IBV, though they have the same linear word order.

- (141) Má-biríki taará káá-fúum-i ku mfaí.
 6-brick 1.father 1SM.PST-buy-PST 17.LOC 9.capital
 ‘It was some bricks that father bought from Brazzaville.’
 ‘...and also a motobike’ ✗

Some more examples of this reduced cleft construction are given in (142) and (143). The construction is most commonly attested as a *wh*-question as (142), in which the speakers place the interrogative word at the start of the sentence. In (143) the focus is on the quantifier of the initial NP, while the whole NP occurs in the initial position. The reduced cleft is only discernible when the initial focused element is a non-subject, since a reduced subject cleft cannot be distinguished from the canonical word order when there is no relative marker, no subject marking allomorphy or word order change.

- (142) Munkí mwáana káá-dzí ntsúí?
 when 1.child 1SM.PST-eat.PST 1.fish
 ‘When did the child eat fish?’

(143) (*The thief would have stolen more goats, but it was only a FEW.*)

Baa-ntaba bá-bíibi njíibi ká-túr-i.

2-goat 2-few 1.thief 1SM.PST-steal-PST

'The thief stole FEW goats.'

In this section I have presented three main types of cleft constructions in Kukuya, namely the basic cleft, (reverse) pseudo-cleft and the reduced cleft. I showed that all these constructions express exclusive focus on the clefted element. Some further research need to be carried out on the pragmatic distinctions on the cleft construction and the IBV focus strategy when they both express identificational/exclusive focus.

* * *

This chapter is devoted to providing a description on the word order variation and topic/focus expressions in Kukuya. In the first part, I have demonstrated that the language has a canonical SVO word order, while any deviation from this word order is produced for the purpose of information packaging. I have shown that a focused constituent, be it an argument or an adjunct of the verb, can be placed in its canonical position or in the IBV position, while the IBV position is preferred. VP focus and verb focus can also be expressed through the canonical SVO word order or by placing the object/infinite verb in the IBV position. Based on these facts and some additional tests, I concluded that the IBV position is really a dedicated focus position in the language, even though the focused elements are not obligatorily placed there. There is some interpretational difference between the IBV and *in situ* focus strategies, in which the IBV focus site is more often associated with identificational focus, and the other often expresses assertive focus. As for topical elements, they all tend to occur in the preverbal domain as in most other Bantu languages, and several types of topical elements can be distinguished, namely the scene-setting topics, primary and secondary topics. Interestingly, the occurrence of some topical elements in the preverbal domain depends on whether the IBV focus position is occupied. I also gave a detailed introduction on two particular constructions that can function as translational equivalents of the passive construction. Different

types of cleft constructions were also discussed. In the next chapter, I will show some shared grammatical properties between the cleft and the IBV focus construction, claiming that the IBV focus strategy, which characterises the expression of information structure in this language, has its origin in a cleft construction.

CHAPTER 4

On the origin of IBV focus

In many examples in the previous chapter we have noticed that the shape of the class 1 subject marker on the verb can vary according to word order. A straightforward example of this morphological alternation is given in (1).

- (1) a. Taará ná á-béer-i?
1.father 1.who 1SM.PST-beat-PST
‘Who beat father?’ [OSV subject focus]
- b. Taará ná ká-béer-i?
1.father 1.who 1SM.PST-beat-PST
‘Whom did father beat?’ [SOV object focus]

In (1a) the word order is OSV in which the subject is focused in IBV, and the class 1 subject marker takes the shape *á-*; while in (1b) it is the object that is focused and the subject marker shifts to *ká-*. Since in the canonical SVO word order, the class 1 subject marker always takes the form *á-* in past tense, I assume the prefix *á-* to be the default and unmarked form

of the class 1 subject marker, while *ká-* is considered to be a marked allomorph. As observed from (1), it seems that the SM alternation takes place when some element other than the subject is fronted to the IBV position. Considering what I have presented in chapter 3 that the exploitation of the IBV position is always related to some type of focus expression, we may want to know whether this SM allomorphy is actually a dedicated morphological device to encode focus or it is just indirectly associated with focus. In fact, this SM alternation is not only attested in IBV focus constructions but also in relative clauses, which may imply a connection between the two constructions. I take the class 1 SM alternation as the starting point of this chapter on investigating the origin of the IBV focus strategy.

In this chapter, I corroborate the hypothesis that the IBV focus strategy originates from a cleft, based on segmental and tonal evidence connecting these two constructions. I first present the distribution of the subject marking allomorphy in different tenses and aspects and with subjects of different persons/noun classes. I show that this allomorphy only occurs with subjects of class 1 and class 3¹ and speech participants, and it is attested in the past and future tenses as well as in negative sentences. This morphological alternation is not directly related to focus marking but may have its precursor in a cleft construction which always involves a relative clause. In addition, some tonal evidence also suggests that the verbal tones in the IBV focus construction pattern with relative constructions. I attempt to propose a possible grammaticalisation path of the IBV focus construction and give the diachronic motivation for it. Section 4.5 makes reference to some other West-Coastal Bantu languages in which the IBV focus position is also attested, displaying some micro-variation with regard to this focus strategy.

¹In many Teke varieties, class 1 and class 3 have merged, which may be due to their identical prefix shape (Hyman et al. 2019, also see chapter 2 section 2.3.2). The distinction on class 1 and class 3 nouns can only be distinguished in Kukuya from the different tone pattern of the connective marker, e.g. *mu-lúmi* (cl.1) *aa me* (my husband) and *mu-tímá* (cl.3) *áá me* (my heart). For simplicity, I only refer to class 1 nouns thereafter in this section.

4.1 Hypothesis: IBV focus originates from a cleft

I hypothesise that diachronically, the IBV focus strategy has its origin in a biclausal cleft construction and has been grammaticalised to become a monoclausal focus construction. That a focus construction can originate from a (pseudo-)cleft construction is not a brand new proposal in the studies of Bantu languages. In some previous studies of Bantu focus constructions, similar hypothesis was made for the IBV focus position in Kisikongo (Ndonga Mfuwa 1995; De Kind 2014), for the focus expressions in canonical word order in Hungan (H42, Takizala 1972, 1974). In Luganda there is also a preverbal focus construction that has developed from a biclausal cleft to a monoclausal focus construction, showing characteristics of both (van der Wal and Namyalo 2016). In Lingala and Kikongo speaking areas there is a *moto* construction expressing subject focus which is also analysed to have grammaticalised from a cleft (van der Wal and Maniacky 2015). Some northeastern Bantu languages such as Kikuyu have a verb-initial *ni-* that can be used to express term focus and predicate focus, which was considered to be derived from a copula in a cleft construction (Bennett et al. 1985; Güldemann 2003; Nurse 2006).

Particularly, in De Kind's (2014) analysis of the IBV focus strategy in Kisikongo (H16a), he observes that the large variation in cleft constructions in Kisikongo actually forms a continuum on the word order level from a biclausal inverted pseudo-cleft to a monoclausal SOV focus structure. The same class 1 SM *ka-* is used in SOV and non-subject relatives, alternating with other SMS in SVO or subject relative clauses, which can also show the connection between IBV focus and cleft. Nonetheless, he also admits that this hypothesis needs further corroboration by more tonal data. He points out that it should be investigated whether the tonal pattern of preverbal objects in monoclausal focus constructions correlates with that of focused constituents in biclausal clefts, and the verbal tone pattern in these two constructions should also be compared.

Now we consider some more grammatical properties of the IBV focus construction in Kikuyu. One minimal pair of sentences is given in (2): the canonical word order SVOO in (2a) and the IBV focus construction with

SOVO order in (2b). If we compare these two sentences which superficially differ only in word order (and interpretation), some segmental and tonal variation on the focused element and the verb can also be noticed. We see the *a-* versus *ka-* alternation of the class 1 SM on the verb. The IBV focused element *má-désu* “beans” has a H toned prefix (2b), while the postverbal object in (2a) takes the default L toned prefix. In addition, the verb *ki-wâ* “to give” is realised as its unmarked HL tone pattern *-wî* in the canonical word order (2a), while its tone becomes H as *-wí* in the IBV focus construction (2b). In the IBV focus construction there is also a H tone observed on the postverbal noun prefix *bá-*, and it seems that there is a grammatical H tone occurring between the verb and the following prefix.

- (2) a. Taará **áá-wî** **baa-ndzulí** **ma-désu**.
 1.father 1SM.PST-give.PST 2-cat 6-bean
 ‘Father gave the cats some beans.’ [SVOO]
- b. Taará [**má-désú**]_[FOC] **káá-wí** **báa-nzulí**.
 1.father 6-bean 1SM.PST-give.PST 2-cat
 ‘Father gave some BEANS to the cats.’ [SOVO]

In summary, in the SOVO order we have observed the class 1 SM *ka-* instead of class 1 SM *a-*, the H tone prefix on the IBV focused element, and the postverbal H tone. Given these characters of the IBV focus construction, next I discuss them in turn to investigate whether they can also be attested in cleft sentences.

4.2 Subject marking alternation in IBV focus and relatives

4.2.1 Subject marking alternation in IBV focus

First I introduce in which situations the subject marking alternation occurs and what shapes the allomorphs can have. This alternation does not only correlate with word order variation, but it is also entangled with tense/polarity of the sentence and is restricted to only speech participants (1/2SG/PL) and class 1 subjects. I will present different combinations of these determining factors of the SM in turn.

For class 1 subjects, the SM alternation with regard to word order and tense/polarity is summarised in Table 4.1. From the table we see that the default class 1 subject marker in both past and future tenses takes the form *a-* in SVO and OSV; while it alternates to *ka-* in a negative sentence and in SOV/S(O)XV, namely when a non-subject constituent is fronted to the IBV position. Some examples to help better understand the table are illustrated

word order	recent past	remote past	future
SVO and OSV	á	â	â
SOV and S(O)XV	ká	kâ	kâ
negative	ká	kâ	kâ

Table 4.1: The allomorphy of class 1 subject markers in Kukuya

in (3)-(11). In (3a) in the recent past tense, the subject marker occurs as the H toned prefix *á-* when the sentence surfaces in SVO; and in (3b) the prefix shifts to *ká-* when the theme object is fronted to the IBV position and gets focused. Similarly in (4a,b) when an adjunct is focused and placed in the IBV position, the subject prefix on the verb also appears as *ká-*, regardless of the position of the object.

- (3) a. Mu-loí á-wí báana wúna maa-nkúru.
 1-teacher 1SM.PST-give.PST 2.children only 6-pen
 ‘The teacher gave the children only pens.’
- b. Mu-loí wúna maa-nkúru ká-wí báana.
 1-teacher only 6-pen 1SM.PST-give.PST 2.children
 ‘The teacher gave the children only pens.’
- (4) a. Ndé nká bu-ní ká-dzwí?
 1.PRO 1.antelope 14-which 1SM.PST-kill.PST
 ‘How did s/he kill the antelope?’
- b. Ndé mu mbielé ká-dzwí nká.
 1.PRO 18.LOC 9.knife 1SM.PST-kill.PST 1.antelope
 ‘S/He killed the antelope with a knife.’

In example (5a) I show that in the near future tense, the class 1 SM takes the *â-* prefix with a falling tone in the SVO word order; in (5b) when there is a focused object in the IBV position, the SM alternates to *kâ-*. From the examples presented so far, we see that this SM allomorphy underlyingly alternates between two segmental forms *a-* and *ka-*. The alternation does not trigger tonal change and the tone on the prefix is only determined by the grammatical tone of the tense (see chapter 2 section 2.4.2 on the tone patterns in different tenses).

- (5) (*What will father buy at the market?*)
- a. Ndé â-fúúm-á má-láala.
 1.PRO 1SM.FUT-buy-FV 6-orange
 ‘He will buy some oranges.’
- b. Ndé má-láálá kâ-fúum-a.
 1.PRO 6-orange 1SM.FUT-buy-FV
 ‘He will buy some oranges.’

Example (6) shows that in the OSV order where the subject is focused and the object(s) is/are topicalised, the canonical subject marker *á-* is used. Assuming that the subjects in (6a,b) are focused in the IBV position, here we

see that the SM alternation does not only correlate with IBV focus, but that there seems to be a dichotomy between subject and non-subject (object, adjunct) which can influence the choice of the prefix.

- (6) a. Maa-ntséke ná á-yi-pfuk-á má-dza?
 6-field 1.who LSM.PST-IMPF-water-FV 6-water
 ‘Who watered the fields?’ [OSV subject focus]
- b. Mó ma-dzá taará á-pfuk-í.
 6.PRO 6-water 1.father LSM.PST-water-PST
 ‘They were watered by FATHER.’ [OSV subject focus]

The SM in negative conjugations also participates in the alternation without word order change. In the affirmative SVO sentence (7), the SM is *á-* as expected, while its negative counterpart takes the *ká-* prefix. Similarly in (8), the SM following the negative marker also shows up as *ká-*. In (9) the negative verb in the subject relative clause also takes the SM *ká-*. In these negative sentences there is no word order change at the clause level, in particular no object or adjunct is fronted from the postverbal domain to the IBV position as in the SOV/S(O)XV examples above.

- (7) a. Mwá wúa á-kwî.
 1.dog 1.DEM.II LSM.PST-die.PST
 ‘That dog died.’
- b. Mwá wúa ka-ká-kwî ni.
 1.dog 1.DEM.II NEG-LSM.PST-die.PST NEG
 ‘That dog did not die.’
- (8) Káli mvúla ka-ká-nók-i ni, kéne bhíí
 if 1.rain NEG-LSM.PST-rain-PST NEG CF IPL.PRO
 líi-báan-i mu-sálá.
 IPL.SM.PST-begin-SBJV 3-work
 ‘If it did not rain, we would have started the work.’

- (9) Mbuurú wu-kítĩ ka-ká-sál-i ni kâ-bvǎ
 1.person 1-COMP NEG-ISM.PST-work-PST NEG ISM.IMPF-fall
 ntsíina ŋa mu-ŋwá áá nzó aa ndé ná
 9.ground 16.LOC 3-mouth 3.CONN 9.house 9.CONN 1.PRO every
 tsúku.
 5.day
 ‘The person who never worked sits in front of his house everyday.’

Examples (10) and (11) display sentences in the remote past tense. Recall that the remote past tense with a class 1 subject is expressed by an auxiliary-like copula *âli* and an inflected verb with a HL toned SM (see chapter 2 section 2.4.2). Here we see that when an element, in this case the object, is fronted to the IBV focus position, both the copula and the main verb takes the *kâ-* prefix.

- (10) Ndé **báa-ntsúí kâ-li** **kâ-yi-fúum-a**.
 1.PRO 2-fish 1SM.RPST-COP 1SM.RPST-IMPV-buy-FV
 ‘S/He bought some FISH.’
- (11) Mu-káli aa mu-kokó ná kâ-li kâ-béer-i?
 1-wife 1.CONN 1-king 1.who 1SM.RPST-COP 1SM.RPST-beat-PST
 ‘Whom did the queen beat?’

In (12) and (13) the sentences are expressed in OSV order and the SM on the verb appears as *kâ-*, which runs counter to the OSV examples in (6). However, I suppose that (12) and (13) are intrinsically different constructions. As seen from the context and the translation, it is the sentence-initial rather than the IBV element that is in focus. These examples are in fact the reduced biclausal cleft construction (see chapter 3 section 3.4), which also surfaces in OSV order. I will discuss by the end of the chapter that I treat the reduced cleft as an intermediate construction between the basic cleft and the IBV focus construction. So here it should be distinguished from the OSV order expressing subject focus.

- (12) (*Did the thief steal many goats?*)

Baa-ntaba **bá-bíibi** ɲíibi ká-túr-i.

2-goat 2-few 1.thief LSM.PST-steal-PST

'It was a FEW goats the thief stole.' [reduced cleft]

- (13) Wúna **ma-biriki** taará ká-fúum-i ku mfaí.

only 6-brick 1.father LSM.PST-buy-PST I7.LOC 9.capital

'(It was) only BRICKS that father bought in Brazzaville.' [reduced cleft]

From the many examples with class 1 subjects illustrated above, we have a preliminary observation that the SM alternation occurs 1) when some focused element other than the subject is fronted to the IBV position; 2) in a reduced cleft OSV where a non-subject constituent is focused in the sentence-initial position; and 3) in negative sentences. So far the only generalisation that can be made is that this morphological alternation seems to encode the focus-related movement of a postverbal element to the preverbal domain, either to the IBV position or the sentence-initial position. Nonetheless, this generalisation cannot account for its occurrence in negative sentences.

For all the languages in the literature that have reported this SM alternation (Takizala 1972; Gueldemann 1996; Bostoen and Mundeke 2012; De Kind 2014), as far as I know, the allomorphy is restricted to class 1 subjects. Interestingly, in Kukuya the SM alternation also happens with 1SG, 2SG, 1PL and 2PL subjects, namely the speech participant subjects, but the alternation becomes purely tonal. We first consider the 1SG subject. It should be noted here that the 1SG SM always has a nasal part which assimilates with the following consonant, and the nasal part can be preceded by another prefix encoding tense information. In the glossing I parsed the two parts as separate prefixes. However, I would treat the tense prefix and the nasal prefix together as the 1SG SM to unify the formation of subject markers in this language, and the nasal prefix is placed after the tense prefix for morphophonological reasons (see chapter 2 section 2.4.1). The 1SG SM alternation is summarised in Table 4.2.

1SG	remote past	future
SVO and OSV	aN	aN
SOV and S(O)XV	âN	âN
negative	âN	âN

Table 4.2: The tone pattern of 1SG subject markers in Kukuya

We see from the table that the 1SG SM alternation does not involve segmental allomorphy as the *a-/ka-* opposition for class 1 subjects, but only the tonal change as a H tone insertion rule. Since the 1SG SM bears a L tone only in the remote past and future tenses in canonical word order, so the addition of the H tone can only be detected in these tenses. Two examples are given in (14) and (15).

- (14) a. Me lía a-m-fúúm-á bí-ko.
 1SG.PRO FUT FUT-1SG.SM-buy-FV 8-clothes
 ‘I will buy the clothes.’ [SVO]
- b. Me lía bí-ko â-m-fúum-a.
 1SG.PRO FUT 8-clothes FUT-1SG.SM-buy-FV
 ‘I will buy the CLOTHES.’ [SOV]
- (15) a. Me âli a-ma-í ki-ko.
 1SG.PRO RPST PST-1SG.SM.put.away-PST 7-clothes
 ‘I had put away the clothes.’ [SVO]
- b. Me âli kâ-ma-í kí-ko ni.
 1SG.PRO RPST NEG.PST-1SG.SM.put.away-PST 7-clothes NEG
 ‘I had not put away the clothes.’ [SNegVO]

In (14a) in the near future tense with the SVO order, the 1SG SM occurs as a L tone prefix, while in SOV (14b) in which the object is focused in IBV, the SM bears a falling tone. In (15a) the subject prefix with L tone is attested in the remote past tense, while in its negative counterpart (15b) the negative prefix *ka-* is fused with the HL toned subject prefix *â-* (which can only be

perceived when the speakers pronounce it morpheme by morpheme), and phonetically they are jointly realised as a mid tone prefix [kā].

The alternation on the 2SG SM is summarised in Table 4.3. In the recent past tense, the 2SG SM always occurs as a null prefix in SVO/OSV, while it takes the form á- in SOV/S(O)XV and negative sentences. In the remote past and the future tenses, the 2SG SM is realised as the L tone prefix a- in SVO/OSV, and as the HL tone prefix â- in the SOV/S(O)XV and negative sentences.

2SG+tense	recent past	remote past	future
SVO and OSV	∅	a	a
SOV and S(O)XV	á	â	â
negative	á	â	â

Table 4.3: The tone pattern of 2SG subject markers in Kukuya

Some examples on the 2SG SM alternation are given in (16) and (17). Example (16) shows the ∅- versus á- alternation in the recent past tense with regard to different word orders and polarity.

- (16) a. We Ø-yáab-i ki-líira ya kí-kila?
 2SG.PRO 2SG.SM-know-PST INF-read with INF-write
 ‘You know reading and writing?’ [SVO] (Paulian 2001: 7,
 glossing adapted)
- b. We ka-Ø-á-mún-i Zacharie ni?
 2SG.PRO NEG-2SG.SM-PST-see-PST 1.Zacharie NEG
 ‘You did not see Zacharie?’ [SNegVO]
- c. We kí-má ké Ø-á-min-í ŋaɔŋaa?
 2SG.PRO 7-what 7.PRO 2SG.SM-PST-swallow-PST just.now
 ‘What did you swallow just now?’ [SOV]

It should be noted here that for the 1SG subject *me* and the class 1 subject pronoun *ndé*, their subject prefix á(N)- in the canonical word order can sometimes be suppressed on the verb while compensated by lengthening the

and the far future tenses, the 1/2PL SM also bears a falling tone when an object is focused in the IBV position.

- (18) a. (*How are we going to celebrate the new year?*)
 Bhií yǎ we lía líi-dzwá ntaba.
 1PL.PRO with 2SG.PRO FUT 1PL.SM.FUT-kill.PST 1.goat
 ‘We are going to kill a goat.’ [SVO]
- b. Bhií yǎ we lía ka-líi-dzwá ntaba ni.
 1PL.PRO with 2SG.PRO FUT NEG-1PL.SM.FUT-kill.PST 1.goat NEG
 ‘We are not going to kill a goat.’ [SNegVO]
- c. Biábe ka-líi-wol-í baarí ba-kítĩ
 1PL.PRO NEG-1PL.SM.RPST-take-PST 2-people 2-COMP
 ka-báá-yáab-i ki-líra ni.
 NEG-2SM.PST-know-PST INF-read NEG
 ‘We did not employed people who did not know reading.’
 [SNegVO]
- (19) a. Bhií lía bí-ko líi-fúum-a.
 1PL.PRO FUT 8-clothes 1PL.SM.FUT-buy-FV
 ‘We will buy some CLOTHES.’ [SOV]
- b. Bé âli kí-má líi-fúum-i?
 2PL.PRO RPST 7-what 2PL.SM.RPST-buy-PST
 ‘What had you(pl.) bought?’ [SOV]

So far in this subsection I have presented the distribution of the SM alternation with class 1 and [+PARTICIPANT] subjects in different word orders, tenses and polarities. I have shown that the SM allomorphy can be either segmental as the *a-/ka-* opposition for class 1 subjects, or tonal as the L versus HL alternation for the [+PARTICIPANT] subject prefixes. In the next subsection I will show that almost the same SM alternation patterns are found in relative constructions, which forms one of the arguments for the hypothesis in section 4.1.

4.2.2 Subject marking in relative (cleft) constructions

In the grammar sketch provided in chapter 2, I have presented the formation of relative constructions in Kukuya. In this subsection I focus on the SM alternation in cleft constructions which always involve a relative clause. I will demonstrate that the alternation has almost the same distribution in relative and IBV focus constructions for class 1 and [+PARTICIPANT] subjects, pointing towards an origin in the relative verb form. I will also point out the cases in which the SM occurrence in these two constructions does not match.

In a subject relative with a class 1 subject in the recent past tense, the relative marker on the verb agrees with the head NP in noun class and is spelled out as the same shape as the class 1 demonstrative marker *wu-*. The tone of the SM slot on the relative verb is always realised as H, and if the SM takes a null form, this H tone just joins to the preceding L tone on the relative marker, which is realised as a rising tone. In (20) we see that the relative marker agrees with the class 1 subject NP and is realised as a rising tone, which is the most common way of class 1 subject relative marking.

- (20) a. Ndé *wũ*-banám-i á-tok-í ndziimi.
 1.PRO IREL-wake.suddenly-PST 1SM.PST-sweat-PST much
 ‘S/He who waked up suddenly sweated a lot.’
- b. Mbuuru *wũ*-yi-tsék-e me á-bvĩ mu
 1.person IREL-IMPF-mock-FV 1SG.PRO 1SM.PST-fall.PST 18.LOC
 ntsá dzuná kii-ntsiiba.
 9.inside 5.hole 7-sudden
 ‘The person who was mocking me suddenly fell into a hole.’

However, it remains to be investigated whether this *wũ-* prefix should be analysed as a relative prefix plus a floating H tone, or the relative marker *wu-* is fused with a H tone class 1 subject prefix *ú-* as an anti-agreement marker when the subject is extracted, which is documented in some other WCB languages (e.g. in the Kikongo cluster H10 and in Hungan H42) (Gívon 1975). For the latter account, recall that the class 1 SM has a canonical

shape *a-* in past tenses, but here it does not appear in the subject relative. Nonetheless, for some consultants the class 1 SM *a-* can still occur, and the REL-SM sequence can be realised as *wu-á-*, which depends on geographical locations of the speakers, upper or lower side of the Kukuya plateau. Therefore, I suppose that the occurrence of the *a-* prefix in subject relatives may have regiolectal variation, and whether there is the anti-agreement effect in subject relativisation needs further exploration.

Importantly, we see that the use of the allomorphic class 1 SM *ká-* in a subject relative is always ungrammatical, as illustrated in (21).

- (21) Mu-kái ná ndé wu-(*ká)-mún-i Gilbert?
 1-woman 1.who 1.PRO 1REL-1SM.PST-see-PST 1.Gilbert
Int: ‘Which woman (is the one who) saw Gilbert?’

As for non-subject relatives with a class 1 subject, recall the differential subject marking of the pronominal and the lexical subject (see chapter 2 section 2.5.2). In (22) we see that when the postverbal subject in a non-subject relative is a class 1 pronoun *ndé*, the SM is realised as *ká-*, which is of the same shape as the class 1 SM allomorph in the IBV focus construction. Example (23) shows that subject marking for *ndé* can also appear as the default class 7 SM *ki-*, which is further discussed in chapter 5 section 5.3.

- (22) Ndé kal-í yă nyama wu-ká-dzwí ndé,
 1.PRO 1.NARR.stay-PST with 1.animal 1REL-1SM.PST-kill.PST 1.PRO
 ŋa kí-kíni ma-táli mu ki-yá naama.
 16.LOC 7-period 6-sunshine 18.LOC INF-go 9.top
 ‘He stayed with the animal that he killed, (until) when the sun rose to the top.’
- (23) Bhií líi-tsuk-í mu ku-ká/kí-túr-í ndé
 1PL.PRO 1PL.SM.PST-talk-PST 18.LOC 17REL-1/7SM.PST-steal-PST 1.PRO
 mi-pará.
 4-money
 ‘We talked about the fact that s/he stole the money.’

Example (24) shows that when the postverbal subject is a lexical DP, the SM on the relative verb can only take the default marker *kí-* and the class 1 SM *ká-* cannot be used.

- (24) *ɲa-kí/*ká-yĩ* *múu-ndziá, mu-kái* *bvi-kídzá*
 16REL-7/*1SM.PST-come.PST 1-foreigner 1-woman 8-food
ká-yî-télek-e.
 1SM.PST-IMPF-prepare-FV
 ‘When the guest came, the woman was preparing food.’ (Paulian 2001: 16, glossing adapted)

In cleft constructions, we find that in the reverse pseudo-cleft (25a) which exploits a non-subject relative, the relative prefix is present on the verb and the subject is postverbal. Here the SM can only be the default marker instead of the class 1 SM. As for (25b) which is a reduced cleft (see more in chapter 2 section 2.5.3 on nominal predication and chapter 3 section 3.4 on clefts), the relative marker on the verb is deleted and the subject is no longer inverted, and here only the class 1 SM can be used whereas the default marker is impossible. Here I consider the subject prefix *ká-* to be the same SM as in non-subject relatives. In this sense, the verb in the reduced cleft (25b) still maintains some properties of a relative construction.

- (25) a. *Kí-taabí ki-kí/*ká-fúúm-í* *taará* *ku* *dzándu.*
 7-shelf 7REL-7/*1SM.PST-buy-PST 1.father 17.LOC 5.market
 ‘A shelf was what father bought at the market.’
 b. *Ka-kí-li* *kí-taabí taará* *ká/*kí-fúúm-i* *ni.*
 NEG-7SM-COP 7-shelf 1.father 1/*7SM.PST-buy-PST NEG
 ‘It was not a shelf that father bought.’

From the examples above, three types of asymmetries pertaining to class 1 subject marking in relative constructions can be summarised, which are the asymmetries between subject and non-subject, preverbal and postverbal placement of the subject, as well as pronominal and lexical postverbal subjects, as illustrated in Table 4.5 (for recent past tense).

SM form	relativised element		subject type and position ^a			
	subject	non-subject	pronominal		lexical	
			preverbal	postverbal	preverbal	postverbal
â-	✓	✗	✓	– ^b	✓	–
ká-	✗	✓	✓	✓	✓	✗
kí-	✗	✓	✗	✓	✗	✓

^a The values under “subject types and position” are only applicable in the presence of a true value under “relativised element”;

^b “–” means this condition never happens, in this case the subject in a subject relative is never postverbal;

Table 4.5: SM of class 1 subjects in relative constructions (*recent past tense*)

For relative constructions in other tenses, the class 1 SM alternation behaves similarly to that in the recent past tense, with only the prefixal tone being modified for that tense. For lexical class 1 subjects, example (26a) shows a remote past tense subject relative in which the SM takes the canonical form *â-* with a falling tone. This *â-* prefix can be contracted with the preceding relative marker on the auxiliary, and it also appears on the lexical verb. Example (26b) illustrates that in object relatives the default marker *kíi-* occurs both on the auxiliary and the main verb, with the falling tone encoding the remote past tense.

- (26) a. taará wǔ-(â)-li â-ték-i mfú
 1.father 1REL-1SM.RPST-COP 1SM.RPST-sell-PST 10.hair
 ‘father who had sold hair’
- b. mfú yi-kíi-li kíi-ték-í taará
 10.hair 10REL-7SM.RPST-COP 7SM.RPST-sell-PST 1.father
 ‘the hair that father had sold’

For the pronominal class 1 subject, in a subject relative (27a) the canonical SM is used; in the object relative (27b), the class 1 SM appears as *káa-* with a falling tone marking future tense; in (27c) both the SMS on the remote past auxiliary and the lexical verb occur as *káa-*. We also notice that in relative constructions of compound tenses in which an auxiliary is employed, the relative marker only appears on the auxiliary but is not reduplicated on the lexical verb.

- (27) a. ndé wũ-(â)-li â-fúum-i baa-ntaba
 1.PRO 1REL-ISM.RPST-COP ISM.RPST-buy-PST 2-goat
 ‘s/he who had sold the goats’
- b. baa-ntaba (líá) ba-káa-fúúm-á ndé
 2-goat FUT 2REL-ISM.FUT-buy-FV 1.PRO
 ‘the goats that s/he will buy’
- c. bi-ko bi-káa-li káa-ték-i ndé
 8-clothes 8REL-ISM.RPST-COP ISM.RPST-sell-PST 1.PRO
 ‘the clothes that s/he had sold’

So far I have presented the SM alternation in subject and non-subject relatives with class 1 subjects. Once we compare the alternation pattern with that attested in the IBV focus constructions in 4.2.1, the connection between them can be observed. If we map the subject focus expressions in SVO/OSV word order with subject relatives, and SOV/SO(X)V word order expressing non-subject focus with non-subject relatives, the *a-* versus *ka-* allomorphy has almost the same distribution in these constructions, as summarised in Table 4.6. What remains to be explained is why the use of the default class 7 SM *ki-* is impossible in the IBV non-subject focus strategy (see the analysis in chapter 5 section 5.3), but is obligatory in non-subject relatives for lexical subjects. We will take up this question in the next chapter, and for now we continue the presentation on SMS by considering subjects other than class 1.

class 1 SM	IBV focus	relative
subject	a-	a-
non-subject	ka-	ka-/ki-

Table 4.6: The *a-/ka-* alternation in IBV focus and relatives

Tonal variation on the [+PARTICIPANT] SMS is also attested in relative constructions, which again correlates with the IBV focus constructions. Since 1PL and 2PL subject prefixes surface in the phonologically identical CV-shape *li-* which is not fused with any preceding prefix, it is easier to capture the tonal change on this SM prefix without influence from the preceding

prefix even in rapid speech. Some examples of the 1/2PL SM in subject (28) and non-subject (29)(30) relatives are given below. In (28a) in the present tense, the 1PL SM bears a H tone on the relative verb and a L tone on the matrix verb; in (28b) the SM which otherwise appears as the L-toned *líi-* on the future-tensed matrix verb, occurs as a HL-toned prefix *líi-* in the subject relative; (28c) is in remote past tense, and the SMS on both the auxiliary and the lexical verb have a falling tone, which are otherwise realised as L in the non-relative form. In all the sentences in (28) the relative marker takes the class 2 *ba-* prefix, and it can only occur on the auxiliary in a compound tense.

- (28) a. Bhií ba-*lí*-ká-sál-á máa-ntséke
 IPL.PRO 2REL-1PL.SM-IMPF-work-FV 6-field
 lí-ká-sílik-a bú-su.
 IPL.SM-IMPF-wake-FV 14-front
 ‘We who work in the fields wake up early.’
- b. bhií lía ba-*líi*-fúúm-á báa-ntaba
 IPL.PRO FUT 2REL-1PL.SM.FUT-sell-FV 2-goat
 ‘we who will buy the goats’
- c. bhií ba-*líi*-li líi-ték-i bi-ko
 IPL.PRO 2REL-1PL.RPST-COP IPL.SM.RPST-sell-PST 8-clothes
 ‘we who sold the clothes’

In non-subject relatives, the relative verb agrees with the postverbal 1/2PL subject and the default marker *ki-* cannot be inserted, as shown in (29). In (30) in the remote past tense, both the auxiliary and the main verb agree with the inverted 1/2PL subject, and the SM bears a falling tone.

- (29) Ma-lúa ma-**líi**/***kíi**-lak-í **bhíi** kína
 6-disease 6REL-1PL/*7SM.PST-say-PST 1PL.PRO yet
 ka-báá-bák-í bu-báa-sá-a mó ni.
 NEG-2SM.PST-get-PST 14REL-2SM.FUT-conquer-FV 6.PRO NEG
 ‘The diseases that we talked about, (people) did not get to conquer them yet.’

- (30) a. mfú yi-**líi**-li **líi**-ték-í **bé**
 10.hair 10REL-2PL.SM.RPST-COP 2PL.SM.RPST-sell-PST 2PL.PRO
 ‘the hair that you(pl.) had sold’
 b. ma-meé ma-**líi**-li **líi**-tí
 6-stone 6REL-1PL.SM.RPST-COP 1PL.SM.RPST-launch.PST
bhíi bví
 1PL.PRO 9.falling
 ‘the stones that we threw away’

As for the SM morphology of 1SG and 2SG subjects, when overtly realised, it always takes the V(N) shape. In a relative construction the vowel-initial SM can always converge with the preceding relative marker and sometimes triggers vowel coalescence. Thus the tonal variation of the SM prefix can only be attested in elicited slow utterances. In a non-relative sentence in the remote past, the 1SG SM bears a L tone. The sentences in (31) illustrate relative constructions with a 1SG subject. In the subject relative (31a) the nasal 1SG SM occurs on both the auxiliary and the lexical verb, and the tense prefix bears a falling tone. In contrast, in the object relatives (31b,c) the nasal prefix cannot occur on the auxiliary but only on the lexical verb, while the tense marker also bears the falling tone. Example (31c) also shows that the tense marker *â-*, on which a H tone is inserted, is fused with the relative marker *ki-* on the auxiliary, and they are realised as *keé-* in which the two adjacent vowels undergo centralisation.

- (31) a. me wu-**â-n**-li **â-n**-ték-i bi-ko
 1SG.PRO 1REL-RPST-1SG.SM-COP RPST-1SG.SM-sell-PST 8-clothes
 ‘I who had sold clothes’

- b. mu-ti wu-â-(*n)-li â-n-kwá-í me
 3-tree 3REL-RPST-1SG.SM-COP RPST-1SG.SM-cut-PST 1SG.PRO
 ‘the tree that I had cut down’
- c. ki-ko keé-li â-n-ték-í me
 7-clothes 7REL.1SG.SM.RPST-COP RPST-1SG.SM-sell-PST 1SG.PRO
 ‘the clothes that I had sold’

Relative constructions with a 2SG subject are shown in (32), in which the same H tone insertion rule is observed. Since the 2SG SM always takes a null form, the tonal alternation is observed on the following tense prefix. In (32) the SM-tense markers on the auxiliary and the lexical verb both have a falling tone. For both 1SG and 2SG subjects, the relative marker on the auxiliary verb takes the class 1 form *wu-* which is underspecified for the [PERSON] distinction.

- (32) a. we wu-Ø-â-li Ø-â-ték-i bi-ko
 2SG.PRO IREL-2SG.SM-RPST-COP 2SG-RPST-sell-PST 8-clothes
 ‘you who had sold clothes’
- b. li-meé leé-li Ø-â-tí we
 5-stone 5REL.2SG.SM.RPST-COP 2SG.SM-RPST-throw.PST 2SG.PRO
 bví
 9.falling
 ‘the stone that you had thrown away’

From the examples above, I have shown that there is a consistent correlation between the tone of the SM in IBV focus and relative constructions with [+PARTICIPANT] subjects, as summarised in Table 4.7. We observe that there is always a floating grammatical H tone on the SM slot of the relative verb, which can be considered as part of a compositional tonal means of relative marking. Since the [+PARTICIPANT] SMS bear an underlying L tone, when they occur on a relative verb the grammatical H tone attaches to the left of the L tone on the SM and their combination is realised as a falling tone. For all the other 3rd person SMS which always have underlying H tone, I assume that the grammatical H tone attachment on the relative verb is still applied but does not affect the surface tone realisation, which is still H. Therefore

the H toneme on the SM in IBV focus/negative and relative constructions can be extended to all kinds of subjects.

SM+tense	canonical	IBV focus	relative
1SG	aN-	âN-	âN
2SG	Øa-	Øâ-	Øâ-
1/2PL	li-	lí-	lí-

Table 4.7: The tonal alternation of 1/2/SG/PL subjects in IBV focus and relatives (*remote past/future tense*)

In this subsection, I have shown that the SM alternation in relative constructions, whether segmental or tonal, correlates with those attested in the IBV focus strategy. Specifically, the *a-/ka-* alternation of class 1 SM is attested in the opposition between canonical word order/IBV focus as well as between subject/non-subject relatives. The grammatical H tone on the SM in relative constructions is also present on the SM in the IBV focus construction. These facts indicate that the verb in the IBV focus construction actually still manifests some relative marking properties, which can provide preliminary evidence that the IBV focus is associated with the cleft construction which always engages relative marking. Next I continue to consolidate this hypothesis with some more tonal evidence.

4.3 More tonal evidence on the connection between IBV focus and cleft

In this subsection I provide an overview of the tonal variation on the nominal prefix with regard to its position in a sentence as well as its information structural status. I first distinguish all the different environments in which a H tone occurs on the nominal prefix, identifying the primitive H tones such as the predicative H tone, and epiphenomenal H tones such as metatony effect and H tone spreading. Then I show that some tonal properties of the IBV focus construction can provide further evidence on its origin in a cleft.

In many Bantu languages, tonal morphology on NPs has much relevance for syntactic relations, one example is the so-called “tone case” that has been reported in languages such as Kikongo (Daelemann 1983, Blanchon 1998) Umbundu (Schadeberg 1986), Giphende (Hyman and Ngalasso 1998) and Herero (Kavari et al. 2012). In Kukuya only the nominal prefix but never the stem can undergo tonal change under different syntactic environments. The nominal prefixes of all noun classes in Kukuya have a default L tone in the citation form of a noun, or when the noun is non-focal or dislocated; a H tone on the prefix is considered to be marked. A nominal prefix usually contains only one mora, but is realised as two moras if the stem starts with a nasal, and in this case only the first mora is subject to tone change rules. A monomoraic nominal prefix never bears a rising or falling tone. I first list all the possible situations in which the tone on the NP prefix shifts to H, as in Table 4.8, and then I disentangle different types/functions of these tonal variation cases in turn.

In the phonology sketch in chapter 2, I have shown that in Kukuya a phonological domain contains a stem plus any following prefix and five fixed tone patterns can be mapped onto this domain. Here I suppose that the occurrence of the H tone in the first two situations in the table, namely when the NP occurs after the possessive marker or after the particle *yă* “with”, is an epiphenomenon triggered by the spreading of the H tone on the possessive marker, rather than pertaining to an independent type of grammatical H tone. In (33a,b) the possessive marker has an underlying

Situation	Example	Translation
after possessive marker	bi-síkí bí mí-féme	“the organs of pigs”
after <i>yâ</i> “with”	ndé ya mú-lúmi	“she and husband”
object of infinitive verb	ki-maná mú-sála	“to finish work”
object in certain tenses	Taará kâ-nywá má-keé.	“Father smokes tobacco.”
predicative NP	Ndé mú-tsúli.	“He is a blacksmith.”
IBV focused NP	Mvá bí-pfúó ká-dzí.	“The dog ate the BREAD.”
postverbal subject in relatives	ntaba wu-kí-fúú mí-mú-kái	“the goat that the woman bought”
postverbal object in SOVO/SXVO	Ndé mvá ká-wí má-désu.	“S/He gave the DOGS beans.”
NP after negative verb	Ndé kíni ka-ká-bvúúrí mí-pará ni.	“S/He did not return the money yet”.

Table 4.8: Situations of H tone occurrence on the nominal prefix

H tone and it forms a prosodic domain with the preceding stem as seen from the H tone plateauing in (33a) and stem-final lowering in (33b) (to avoid LHH sequence). The H tone on the possessive marker also carries over onto the following prefix, if there is one. In (33c) the class 9 possessive marker bears an underlying L tone, and the tone on the following nominal prefix is also L, which shows that the tone on the nominal prefix is actually determined by the preceding possessive marker.

- (33) a. bi-[bú-bááná bí bá]a-ndzulí
8-DIM-child 8.CONN 2-cat
‘the small children of the cats’ [báana ‘children’]
- b. ki-[bhiima kíí mú]-kóko
7-corpse 7.CONN 1-king
‘the corpse of the king’ [ki-bhiimá ‘corpse’]
- c. [nzó yii mu]-kái
9.house 9.CONN 1-woman
‘the house of the woman’

As for the associative marker *yă* which has an underlying rising tone, when followed by a nominal prefix, the rising tone is realised on two tone-bearing units: the associative marker bears a L tone and the following prefix with a H tone; when there is no following prefix, the rising tone just appears on

- b. ki-mún-á má-kinima
 INF-see-FV 6-pain
 ‘to suffer’ [ki-múna “to see”]
- c. ki-sá mí-táami
 INF-do 4-amusement
 ‘to have fun’ [ki-sâ “to do”]

In certain tenses/aspects, the conjugated verbs also exhibit tonal alternation that is similar to metatony. In (36a,b), in the near future tense and the perfect aspect the tone on the verb-final vowel is realised as H and is extended onto the following nominal prefix. Examples (36c,d) show that the verb-final vowel does not become H in the recent past tense and imperative, where the tone on the nominal prefix of the complement remains L. In (36e) even the verb has an underlying H tone FV, the following prefix still carries a L tone in the recent past tense. If we view metatony in the infinitives as the default transitive VP tone pattern, the modified tone pattern on the inflected verbs can be accounted for by some intervening grammatical tones between the verb and its complement, encoding TAM distinctions. As seen in (36), this floating grammatical tone can be a H tone or zero in the future tense (36a), and L in the past tense and imperative (36c-e). An intervening L tone can block the metatonic H tone on the FV which may have originated from the following prefix, but does not impact a lexical H tone (36e).

- (36) a. Bó báa-bvúúr-á mí-para.
 2.PRO 2SM.FUT-return-FV 4-money
 ‘They will return the money.’ [ki-bvúura “to return”]
- b. Bó báa-maa-bvúúr-á mí-para.
 2.PRO 2SM.PST-PERF-return-FV 4-money
 ‘They have returned the money.’
- c. Bó báa-fúum-i baa-ntaba.
 2.PRO 2SM.PST-buy-PST 2-goat
 ‘They bought some goats.’ [ki-fúuma “to buy”]
- d. Sâ mu-tere áá me taataa.
 do.IMP 3-basket 3.CONN 1SG.PRO well
 ‘Make a good basket for me!’ [ki-sâ “to do”]

- e. Bó báá-swaak-í ma-sáani.
 2.PRO 2SM.PST-wash-PST 6-plate
 ‘They washed the plates.’ [ki-swaakí “to wash”]

The table above also shows that predicative NPs have a H tone prefix. In Kukuya the copula is usually absent in affirmative sentences, therefore a predicative construction is mostly expressed by juxtaposition of two NPs, as illustrated in (37) in which the predicative NP takes a H tone prefix.

- (37) Ndé mú-tsúli.
 1.PRO 1-blacksmith
 ‘He is a blacksmith.’

A construction like “it is...” is expressed by simply placing a H tone on the nominal prefix, as shown in (38a); when there is no segmental prefix attached to the stem, the stem-initial consonant is geminated and the H tone is realised on the first part of the geminated consonant (Paulian 1975; Hyman 1987) as in (38b); in (38c) the H tone is placed on the adnominal prefix of a predicative adjective.

- (38) a. Mú-ti.
 3-tree
 ‘It is a tree.’ [mu-ti “tree”]
 b. Ílege.
 1.weaver
 ‘S/He is a weaver.’ [lege “weaver”]
 c. Ma-téme má-bvé.
 6-hoe 6-good
 ‘The hoes are good.’ [ki-bvé “good”]

In (39) and (40) two more examples on predicative NPs with the H tone prefix are shown. In (39) the predicative expression is realised by juxtaposing two NPs whereby the prefix of the second NP receives the predicative H tone. Example (40) is a pseudo-cleft which is formed by a free relative plus a predicative NP.

- (39) Ki-báka **ki**-báka, bu-bila.nkele **múu**-nkwáára.
 7-obtain 7-obtain 14-question 3-keeping
 ‘To obtain is to obtain, the question is (how) to keep.’
 (Paulian 1975: 194, glossing added by the author)

- (40) Kĩ-n-dzií me ki-nywâ **má**-dzá maa-mfé.
 7REL-1SG.SM-please 1SG.PRO INF-drink 6-water 6-cold
 ‘What I like to drink is cold water.’

It remains to be investigated where this predicative H tone originates. In some previous studies on the tonal marking of predication in Bantu, there are constructions that express predication by tonal replacement. In some Bantu languages, a H tone can replace the inherent L tone on the noun class prefix or the augment. For example in Shona (S10, Zimbabwe), a H tone occurs on the nominal prefix to express identification as shown in (41); while in Herero (R30, Namibia) predication is achieved by the H tone being attached onto the the augment, which is the case both for nominal predicates (42b) as well as in adjectival predicates (42d). This predicative H tone may come from the H tone on a historical copula, which Meeussen (1967: 115) reconstructed as **ní*.

- (41) a. Shona (Welmers 1973: 323)
 mù-nhù
 1-person
 ‘person’
 b. **mú**-nhù.
 1-person.PRD
 ‘It is a person.’
- (42) a. Herero (Möhlig and Kavari 2008: 122, Kavari et al. 2012)
 ò-tjì-hávèrò
 AUG-7-chair
 ‘chair’

- b. ó-tjì-hávérò
AUG-7-chair
'It is a chair.'
- c. ò-zò-ngòmbè ò-zò-néné
AUG-10-cow AUG-10-big
'big cows'
- d. ò-zò-ngòmbè ó-zò-néné
AUG-10-cow AUG-10-big
'The cows are big.'

Through many examples in chapter 3 we have already seen that the prefix of a focused NP in the IBV position always carries a H tone, while the post-verbally focused NPs do not. In (43) both sentences in SVO and SOV order can be felicitous answers to an object question, but only the IBV focused NP has the H tone prefix. In (43b) the stem of the focused NP *ma-láala* "orange" also undergoes tonal change from HL to H, which is due to the H tone plateauing effect triggered by the immediately following H tone SM *ka-* on the verb.

- (43) (*What did father buy?*)
- a. Ndé á-fúum-i **ma-láala.**
L.PRO 1SM.PST-buy-PST 6-orange
'S/He bought some oranges.' [SVO]
- b. Ndé **má-láálá** ká-fúum-i.
L.PRO 6-orange 1SM.PST-buy-PST
'S/He bought some ORANGES.' [SOV]

When there are multiple NPs in the preverbal domain, only the IBV NP can and must have the H tone prefix (if they have a segmental prefix at all), as shown in the examples below. In (44) the correctively focused recipient object is placed in IBV with the H tone prefix. In (45) and (46) what occurs in IBV is a focal subject which also takes the H tone prefix. In all these examples the other preverbal non-focal elements cannot have the H tone prefix, which again shows that the occurrence of the H tone prefix correlates with the IBV position.

- (44) (*Did the woman give the fish to the DOGS?*)

Mu-kái baa-ntsúi báa-ndzulí ká-wî.
 1-woman 2-fish 2-cat 1SM.PST-give.PST
 ‘The woman gave the fish to the CATS.’

- (45) (*Who gave the child the oranges?*)

Mwáana ma-láala bí-búru bíi-wî.
 1.child 6-orange 8-parent 8SM.PST-give.PST
 ‘The child was given the oranges by PARENTS.’

- (46) Mfúúlá yi-ká-n-yé me yi báa-sinwá
 9.road 9REL-IMPF-1SG.SM-go 1SG.PRO 9.REL 2-Chinese
 báa-sí (yó).
 2SM.PST-do.PST 9.PRO

‘The road on which I am walking was built by the CHINESE people.’

In the reduced cleft (47) the focused NP is not placed in IBV but occurs sentence-initially, but it also takes the H tone prefix. As seen from the context in (48), it is not the whole NP but only the numeral modifier that is correctively focused. Interestingly, here only the agreeing adnominal prefix bears the H tone, but not the prefix of the head NP. So it could be the case that the H tone prefix occurs on a preverbally focused NP or a subpart of it, whether it is in the IBV position or sentence-initial; or the modifier could be used pronominally (and predicatively), and (48) is effectively interpreted as ‘the woman, as for knives, it’s THREE that she is holding’.

- (47) (*What did father buy?*)

Bí-ká taará ká-fíum-i.
 14-cassava 1.father 1SM.PST-buy-PST

‘It was the cassava that father bought.’

(48) (*Is the woman holding TWO knives?*)

Ndé maa-mbhiélé **má**-tíri kâ-kwaal-a.

1.PRO 6-knive 6-three 1SM.IMPf-hold-FV

'She is holding THREE knives.'

The occurrence of a H tone on the nominal prefix is also attested on the postverbal subject in non-subject relatives. In (49), in the free relative clause the verb-final vowel and the following prefix of the postverbal subject both bear a H tone, which is at first glance reminiscent of the metatony effect mentioned above.

(49) Me kâ-n-kín-a ŋa-kí-yím-á **mú**-kálí aa

1SG.PRO IMPF-1SG.SM-dance-FV 16REL-7SM-sing-FV 1-wife 1.CONN

me.

1SG.PRO

'I am dancing while my wife is singing.' [*ki-yúma* "to sing"]

However, this should be an additional grammatical H tone rather than metatonic. In (50) and (51), according to my earlier analysis there should be a verb-final floating grammatical L tone encoding past tense and triggering a L tone on the following prefix (see (36) above). However in these examples the postverbal subject has a H tone prefix, which indicates that there should be another verb-final H tone occurring after the past tense L tone and spreading onto the following prefix.

(50) Mbuká yi-kíi-sweek-í **mú**-kái ntséke ŋíibi

9.place 9REL-7SM.PST-hide-PST 1.woman 9.product 1.thief

á-swool-í (yó).

1SM.PST-find-PST 9.PRO

'The place where the woman hid food was found by the thief.'

- (51) Baa-ntsúú ba-kíí-ká-í báa-ndukú báá-bol-í
 2-chicken 2REL-7SM.PST-grill-PST 2-friend 2SM.PST-get.wet-PST
 mu mvúla.
 18.LOC 3.rain
 ‘The chicken that the friends grilled got wet due to the rain.’

In (52) a minimal pair of subject and non-subject relatives is displayed. Despite the asymmetry in the subject agreement morphology on the verb, we notice that in the subject relative (52a) the FV on the verb together with the following nominal prefix of the object both carry a L tone as in SVO sentences in past tense; while in the object relative (52b) the tone pattern on the verb shifts from HL to H (H tone plateauing) and the tone on the prefix of the postverbal subject NP is also realised as H, which indicates the emergence of a verb-final H tone.

- (52) a. mu-kái wǔ-fúum-i **mi-féme**
 1-woman 1REL-buy-PST 4-pig
 ‘the woman who bought the pigs’
 b. mi-féme mi-kíí-fúúm-í **mú-kái**
 4-pig 4REL-7SM.PST-buy-PST 1-woman
 ‘the pigs that the woman bought’

Further research needs to be carried out to explain why this verb-final H tone is limited to non-subject relatives but absent in subject relatives. This “H grammatical tone occurring between the verb and the (postverbal) subject” (Hyman 2012: 109) in non-subject relatives is also reported in languages such as Nzadi (B865, Hyman 2012), Haya (JE22, Hyman & Byarushengo 1984) and Giphende (L11, Hyman 2017). Hyman (2012) conjectured the H tone to be a trace of a postverbal relative marker or pronoun. The grammatical H tone may have replaced the past tense L tone between the non-subject relative verb and the postverbal subject and carries over onto the following prefix. The H tone spreading onto the prefix of the subject NP also suggests that the relative verb and the postverbal subject are in the same phonological domain. I will discuss more on the syntactic derivation of non-subject relatives and its surface tone pattern in chapter 5.

Notably, the tonal distinction discovered in subject/non-subject relatives has a similar occurrence in subject/non-subject focus in the IBV strategy. Example (53a) shows that in an SVO sentence in the past tense, the FV of the verb *ki-wâ* “to give” appears in its citational HL tone pattern, and the following nominal prefix bears the canonical L tone; while in (53b) when the IBV position is occupied by a focused object, the tone on the verb-final vowel and the prefix of the postverbal non-focal object both become H. The same tonal alternation occurs in (54) where the focused element in IBV is a locative adjunct. From these examples we see that similar to the non-subject relatives, a verb-final H tone occurs when a non-subject element is focused in the IBV position.

- (53) a. Me á-m-wî mu-káli mi-pará.
 ISG.PRO PST-1SG.SM-give.PST 1-wife 4-money
 ‘I gave my wife the money.’
- b. Me mú-káli á-m-wí mí-pará.
 ISG.PRO 1-wife PST-1SG.SM-give.PST 4-money
 ‘I gave my WIFE the money.’
- [*ki-wâ* “to give”]

- (54) (*Where did father buy the wine?*)
 Ndé ku dzándú ká-fúúm-í má-lí.
 1.PRO 17.LOC 5.market 1SM.PST-buy-PST 6-wine
 ‘He bought the wine AT THE MARKET.’ [*ki-fúuma* “to buy”]

If the tone pattern of the IBV focus construction indeed reflects a mapping with relatives, we would expect that the emergence of the verb-final H tone should not occur in subject focus constructions since it is not attested in subject relatives (see (52a)), and this is borne out in Kukuya. Both sentences in (55) show that when the subject is focused in IBV and there is a postverbal object, no H tone appears on the verb-final vowel or the following nominal prefix. However, there is one counterexample in my corpus, i.e., the one in (56), in which the subject is in focus and a grammatical H tone is observed

between the verb and the following prefix of the object. I leave this for further research.

- (55) a. Mwáana ná á-wî ma-láala?
 1.child 1.who 1SM.PST-give.PST 6-orange
 ‘Who gave the child the oranges?’
 b. (*Who watered the tree?*)
 Mu-tí taará á-mwáal-i ma-dzá.
 3-tree 1.father 1.SM.PST-water-PST 6-water
 ‘FATHER watered the tree.’
- (56) Mwáana lí-meé lí-búl-í mú-tswê.
 1.child 5-stone 5SM.PST-hurt-PST 3-head
 ‘The child, the STONE hit her/his head.’

One question here is whether this grammatical H tone arises verb-finally and spreads onto the following prefix, or it starts from the nominal prefix itself. In the IBV focus construction, the emergent H tone can only be perceived when there is a complement or an adjunct following the verb, but it cannot be observed when the verb is final, since there is a general sentence-final H tone lowering rule in this language. In a non-subject relative, there is always a postverbal subject, so the H tone can be easily detected on the prefix of the subject NP. If the H tone were to start from the prefix of the postverbal NP and marks a “tone case”, it is problematic to assume that in the IBV focus strategy the H tone originates from the prefix of an object NP while it comes from a subject NP in a non-subject relative. So here I suppose that the H tone emerges verb-finally, and the H tone on the prefix of the following NP in both constructions is spread from the verb-final H tone.

The last type of H tone occurrence on a nominal prefix is found with an object NP following a negative verb. For some consultants but not all, there is also a grammatical H tone occurring between the negative verb and the following object NP, as shown in (57). However, the H tone emergence is unexceptionally observed on a negative verb when there is

another non-subject element preceding the verb, in most cases it is the class 7 NP *ki-ni* “period” as in (58), which looks quite similar to the IBV focus construction in the linear word order. Recall that the segmental and tonal SM alternation are attested in IBV focus as well as in negative sentences (see section 4.2.1), here the presence of the verb-final H tone can add evidence on the diachronic connections among IBV focus, relatives, and negative sentences.

- (57) %Ndé ka-ká-fúúm-í má-sáani ni.
 1.PRO NEG-ISM.PST-buy-PST 6-plate NEG
 ‘S/He did not buy the plates.’

- (58) Me kí-ni ka-á-m-bvúúr-í mí-pará ni.
 1SG.PRO 7-period NEG-PST-1SG.SM-return-PST 4-money NEG
 ‘I did not return the money yet.’

In this section, I have presented various types of H tone occurrence on the nominal prefix. Crucially, we have found some tonal connections, both nominal and verbal, between the IBV focus and the relative constructions in addition to the SM alternation. The predicative and IBV focused NPs both have a H tone prefix; and there is a verb-final grammatical H tone occurring in the non-subject IBV focus context as well as in non-subject relatives. There are also some H tone occurrences that are attributed to the H tone spreading from a preceding element. A comparison of all these grammatical features between the cleft and IBV focus constructions is illustrated in (59), from which we can see the correspondences of each feature that point to the close connections between the two constructions. I take the SM alternation and the verb-final H tone as the features of a relative verb form. Next I try to provide a possible grammaticalisation pathway of the IBV focus strategy from the cleft in detail.

- (59) a. (Kí-li) [má-désu]_[FOC] ma-*áá/káá-wí ndé
 7SM-COP 6-bean 6REL-1SM.PST-give.PST 1.PRO
 baa-ndzulí.
 2-cat
 ‘It was some BEANS that s/he gave to the cats.’ [cleft]
- b. Ndé [má-désú]_[FOC] *áá/káá-wí baa-nzulí.
 1.PRO 6-bean 1SM.PST-give.PST 2-cat
 ‘S/He gave some BEANS to the cats.’ [IBV focus]

4.4 Origin and grammaticalisation of IBV focus

Given that the IBV focus and cleft constructions share many similar grammatical features, and that the innovation of a focus construction from a cleft is attested cross-linguistically, a diachronic scenario seems plausible. In this section, I investigate the origin of the IBV focus, showing that the IBV focus strategy has its precursor in a cleft construction. First I provide evidence on the mono-clausal properties of the IBV focus construction, arguing that it is no longer a cleft but has developed into a dedicated focus construction, though some biclausal characteristics still exist. Then I describe a possible way in which the IBV focus construction might have originated, illustrating each possible intermediate stage of this process.

4.4.1 IBV focus: mono-clausal or bi-clausal?

Talking about the origin of IBV focus, a natural question arises as why it should have an origin in some other constructions, rather than emerged independently. In other words, why diachronically the IBV focus construction should be considered to be innovated at a later stage than the cleft? In this section I show evidence that the IBV focus construction manifests many monoclausal properties, while some of its morphosyntactic features still reflect some residue of a relative/cleft, which implies that the focus construction originates from the latter.

To start, we first compare a pseudo-cleft and an IBV focus construction in (60). The main differences between the two constructions lie in the word order and agreement morphology on the verb. The pseudo-cleft in (60a) consists of a free relative and a predicative NP, as seen from the relative marker on the verb, the postposed subject and the predicative H tone on the nominal prefix. In the IBV focus construction in (60b), an apparent monoclausal property is the lack of (segmental) relative marking on the verb and the preverbal occurrence of the subject.

- (60) a. Wu-**kíí**-fúúm-í taará **múu**-ngwa.
 3REL-7SM.PST-buy-PST 1.father 3-salt
 ‘What father bought was some salt.’
- b. Taará **múu**-ngwa **káá**-fúum-i.
 1.father 3-salt 1SM.PST-buy-PST
 ‘Father bought some SALT.’

Prosodic evidence also suggests the monoclausal status of the IBV focus construction. The IBV focused element is always phrased together with the following verb without a phonological break. In (61) the focused object is phrased together with the following verb which starts with a vowel prefix, and vowel coalescence happens between the two adjacent vowels. In (62) the H tone plateauing effect is attested on the phonological domain that consists of the HL noun stem of the IBV focused object plus the H-toned SM on the verb. If (61) and (62) were actually biclausal cleft constructions, we expect a phonological break between the predicative focused NP and the free relative (Cheng and Downing 2007, 2013) to prosodically separate the two clauses. Therefore I take the conjoint phrasing of the focused element and the verb as evidence for the IBV focus construction to be monoclausal.

- (61) Me máa-lí á-m-fúum-i.
 /me máaláámfúumi/
 1SG.PRO 6-oil PST-1SG.SM-eat.PST
 ‘I bought some OIL.’

- (62) Ndé má-**láálá** **káá**-fúum-i.
 1.PRO 6-orange 1SM.PST-buy-PST
 ‘S/He bought some ORANGES.’ [*ma-láala* “oranges”]

Another crucial piece of evidence on the monoclausality of the IBV focus construction is that an SOV word order can be used to express VP focus, as illustrated in (63). In other words, focus in the IBV position can project onto the whole VP. Since the use of one focus marking for different scopes of foci is considered to be a monoclausal property (Jendraschek 2009, van der Wal and Maniacky 2015) and the focus interpretation expressed in a cleft can

never be extended to the VP, I take the VP focus expression in SOV in (63) as evidence for its monoclausal status.

- (63) a. (*What did father do in the morning?*)
 Ndé mí-fémé ká-dzwí.
 1.PRO 4-pig 1SM.PST-kill.PST
 ‘He [killed some pigs]_{VP focus}.’
- b. (*Did you wash the plates or do your homework?*)
 Me má-sáání á-n-swaak-í.
 1SG.PRO 6-plate PST-1SG.SM-wash-PST
 ‘I [washed the plates]_{VP focus}.’

The fact that an IBV focused element is mostly preceded by some other topical NPs may also indicate that this construction is more towards monoclausal, as shown in (64). An IBV focused subject/object/adjunct/infinite can always be preceded by multiple other primary and secondary topical elements (see chapter 3 section 3.3). Since a monoclausal focus construction usually allows fronting of other topical elements or modifiers while this is degraded in a biclausal cleft (Schwarz 2013; Abels and Muriungi 2008), the occurrence of multiple topics in the preverbal domain also suggests that the IBV focus construction is monoclausal rather than a cleft.

- (64) Ngúku lóoso munkí káá-dzí?
 1.mother 5.rice when 1SM.PST-eat.PST
 ‘When did mother eat the rice?’

However, there are still some residual properties of the cleft observed in the IBV focus strategy. Since a cleft always involves a predicative part and a relative clause, the SM alternation and the verb-final H tone that are only attested elsewhere in a (non-subject) relative, and the predicative H tone prefix on the IBV focused element could provide arguments for the cleft origin of IBV focus. In addition to these, there are also some other biclausal properties of the IBV focus construction. An example is that when negating the IBV focused element, a copula can optionally appear with the negative prefix

attached to it, which is very typical of the *it*-cleft construction, as shown in the negation on IBV subject focus (65a) and object focus (65b).

- (65) a. Ngwangúlu ka-kí-li mvá áá-dzí ni.
 1.gecko NEG-7SM-COP 1.dog 1SM.PST-eat.PST NEG
 ‘The gecko was not eaten by the DOG.’
- b. Me ka-kí-li báa-ntaba áá-m-fúum-i ni.
 1SG.PRO NEG-7SM-COP 2-goat PST-1SG.SM-buy-PST NEG
 ‘I did not buy the GOATS.’

Here one question is whether we should treat the affirmative and negative IBV focus sentences as the same constructions which only differ in polarity, or they are developed separately, namely the affirmative is more grammaticalised to become monoclausal, whereas the negative remains biclausal. I will come back to discuss this question in chapter 5 section 5.4.

In the negative question formation of IBV focus in (66) and (67), we see that the negation on the verb is expressed in the lexical strategy by using the word *ki-bía* ‘refuse’ as in (66a) and (67a), but the canonical negation strategy *ka...ni* that is used in matrix SVO clause is infelicitous as in (66b) and (67b), which is reminiscent of the ban on the use of the negation strategy *ka...ni* in relative constructions (see chapter 2 section 2.5.2). This can also show that the IBV focus construction is at least not completely mono-clausal but does maintain some grammatical properties of a relative clause.

- (66) a. Joní ku-ní káá-bí-í kí-ya?
 1John 17-which 1SM.PST-refuse-PST INF-go
 ‘Where didn’t John go?’
- b. ??Joní ku-ní ka-káá-yení ni?
 1John 17-which NEG-1SM.PST-go.PST NEG
Int: ‘Where didn’t John go?’

- (67) a. Ná áá-bí-í ki-mún-a we?
 1.who 1SM.PST-refuse-PST INF-see-FV 2SG.PRO
 ‘Who didn’t see you?’
- b. ??Ná ka-káá-mún-i we ni?
 1.who NEG-1SM.PST-see-PST 2SG.PRO NEG
Int: ‘Who didn’t see you?’

Another intriguing use of the IBV focus construction that is reminiscent of a cleft comes from example (68). The question in (68) intends to identify the thing that the child broke, which caused mother to beat the child. In (68) there is no (segmental) relative marking on the matrix verb, and the verb needs an argument which is expressed by the SOV order (in brackets). From the intended meaning we see that the SOV order in which a *wh*-object is focused in IBV can function as a cleft, and the whole sentence is also a cleft.

- (68) [Mwáana **kí-má** káá-búl-í] ngúku káá-béer-i?
 1.child 7-what 1SM.PST-break-PST 1.mother 1SM.PST-beat-PST
Int: ‘It was what that the child beat that (caused) mother to beat (him/her)?’

In summary, the IBV focus construction displays many monoclausal properties: 1) absence of the relative marker; 2) focus projection; 3) and conjoint prosodic phrasing. Nevertheless, it has still retained some characters of a biclausal cleft: 1) presence of the copula in negation; 2) H tone insertion on the SM; 3) verb-final H tone in non-subject extraction; 4) the H tone prefix on the IBV element which is suggestive of predication; and 5) unavailability of the *ka...ni* negation strategy. So this construction seems in a transitional stage where it has acquired some characteristics of a monoclausal structure and retains some biclausal properties. Apparently, the IBV focus construction involves less agreement relations (no relative marking) and less complicated clause structure than the cleft, based on which I propose that the IBV focus strategy has emerged at a later stage. Next I attempt to give the diachronic motivation for this grammaticalisation process, presenting the potential intermediate stages step by step.

4.4.2 From cleft to IBV focus: the grammaticalisation path

In this subsection, I discuss how the IBV focus construction may have grammaticalised from a cleft, illustrating what kind of semantic and syntactic changes are involved in different stages of this process. I start by looking into the structural and interpretational properties of the cleft construction, and track the development from it towards the monoclausal IBV focus construction by making reference to a mediating construction, namely the reduced cleft which places a focused element in the initial position. Then I investigate how the IBV focus strategy is ultimately innovated, discovering the motivation behind this diachronic development.

As illustrated in (69), clefts in Kukuya include the basic cleft (69a), pseudo-cleft (69b), and inverted pseudo-cleft (69c) (also see chapter 3 section 3.5). To start, it should be clarified which type of the three clefts the IBV focus is derived from. Here I first exclude the inverted pseudo-cleft in which the copula follows the predicate NP, since we have seen that when negating the IBV focus the copula always *precedes* the focused element, therefore in the original construction the copula should also precede the predicative NP. Deciding between the basic cleft and the pseudo-cleft which can be derived from each other, I take the basic cleft (69a) as the original construction, since in a basic cleft the focused NP is placed preverbally, so we don't need to additionally postulate a fronting rule for the predicative NP if we would consider the pseudo-cleft as the starting point.

- (69) a. (Kí-li) báa-ntaba ba-kíí-fúúm-í mú-kái.
 7SM-COP 2-goat 2REL-7SM.PST-buy-PST 1-woman
 'It was the GOATS that the woman bought.' [basic cleft]
- b. Ba-kíí-fúúm-í mú-kái (báá-li) báa-ntaba.
 2REL-7SM.PST-buy-PST 1-woman 2SM.PST-COP 2-goat
 'What the woman bought were the GOATS.' [pseudo-cleft]
- c. Báa-ntaba (báá-li) ba-kíí-fúúm-í mú-kái.
 2-goat 2SM.PST-COP 2REL-7SM.PST-buy-PST 1-woman
 'The GOATS were what the woman bought.'
 [inverted pseudo-cleft]

Syntactically, a basic cleft usually consists of two clauses: one contains a nominal predicate and the other contains a free relative clause, which are often linked by a copula. It should be noted that in a basic cleft the predicative NP does not have a focus function by itself, but the (exclusive) focus reading arises from the combination of the relative clause and the nominal predicate. The relative part of the cleft is presented as the maximal group of referents to which the predicate applies and is equated to the referent in the nominal predicate, and in this way an identificational and exclusive focus reading arises (van der Wal and Maniacky 2015).

In the next step towards developing into a focus construction, some bi-clausal properties of the cleft construction would be reduced and some characteristics of a monoclausal structure would emerge. In Harris and Campbell's (1995) studies on the universal changes from a cleft to a focus construction, they propose some indicators on the changes from biclausal to monoclausal, from which I list some relevant properties to Bantu languages, as shown in (70).

(70) *Changes biclausal > monoclausal* (Harris and Campbell 1995: 166, 167)

- dropping the copula or relativiser altogether
- reordering of constituents
- (re)introducing agreement according to monoclausal structure
- ceasing to use a special verb form

I suppose that the reduced cleft, as shown in (71), can reflect a diachronically intermediate stage in the development from the cleft to the IBV focus construction, though synchronically it co-exists with the two constructions. In the reduced cleft we see the deletion of the relative marker on the verb and reordering of the constituents. The postverbal subject in the cleft is fronted to a preverbal position in the reduced cleft, leaving the subject postverbal is ungrammatical, as in (72). Since the copula is always optional in affirmative nominal predication in Kikuya, so its absence in the reduced cleft may be irrelevant to the ongoing grammaticalisation. In the reduced cleft, the

focused element occurs sentence-initially and retains the same exclusive focus interpretation as in the basic cleft (also see chapter 3 section 3.4).

- (71) **Báa-ntaba mu-kái káá-fúum-i.**
 2-goat 1-woman LSM.PST-buy-PST
 ‘It was the GOATS that the woman bought.’ [reduced cleft]
- (72) ***Báa-ntaba káá-fúum-i mu-kái.**
 2-goat LSM.PST-buy-PST 1-woman
Int: ‘It was the GOATS that the woman bought.’

Similar constructions that place focus in the sentence-initial position are attested in many other West-Coastal Bantu languages and some other varieties of Teke, in many of which the IBV focus strategy is not observed or is only marginal. I suppose that in these languages the development of focus construction only attains the stage of a reduced cleft or an initial focus construction but the IBV focus has not emerged, which can also suggest that the reduced cleft is an intermediate stage which predates the IBV focus. I will return to present this in the next section.

We also notice that the class 1 subject agreement morphology in the reduced cleft (71) differs from the basic cleft (69a). In the basic cleft the postverbal class 1 lexical subject triggers the default SM *ki-* on the verb, whereas in the reduced cleft the preverbal class 1 subject is co-indexed as *ka-* on the verb, which does not fit into any agreement pattern we’ve seen so far. In all other cases where the class 1 subject is linearly adjacent to the verb, it takes the canonical SM prefix *a-*. So here it should be explained why the allomorph *ka-* is used and how this is related to the grammaticalisation of the focus construction.

I provide one possible account of the SM change in the reduced cleft. Recall that the *ka-* form also appears in a non-subject relative when the postverbal subject is a class 1 pronoun *ndé* (see examples (22) and (23) above). We may wonder whether the SM *ka-* in the reduced cleft can be associated with the postverbal pronominal subject. Here I provide support

from a related construction attested in the Nzadi language, which is a West-Coastal Bantu language spoken by a community of fishermen on the Kasai River in the Democratic Republic of Congo and has been classified as B865 in Maho (2009)'s updated list. In Nzadi non-subject relatives, while the subject is commonly placed postverbally as in (73a), a lexical subject NP can also appear in a preverbal position but only if there is an agreeing pronoun co-occurring immediately after the verb, as in (73b). This was described in Hyman (2012) as the VS/SVs alternation in this language. Similar to Kukuya, in Nzadi non-subject relatives, there is also a floating H tone occurring between the verb and the postverbal subject, which is absent in subject relatives and in the main clause. Hyman (2012) suggested the H tone to be a trace of a postverbal relative marker or pronoun, i.e. perhaps an older *SVs.

- (73) a. mwaán (na) (ŋg) o món **okáar**
 1.child that which PST see woman
 'the child that the woman saw' [Nzadi, VS] (Hyman 2012: 8a)
- b. mwaán (na) **okáar**_i o món **ńdé**_i
 1.child that woman PST see she
 'the child that the woman saw' [Nzadi, SVs] (Hyman 2012: 10a, index added)

Nzadi has developed a sentence-initial focus strategy, as exemplified in (74)-(76), where the non-subject *wh*-words occur in the initial position. The subject can be postverbal (74), or preverbal when there is the co-referring pronoun after the verb as in (75). In this sense, the initial focus strategy in Nzadi is analogous to the reduced cleft in Kukuya which involves a (non-subject) relative clause without segmental relative marking, but the SVs/VS alternation is still attested. It is noteworthy in (76) that in the presence of an initial focus, the subject can also be preverbal *without* the postverbal pronominal copy, which resembles the Kukuya reduced cleft in word order and may suggest a further grammaticalisation stage of an initial focus strategy than (75). In this stage, the postverbal pronominal copy, as a remaining indicator of the relative clause, is deleted. However, since in Nzadi there is a systematic lack of subject-verb agreement, the grammaticalisation pathway cannot be corroborated by its agreement morphology.

- (74) *ně ò món báàr?*
 who PST see people
 ‘Who did the people see?’ [Nzadi] (Hyman 2012: 107)
- (75) *oŋgér ŋge okáar_i o pé ndé_i bš?*
 thing what woman PST give she them
 ‘What did the woman give them?’ [Nzadi] (Crane et al. 2011: 10.78,
 index added)
- (76) *ně bààr ó môn?*
 who people PST see
 ‘Who did the people see?’ [Nzadi] (Hyman 2012: 107)

If the development of the initial focus construction in the two languages is indeed comparable, then the class 1 SM alternation in Kukuya may be explained by analogy to the Nzadi pattern. I conjecture that at some historical point there was also a VS/SVs alternation in Kukuya reduced clefts similar to (74) and (75), and the verb always agrees with the postverbal NP, namely with the S in VS and with s in SVs. Later the postverbal pronominal copy in SVs was deleted or became unpronounced for some reason, which is the same as in (76), but the class 1 SM which once agreed with the postverbal pronoun s was retained. One question here is why synchronically there is no VS/SVs alternation in non-reduced clefts and non-subject relatives. It seems that the deletion of the relative marker in the reduced cleft provides the necessary condition for the fronting of the subject, which is not applicable in the non-reduced version. Another question is how to account for the class 1 SM alternation with regard to different positions of the subject NP from the synchronic point of view, and which structural position the preverbal subject occupies in the reduced cleft. These are discussed in the next chapter (chapter 5 section 5.4.1).

The development process presented above can be summarised as in (77).

- (77) *Stage I: Basic cleft > reduced cleft*
 (COP)-NP_[PRED] REL-SM-Verb-SUBJ \implies
 NP_[PRED] SUBJ_i SM-Verb-PRO_i \implies
 NP_[PRED] SUBJ SM-Verb

I suppose the next step in the grammaticalisation to be a “hypoanalysis” process, in which the listener reanalyses a contextual semantic/functional property as an inherent property of the syntactic unit (Croft 2000: 126). The focus reading in the reduced cleft is generated from the whole construction, and at some later point the focus interpretation must have been linked to the predicative NP itself rather than to the whole sentence. In other words, the initial NP has gained a [FOCUS] feature, and the new construction consists of one clause with one verb and a focused constituent, rather than being a combinational focus construction. This hypoanalysis and the introduction of the [FOCUS] feature are also crucial conditions in the transition from a biclausal to a monoclausal focus construction.

To derive the IBV focus construction in (78), the question here is why the focused constituent requires verb-adjacency. Especially, why didn't the reduced cleft just develop towards a monoclausal initial focus construction but instead evolved into IBV focus, since they have similar functions of (exclusive) focus expression? Why should the preverbal subject move further to the sentence-initial position when the object is in focus?

- (78) Mu-kái báa-ntaba káá-fúum-i.
 1-woman 2-goat 1SM.PST-buy-PST
 ‘The woman bought some GOATS.’ [IBV focus]

In Bostoen and Mundeke's (2011) analysis of the functional passive OSV construction in Mbuun (B87), in which the subject is focused in IBV position, they claim that the patient NP is fronted in order to “make the focused agent NP less topical”. I agree with their proposal in that topic fronting is an important factor in deriving the word order. I also propose that, to place the focused element in the IBV position is a further step towards monoclausality,

in which the IBV focused element becomes clause-internal. This step also creates the necessary precondition for further reanalysis such as focus projection and pragmatic neutralisation. The exclusive focus reading inherited from the original cleft sentence was retained in the early stage of reanalysis, and becomes pragmatically neutral later, namely the focus expression is not necessarily exclusive/contrastive (see chapter 3 section 3.2.4). The development from the reduced cleft to the IBV focus construction is schematised in (79).

- (79) *Stage II: Reduced cleft > IBV focus*
 NP_[PRED] SUBJ SM-Verb \implies
 NP_[FOC] SUBJ SM-Verb \implies
 SUBJ_[TOP] [NP_[FOC] SM-Verb]

In the above analysis I have presented a plausible grammaticalisation path of non-subject focus in the IBV position. As for subject focus, a similar grammaticalisation process starting from a subject cleft can be proposed, in which the relative marking deletion, hypoanalysis, and topic fronting also occur and ultimately the subject is focused in the IBV position. The whole pathway is summarised in Table 4.9 with illustrative examples.

Structure	Scheme	Example
Basic cleft	(COP)-NP _[PRED] REL-SM-Verb-SUBJ	(Kí-li) báa-ntaba ba-kíí-fúúm-í mú-kái.
Reduced cleft	NP _[PRED] SUBJ SM-Verb	Báa-ntaba mu-kái káá-fúum-i.
IBV focus	SUBJ _[TOP] [NP _[FOC] SM-Verb]	Mu-kái báa-ntaba káá-fúum-i.

Table 4.9: Grammaticalisation pathway of the IBV focus construction

One further question here is why the language simultaneously develops IBV focus for both subject and non-subject. Since in an SVO language the subject is often the default topic (Lambrecht 1994), subject focus is usually more marked than non-subject focus, for example in Kukuya a subject question is often expressed in a pseudo-cleft (see chapter 3 section 3.4). If so, why would subject focus be derived from a reduced cleft and to occur in the preverbal position again? How does this process void the rigid constraint against the preverbal subject to be focal as in many other Bantu languages (Morimoto 2000; Zerbian 2006; van der Wal 2009, 2015;

Downing and Marten 2019)? Is there any kind of subject-object asymmetry overlooked here? One possible approach would be that once a dedicated focus position was established in the language, it can just attract the subject to be focused there. I will continue to discuss this in the next chapter.

In this section I investigated the grammaticalisation process of the IBV focus construction. I first showed that there is a mixture of monoclausal and biclausal properties in the IBV focus strategy. Then I presented a possible grammaticalisation pathway from the biclausal cleft to a monoclausal focus construction. I proposed that in this process, relative marker deletion, hypoanalysis and introduction to a [FOCUS] feature, and topic fronting rules occur in sequence, resulting the IBV focus construction, which also shows that Kukuya has moved from more syntax-configurational to more discourse-configurational. Another question is where to situate the cleft origin of the IBV focus strategy, it would be interesting to investigate whether it is an independent development in Kukuya or an older ancestral evolution inherited in Kukuya and its closest relatives.

4.5 Microvariation in preverbal focus in West-Coastal Bantu

In this section, I shed some light on the IBV focus in some other West-Coastal Bantu languages to see if any generalisation on the origin of this focus strategy can be made, and what types of microvariation can be observed.

The class 1 subject marking alternation is also attested in many other West-Coastal Bantu languages neighbouring to Teke. For example in Mbuun (B87), which also employs the IBV focus position, the *á-* versus *ká-* alternation is attested in subject focus and non-subject focus in the past/perfective tense as well as in some other tenses/aspects (Bostoen and Mundeke 2012). Where it differs from Teke is that in Mbuun the canonical class 1 SM in the past/perfective tense is *ká-* when there is no focused argument, as in (80a), while it keeps the form *ká-* when the object is focused (80b) and shifts to *á-* when the subject is focused (80c). In other tenses/aspects, the canonical subject marker is *á-* and shifts to *ká-* in the context of object focus. For adjunct focus in the IBV position, the SM alternation only optionally occurs. The OSV functional passive construction in Mbuun does not only involve object topicalisation but the fronted topical element also shows many subject properties, which suggests that Mbuun is in a further stage of grammaticalisation of the IBV focus construction (Bostoen and Mundeke 2011).

- (80) a. *ngwén ká-wó-kér i-sal ka kwil*
his.mother ISM-PST-do 7-work LOC Kikwit
'His mother worked in Kikwit.'
- b. *mo-an ná ká-mwén-ii?*
1-child which ISM-see-PERF
'Which child did she see?'
- c. *ná á-wéén le ndza?*
who ISM-go.PERF with him
'Who has accompanied him?'

[Mbuun B87] (Bostoen and Mundeke 2012)

In Kisikongo (H16a) which has also been reported to have the IBV focus position, the class 1 SM takes a null form when there is no argument focus, as in (81a); in (81b) the adjunct is focused preverbally, and the SM appears as *ka-* (De Kind 2014). The same class 1 SM *ka-* is used in SOV and non-subject relatives, alternating with other SMS in SVO or subject relative clauses, which can also show the connection between IBV focus and cleft.

- (81) a. *yandi Ø-zol-idi zay-a vo kuna N-banza Kongo*
 1.PRO 1SM-want-PERF know-FV if 17.DEM 9-city Kongo
ma-dyoko tu-lamb-ang-a
 6-cassava 1PL.SM-cook-IMPF-FV
 ‘She wants to know if we prepare cassava in Mbanza Kongo.’
- b. *o-Ø-se ve ka-vat-idi?*
 1AUG-1SM-father where 1SM-cultivate-PERF
 ‘WHERE did the father cultivate?’

[Kisikongo H16a] (De Kind 2014)

In Hungan (H42), the class 1 subject marker surfaces in the form *a-* when expressing subject focus (82a), and appears as *ká-* when an object or adjunct is focused (82b,c). It is also noteworthy that in Hungan, argument or adjunct focus does not trigger word order change, as we see in (82b) that the focused object is placed in its canonical postverbal position rather than fronted to a preverbal position (Takizala 1972, 1974).

- (82) a. *Kipés á-swíim-ín kit zóónó.*
 Kipese 1SM-buy-PST 7.chair yesterday
 ‘KIPese bought a chair yesterday.’
- b. *Kipes ká-swíim-ín kít zóónó.*
 Kipese 1SM-buy-PST 7.chair yesterday
 ‘Kipese bought a CHAIR yesterday.’

- c. Kipes **ká**-swíim-ín kit zoon.
 Kipese 1SM-buy-PST 7.chair yesterday
 ‘Kipese bought a chair YESTERDAY.’
 [Hungan H42] (Takizala 1972: 11-13)

The grammaticalisation process synchronically shows different levels reached in the various languages in West-Coastal Bantu and even within the Teke cluster itself. In Tege (B71) (Linton 2013) (non-subject) *wh*-words are commonly placed postverbally, as shown in (83a); in (83b) the *wh*-word occurs sentence-initially and the subject is inverted; in (83c) the object *wh*-word is placed in the IBV position; the subject is focused preverbally in (83d); in (83e) adjunct focus occurs in the IBV position and the topical object is fronted to the preverbal domain.

- (83) a. Taará **ká**-lag-a na?
 1.father 1SM.PROS-talk-FV who
 ‘Who is father going to talk with?’ [object focus]
- b. **Ngondo ó**-ma **ká**-ye taará Ngabon?
 1.month 1AGR-which 1SM.PROS-go 1.father Libreville
 ‘In which month will father go to Libreville?’ [adjunct focus]
- c. Mbali brĩ **é**-má lê-yíríg-a?
 tomorrow 1PL.PRO 8-what 1PL.SM.FUT-teach-FV
 ‘What will we teach tomorrow?’ [object focus]
- d. **O-ngébé ó**-má â-bíl-á bila?
 1-child 1AGR-which 1SM.FUT-bring-FV 8.food
 ‘Which child will bring the food?’ [subject focus]
- e. We a-mbílí **kákuní** â-sur-a?
 2SG.PRO 6-food where 2SG.SM.FUT-deposit-FV
 ‘Where will you deposit the food?’ [adjunct focus]
 [Teke-Tege B71] (Linton 2013: 5-8, glossing adapted)

In Teke-Boma (B74), the IBV focus strategy is also observed, as shown in the sentences in (84). In these examples subject and object focus are placed in the IBV position, and we see that a H tone occurs on the nominal prefix or the first stem syllable of the focused NP.

- (84) a. **Bihí bá-káhá** líi-luó.
 1PL.PRO 2-wife 1PL.SM.PST-teach
 ‘We taught the WIVES (not the husbands).’ [object focus]
- b. **Wé aá-béére** ηomo.
 2SG.PRO 2SG.SM.PST-beat 9.drum
 ‘YOU (not me) played the drum.’ [subject focus]
- c. **We ηómo** aá-béére.
 2SG.PRO 9.drum 2SG.SM.PST-beat
 ‘You played the DRUM (not the guitar).’ [object focus]
 [Teke-Boma B74] (Raharimanantsoa: p.c., glossing added)

In Iyaa (B73, Mouandza 2001) and Fumu (B78, Makouta Mboukou 1976) *wh*-words occur sentence-initially but are not seen in the IBV position. Two interrogative sentences of Iyaa are given in (85).

- (85) a. **Ná we** món-i?.
 1.who 2SG.PRO see-PST
 ‘Who did you see?’ [object focus]
- b. **Bû-ni we** díbili i-ku:ku:?.
 14-which 2SG.PRO cook-PST 7-meal
 ‘How did you cook the meal?’ [adjunct focus]
 [Teke-Iyaa B73] (Mouandza 2001: 323, 326, glossing added)

The examples above show that the grammaticalisation process of the IBV focus strategy is in different stages in West-Coastal Bantu languages. Notably, we see that the languages that display IBV focus all make use of the initial focus strategy, but no language only employs the IBV focus position alone, which can corroborate the hypothesis that the initial focus, namely the reduced cleft construction, is prior to IBV focus in this process.

In Table 4.10 below, I list several checkpoints on some grammatical features in Mbuun, Kisikongo and Kukuya that are generalised from some available corpus and literature. It is interesting to see that even within the three languages that belong to the West-Coastal Bantu and that all have been reported to have the IBV focus strategy, there is much variation on

the other grammatical properties which are associated with IBV focus. It is worthwhile investigating if there are any dependencies or correlations among these features, and why some constructions especially subject inversion is allowed in languages like Kisikongo and Tege (see (83)b above) but not in the others.

	Mbuun	Kisikongo	Kukuya
object marking	✓	✓	✗
subject inversion	✗	✓	✗
augment	✗	✓	✗
use of focus particle	✓	✓	✗
focus projection	✗	✓	✓
<i>in situ</i> object focus	✗	✗	✓
anti-agreement effect	✗	✓	✗
preverbal subject in non-subject relatives	✗	✓	✗
agreement with postverbal lexical subject	✗	✓	✗
subject properties of the fronted topic	✓	✗	✗
class 1 SM <i>ka-</i> in SVO	✓	✗	✗

Table 4.10: Some microvariation with regard to the IBV focus strategy in three WCB languages

* * *

This chapter is dedicated to investigating the possible origin of the IBV focus construction. In the first two sections, I introduced the connections between the IBV focus and relative/cleft constructions on the SM alternation and tonal variation including the H tone on the SM, predicative H tone on the focused element, and the verb-final H tone, corroborating the hypothesis that the IBV focus strategy is very likely to have been derived from a basic cleft. I also showed that the IBV focus construction exhibits both monoclausal and biclausal properties, and proposed a grammaticalisation process of this focus strategy. In the next chapter, I will give a fine-grained analysis on the structural derivation of the IBV construction

from the synchronic point of view, and will also provide an explanation on the mechanism of the class 1 subject marking allomorphy.

CHAPTER 5

The syntax of IBV focus and subject marking asymmetry

In the previous chapter, I proposed that diachronically the IBV focus construction has its origin in a cleft sentence, relying on evidence from the H tone prefix on the focused NP, segmental and tonal SM alternation as well as the verb-final H tone in both constructions. I argued that while bi-clausal (pseudo-)cleft constructions are still used in Kukuya for the expression of focus, an IBV focus strategy has been innovated from the cleft and has been grammaticalised *towards* a monoclausal construction dedicated for focus marking. The grammaticalisation process can be observed from the synchronically co-existing constructions shown in (1).

- (1) a. (Kí-li) báa-ntaba ba-kí-fúúm-í mú-kái.
7SM-COP 2-goat 2REL-7SM.PST-buy-PST 1-woman
'It was some goats that the woman bought.' [cleft]

- b. Báa-ntaba mu-kái ká-fúum-i.
 2-goat 1-woman 1SM.PST-buy-PST
 ‘(It was) some goats that the woman bought.’ [reduced cleft]
- c. Mu-kái báa-ntaba ká-fúum-i.
 1-woman 2-goat 1SM.PST-buy-PST
 ‘The woman bought some GOATS.’ [IBV focus]

I proposed that the diachronic derivation starts from a basic cleft construction like in (1a) in which the relative clause and the predicative part of the sentence can be clearly distinguished. In (1b) the relative marker on the verb is deleted, which is accompanied by the fronting of the postverbal subject to a preverbal position, and the focus interpretation in (1b) is restricted to the sentence-initial NP. One step further in the derivation is seen in (1c) where the preverbal subject has moved to a higher position preceding the focused object, while the previously extra-posed focused object is now “integrated” into the predicate and becomes a clause-internal constituent, and this is the point where we arrive at the IBV focus construction.

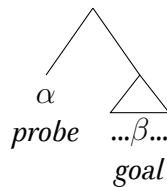
There are many questions on the constructions in (1) that were not explained in the previous chapter: for example why in (1b) the deletion of the relative marker triggers the fronting of the subject and the subject cannot just be left postverbal; what motivates the subject in (1c) moving higher and where is this higher position; and if we consider the class 1 subject agreement on the verb in (1), we may wonder why subject marking differs with regard to whether the subject is preverbal or postverbal. In addition, since we have seen that there is still some “trace” of relative marking on the verb in the IBV focus construction, which I took as evidence for its connection with the cleft construction, we want to know if there is still any influence from the original structure on the synchronic syntactic structure of the innovated structure. All these questions need to be considered and explained in a synchronic account by looking into the syntactic structure of each construction in (1), especially the IBV focus construction, and this is what this chapter is dedicated to.

In this chapter, I follow the generative Minimalist generative approach to account for the syntax of IBV focus and especially the subject agree-

ment alternation. Here I briefly introduce some relevant terms for the readers who are less familiar with this framework. Subject agreement has been captured in the Minimalist framework by the syntactic operation Agree (Chomsky 2000, 2001), as stipulated in (2). Under Agree, a head and a phrase share features which include ϕ -features: [Person], [Number] and [Gender]. The Agree relation is initiated by a head with unvalued ϕ -features, which is usually referred to as the probe. The probe searches in the c-command domain which is the derivation that has been built up so far; when it encounters an element which has the feature specification that it is looking for, this element counts as a matching goal for the probe and an Agree relation is established between the probe and the goal. The morphological representation of an Agree relation can be a prefix on the verb, for example the subject and object markers in the Bantu languages.

- (2) a. Agree (α, β) if α c-commands β ; α, β have matching features; there is no γ with matching features such that α c-commands γ and γ c-commands β .

b.



This chapter is organised as follows. In section 5.1, I give an account of the structural representation of SVO sentences, investigating the structural position of the verb and the preverbal grammatical subject. In section 5.2, I discuss the structural position of the IBV focus site. I first diagnose the syntactic status of the IBV focused element, discovering whether it is base-generated or it has undergone movement; and if it has moved, what kind of syntactic movement is involved. Then I show that IBV focused subject and non-subject are placed in the same structural position and discuss whether this corresponds to a high or low FocP in the hierarchy. Section 5.3 concentrates on the class I subject marking alternation. First I deviate a bit from IBV focus to discuss the subject agreement asymmetry between preverbal and postverbal subjects in subject and non-subject relative constructions which are closely associated with the IBV focus construction. I adopt the defect-

ive goal approach to provide a featural account of the relativised agreement pattern. I propose that in a non-subject relative, the SM can only be spelled out as an agreement marker when T agrees downwards with a DP that has an extra [Person]-layer, namely the person pronouns. [Gender] is located on an inner *n*P that selects a lexical root thus is not visible to T under Minimal Agree. In section 5.4, I first illustrate the structural representation of the cleft constructions and the derivation of the IBV focus construction. I show that in the IBV focus strategy, at the start of the derivation a ϕ P with a 3rd person value is merged as the external argument in the *v*P, and the topical element that is co-referential with this 3rd person ϕ P is base-generated in the left periphery. The class 1 subject marking alternation actually corresponds to fundamentally different structures. Some remaining questions for further research are highlighted at the end of the chapter.

5.1 Structural representation of SVO

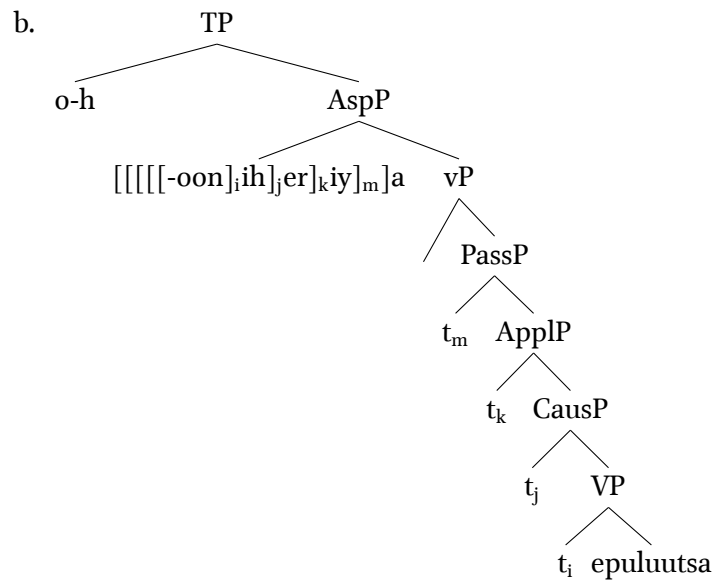
In this section, I investigate how the derivation of the canonical SVO word order should be structurally represented. I first look into the position of the verb, arguing that the verb in Kukuya may stay in different height in different TAM configurations. Then I discuss the structural position of the grammatical subject in SVO, showing that an preverbal A-position is available for hosting the subject, which should be specFinP or specTP.

5.1.1 Position of the verb

For the sake of investigating the structural positions of the clausal arguments, first we need to know the position of the verb in the hierarchy. Here I follow the analysis in Myers (1990), Kayne (1994), Julien (2002), Kinyalolo (2003), Buell (2005) and van der Wal (2009, 2022) to assume that the verb in Bantu languages starts out as a lexical base and head-moves to a higher inflectional position such as Aux or T (Ngonyani 2000; Wasike 2007; Zeller 2013), or it just ends up in a position lower than T such as Aspect or Mood (Julien 2002; Kinyalolo 2003; Carstens 2005; Van der Wal 2009), while incorporating the derivational and inflectional suffixes *en route* by left-adjunction morphology. The inflectional prefixes such as subject marker and TAM markers are instead spelled out in their base positions and form a single word with the verb stem through phonological merger.

To illustrate, an example sentence from Makhuwa (P31) and its derivation (van der Wal 2022: 35) are given in (3). The verb root *-oon-* “to see” head-moves to CausP and incorporates the causative suffix to its left: *-oon-ih-*. This merged complex head then moves to ApplP, PassP and vP step by step and adds the suffixes to its left *en route*, and finally the verb complex terminates in AspP where it takes the final vowel. The suffixes surface in exactly the reverse order of the structural hierarchy. The subject marker and the TAM marker as prefixes, are spelled out in their individual positions in the inflectional domain. The whole verb complex *o-h-oon-ih-er-iy-a* is formed through phonological merger.

- (3) a. Nlópwáná o-h-oón-íh-er-íy-a epuluútsá.
 1.man 1SM-PFV.DJ-see-CAUS-APPL-PASS-FV 9.blouse
 ‘The man was shown the blouse.’ [Makhuwa P31] (Van der Wal 2022: 35)



Now we consider how this head-movement account is applied to the verb in Kukuya. As the language has lost all its verbal derivational suffixes but only maintains some unproductive remnants, I assume that there are no projections to host any verb extensions in the lower part of the vP. Therefore, in order to know which position the verb head-moves to, we only need to consider what the final vowel (FV) of the verb stem encodes. However, the exact function(s) of the FV in Kukuya is not as clear as in many other Bantu languages, which may be due to the historical loss of some inflectional suffixes.

As introduced in chapter 2 (see section 2.4.2) on TAM marking, the FV appears as -a as its default form in infinitives and non-past tenses when the stem vowel is a, i or u, as a vowel copy suffix when the stem vowel is e or o, and it occurs as -i in all past tenses regardless of the quality of the stem vowel, as shown in (4).

- (4) a. *ki-bhúima* “to empty”
ki-kúula “to be older”
ki-sóo “to search”
ki-téke “to sell”
ki-khée “to try”
ki-lheeme “to become clear”
ki-bolo “to rot”
- b. Mwáana áá-tsúúl-í mú-fiémé mu ki-wá
1.child ISM.PST-make.fall-PST 3-bottle 18.LOC INF-give
ndzulí buokó.
1.cat 14.fear
‘The child made the bottle fall to scare the cat.’

Therefore, it seems that the FV alternation in Kukuya is used to encode the past/non-past tense distinction, functioning together with the vowel lengthening tense prefix and grammatical tones like a circumfix. This is also what Nurse (2008) documented for some Bantu languages including Zone B which exploit the combination of two different positions such as TA and FV to carry one tense reference, where he notes that it is often “impossible to split the functions” of the two positions (Nurse 2008: 81). However, in some grammar sketches of other Teke varieties, the FV is usually glossed as an aspect marker whereby -a and the vowel copy prefix encode imperfective and -i encodes perfective (Mouandza 2001 on Iyaa B73; Etsio 1999 and Raharimanantsoa 2012, 2017, 2020 on Eboo B74), or glossed as ambiguous between tense and aspect (Tsoue 2017 and Linton 2018 on Tege B71a; Calloc’h 1911 and Makouta-Mboukou 1976 on Fumu B77b). For Kukuya, in Paulian (1997: 213, 2001) she labelled the FV -a as MA “*marque d’aspect*” and -i as MTA “*marque de temps et d’aspect*”.

An apparent counterargument for the FV in Kukuya to be an aspect marker is that when an aspect *prefix* occurs, whether the prefix is the perfective/resultative *maá-* or the imperfective/durative *yi-*, the FV can only take the default form -a or the vowel copy suffix but can never be -i, which shows that the FV alternation does not function to distinguish aspects. However, this co-occurrence constraint on the aspect markers and the FV may be explained as that diachronically the aspect prefixes such as *maá-*

and *yi-* were very likely to have been grammaticalised from lexical verbs *ki-mana* “to finish” and *ki-ye* “to go”, then they were phonologically reduced and fused together with a following *infinitive* verb stem, which may have an impact on the synchronic realisation of the FV on the verb complex that it can only take the unmarked *-a* as in infinitives.

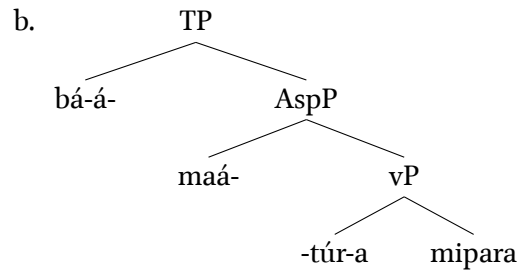
In addition to these inconsistent analyses on whether the function of FV in Teke is more towards tense or aspect marking, it should also be noted that the FV alternation is also attested in conditional and subjunctive clauses (see also chapter 2 section 2.6). One example is given in (5), in which the verb in the subjunctive clause takes the FV *-i* which is not related to past tense marking, so I suppose that the FV can also be used to encode the subjunctive mood.

- (5) Me ki-n-yáab-a pirí me bú-kíá
 ISG.PRO COND-1SG.SM-know-FV ISG.COMP ISG.PRO 14-tomorrow
 â-ŋ-kwá, me n-ték-i bi-lóko bvíí me
 FUT-1SG.SM-die ISG.PRO ISG.SM-sell-SBJV 8-thing 8.CONN ISG.PRO
 bvheí lo.
 8.all today
 ‘If I know that I will die tomorrow, I would sell all my things today.’

Based on the facts introduced above, I propose that the FV in Kukuya, at least for its modified form *-i*, should not be treated as a single suffix with one unique function, but may have multiple homophonic counterparts which may have originated from different historical suffixes, whose functions are not confined to tense (and possibly aspect) but also express mood marking. The suffix *-i* used in the past tenses may originally have represented perfective or anterior, and since the semantic shift from anterior to past is commonly attested cross-linguistically, it can result in the ambiguity on whether the FV encodes tense or aspect or both (Nurse 2008: 95). For simplicity, and also due to the fact that in Kukuya the alternation of FV is overwhelmingly used to encode past/non-past distinction, in this chapter I treat the FV as a tense suffix when aspect and mood are not marked in a sentence.

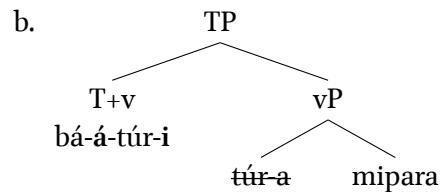
In (6) I propose that in the derivation of a basic SVO sentence, when there is an AspP that hosts a segmental aspect prefix, the verb just moves from V to v where it takes the default FV *-a*, and the perfective aspect prefix *maá-* is attached to the verb by phonological merger.

- (6) a. Bó báá-maá-túr-á mí-pará.
 2.PRO 2SM.PST-PFV-steal-FV 4-money
 ‘They have stolen the money.’



When the FV encodes past tense as in (7), I suppose that the verb head-moves to the head of TP which is just above the vP, incorporating the past tense suffix *-i* and the vowel lengthening prefix there, and here the AspP is not relevant. In (6) and (7) only the derivation of the verb complex is displayed, while the structural position of the subject and the place of subject agreement will be discussed later.

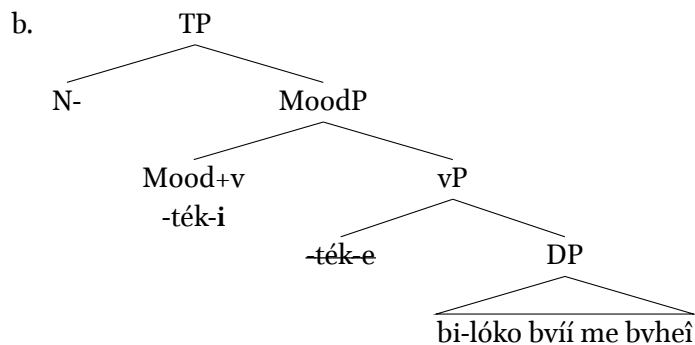
- (7) a. Bó báá-túr-i mi-pará.
 2.PRO 2SM.PST-steal-PST 4-money
 ‘They stole the money.’



The subjunctive clause in (5) is repeated in (8a), and its derivation is shown in (8b). When the FV is used for encoding subjunctive mood, the verb first moves to the head of a MoodP just above vP, incorporating the suffix *-i*. Since

there is no direct evidence on whether the verb moves higher to T/Asp or stays in Mood in a next step, I assume that the verb terminates at Mood in a subjunctive clause.

- (8) a. ...me n-ték-i bi-lóko bvíí me bvheí.
 1SG.PRO 1SG.SM-sell-SBJV 8-thing 8.CONN 1SG.PRO 8.all
 ‘...I would sell all my things.’



Now we see that the height of the verb in Kukuya is variable, and three situations can be distinguished: 1) in a non-past tense or when there is an overt aspect prefix, the verb stays in its base position within the vP; 2) when the FV vowel is marked for encoding past tense, the verb head-moves to T to incorporate this tense suffix; and 3) when the FV is used to express subjunctive mood, the verb moves to the head of MoodP and stays there. Although the exact function of the FV is still unclear, which may blur the structural position of the verb, and there are puzzles such as how the verb can incorporate *both* the tense prefix and suffix when it moves to T, we will see in the rest of the chapter that the exact landing site of the verb in any functional head between TP and vP in SVO is of little influence on the analysis of the IBV focus construction. For consistency, throughout the chapter I display the AspP and MoodP only where necessary. In addition, I do not take the split-InflP account¹ (Pollock 1989; Chomsky 1993) to locate the subject marker and tense marker in separated projections Agr(S)P and TP, but I treat T as the head hosting both the φ -features that are responsible for subject agreement and the tense features as shown in the examples above, in which the

¹For presentation, I still treat NegP and AspP as separate projections, which is inline with a split InflP.

subject prefix and the tense prefix are always phonologically merged and often inseparable. Next we consider the structural position of the grammatical subject in SVO.

5.1.2 Position of the subject in SVO

In this subsection, I discuss the structural position of the preverbal subject in SVO sentences in Kikuyu. From a cross-linguistic point of view as well as in the literature of Bantu syntax, a preverbal subject in SVO can occupy different structural positions. Subject marking in Bantu languages can represent *agreement* with a clause-internal preverbal subject DP in an A-position such as specTP, specAgrSP or specFinP; and the subject marker can also be *pronominal*, with a co-indexing subject DP functioning as the topic of the sentence which occupies a dislocated A'-position such as specTopP. In generative syntax, an A(argument)-position refers to a structural position to which a theta-role can be assigned, which the grammatical subject and object(s) of a sentence usually occupy; an A'-position is a position that is not occupied by an argument. The two English sentences in (9) can help understand this distinction: in (9a) the verb agrees with the grammatical subject “my dog”, which is in an A-position, and takes the 3rd person singular suffix -s which corresponds to the agreeing subject marker in Bantu; while in (9b) the grammatical subject is “they” which is co-referential with the dislocated topic “the children” that is in an A'-position, and is comparable to the pronominal subject marker.

- (9) a. My dog waits for me at home every day.
b. The children_i, they_i got wet in the rain.

In the Bantu literature, the preverbal DP in languages such as Bembe (D54, Iorio 2014) is analysed as always being in an adjoined A'-position and the subject marker is pronominal (also see Givón 1976); while for languages such as Kĩtharaka (E54, Muriungi 2008), Bemba (M42, Mwansa 2011) and Makhuwa (P31, van der Wal 2009), the subject DP can be non-dislocated and situated in an A-position in specTP or specFinP. I argue that in Kikuyu the preverbal DP in SVO can always occupy an A-position, which I consider

to be specFinP, and consequently the SM functions as a marker of syntactic agreement. Note that here I am not defining the structural position for the grammatical subject in all word orders but only in SVO. I will discuss other word orders such as SOV and OSV in section 5.2 and 5.4. I provide my arguments below.

1) One obvious argument for the preverbal subject in SVO to be in an A-position is that, the preverbal subject in Kukuya is in most cases overtly expressed and is only seen omitted in some casual speech, while the subject marker rarely stands by itself without an overt subject DP. The preverbal subject NP and the SM can often phonologically fuse and vowel coalescence can occur when the SM starts with a vowel. This can be taken as evidence that the subject marker functions as an agreement marker rather than as a pronoun, and the overt subject is the true argument of the verb thus cannot be absent.

There are also exceptional cases in which there is no overt subject DP but only the subject marker, notably when the subject is a speech participant, as in the first part of the sentence in (10). In this case I assume that the true subject is a covert *pro* with which the 1st person plural SM agrees.

- (10) [Li-dzwá nyama wúa], biáwe ndé ku-ní
 IPL.SM-kill 1.animal 1.DEM.II IPL.PRO 1.PRO I7-which
 líi-kab-a?
 IPL.FUT-share-FV
 '(As) we kill that animal, where will we share it?'

2) Another crucial diagnosing point is that an indefinite and non-specific subject which is often considered impossible to be dislocated (Rizzi 1986; Cinque 1990; Baker 1996, 2003), can occur in the preverbal position in SVO, as shown in (11). Since Kukuya has lost the augment which often functions as a determiner in many other Bantu languages (de Blois 1970; Van de Velde 2019; Halpert *to appear*), the indefinite interpretation can only be detected in a restricted number of contexts. The example in (11) is actually athetic expression, and I have shown that athetic sentence is usually expressed via the SVO order (see chapter 3 section 3.1.3), in which the preverbal subject is

non-topical/non-focal thus cannot be dislocated. These facts show that an indefinite non-specific subject can be placed preverbally, so there should be at least one A-position in the preverbal domain.

- (11) (*You returned home and found some footprints on the floor; you say to your roommate:*)

Mbuurú (nguumó) á-yení.
 1.person 1.one 1SM.PST-come.PST
 ‘Someone came.’

3) NPs modified by strong quantifiers such as “all” and “every” are also observed as impossible to be dislocated (Rizzi 1986; Baker 1996). Zerbian (2006) and Zeller (2008) also showed that in Zulu and Northern Sotho strongly quantified DPs are not dislocated. In examples (12) and (13) we see that the subject NPs modified by *ná* “every” and *bwě* “all (cl.2)” can both occur in the preverbal position in SVO, which again indicates that there is an A-position in the preverbal domain to host the grammatical subject.

- (12) **Ná ndúku** aa me á-tó-i.
 every 1.friend 1.CONN 1SG.PRO 1SM.PST-arrive-PST
 ‘Every friend of mine arrived.’

- (13) **Báana bhoi** báá-bák-ĩ baa-bonbon.
 2.children 2.all 2SM.PST-get-PST 2-candy
 ‘All the children got the candies.’

4) The agreement pattern in relative clauses can also help diagnose the syntactic status of the subject marker and the lexical subject. In (14a) we see that the 1PL SM co-occurs with the postverbal subject, while in (14b) the class 2 postverbal subject cannot be indexed by the class 2 SM on the verb, but only a class 7 default SM can occur in the SM slot (also see section 5.3). Even if the postverbal subject *baa-kái* is elided, the class 2 SM cannot be used for referring to the “women”, so there is no complementary distribution of the SM and the subject DP here. This differential subject marking strategy also

shows that the subject marker is actually a grammatical agreement marker rather than a referential pronominal marker, otherwise we would expect that subject markers of all noun classes are available for this slot for pronominal use.

- (14) a. mi-féme mi-**lí**-li **lí**-fúúm-í **bhí**
 4-pig 4REL-1PL.SM.RPST-COP 1PL.SM.RPST-buy-PST 1PL.PRO
 ‘the pigs that we had bought’
- b. mwáana wu-***báá/kí**-kú-í **báa-kái**
 1.child 1REL-2SM.PST/7SM.PST-bring.up-PST 2-woman
 ‘the child that the women brought up’

Based on the evidence 1)-4) above, I conclude that the subject marker in Kukuya should be analysed as an agreement category rather than a pronoun, and the preverbal subject NP can be non-dislocated in an A-position in SVO order.

The preverbal subject is elsewhere most commonly attested as the topic of the sentence, which is consistent with the claim made by Givón (1983), Gundel (1988) and Henderson (2006) (on Bantu) on the tendency of the grammatical subject to have a discourse topic function, as shown in example (15a). In (15a) the subject is the topic of the sentence because it is mentioned in the context question. Note that (15a) should be distinguished from (15b): in the latter example *taará* is a dislocated topic, and the grammatical subject is the class 1 pronoun *ndé*.

- (15) (*What did father sell yesterday?*)
- a. **Taará** á-ték-i baa-ntaba.
 1.father 1SM.PST-sell-PST 2-goat
 ‘Father sold some GOATS.’
- b. **Taará**, ndé á-ték-i baa-ntaba.
 1.father 1.PRO 1SM.PST-sell-PST 2-goat
 ‘(As for) father, he sold some GOATS.’

A preverbal subject can also be focal, as shown in (16) where the subject is

focused as the answer to a subject *wh*-question (also see chapter 3 on the expressions of subject focus).

- (16) (“Who gave the child the oranges?”)
Bí-búru bíí-wí mwáana ma-láara.
 8-parent 8SM.PST-give.PST 1.child 6-orange
 ‘The PARENTS gave the child the oranges.’

The structural positions of topical and focal subjects should be distinguished. We see that topical and focal subjects differ in their relative position to the negative prefix, as illustrated in (17). In (17a) the negative marker which scopes over the VP is attached to the verb and follows the topical subject; whereas in (17b) the negative marker precedes the preverbal focal subject and the negation is confined to the subject itself. Assuming that the negative marker is placed in a fixed position in the hierarchy, I take the word order contrast in (17) as evidence that focal and topical subjects occupy distinct structural positions. I will argue in section 5.4 that the two sentences in (17) are in fact very different constructions.

- (17) a. (*Did father kill the leopard?*)
Taará ka-ká-dwí ngo ni.
 1.father NEG-ISM.PST-kill.PST 1.leopard NEG
 ‘Father did not kill the leopard.’
- b. Ngo ka **taará** á-dwí ni.
 1.leopard NEG 1.father ISM.PST-kill.PST NEG
 ‘FATHER did not kill the leopard.’ (*but it was killed by someone or some animal else*)

So far we have seen that the preverbal subject in Kukuya can be topical or focal, or non-topical/non-focal. Since I have shown that the subject marker in Kukuya is an agreement morpheme and the preverbal subject DP can be non-dislocated, there must be some A-position(s) hosting the preverbal subject. I also showed that focal subjects must be distinguished and have a separate position. The next questions are in which exact structural position the subject is placed, is it a high or low A-position, and are the topical and

non-topical subjects structurally distinguished?

Based on the properties of the preverbal subject introduced above, I don't see solid arguments for the topical and non-topical/non-focal subjects to be located in different A-positions, while a focal subject should be placed in a different position. Considering the fact that in Kukuya the subject, whether topical or non-topical, can never occur postverbally in the matrix clause and there are no subject inversion constructions as in many other Bantu languages (Marten and van der Wal 2014), as stipulated in (18), an EPP feature should be activated in the grammar to always raise the subject. Assuming that it is the T head that is equipped with the EPP feature, the specTP position in Kukuya should be available to host the grammatical subject and allows it to be topical or non-topical.

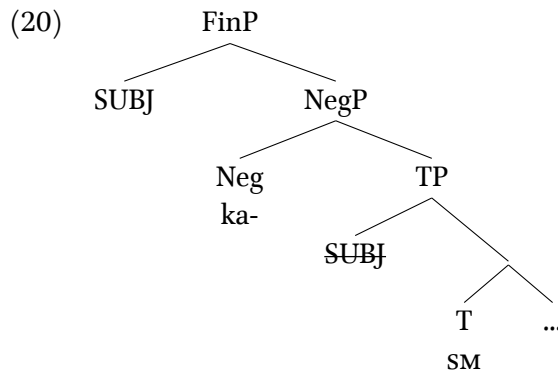
- (18) *Postverbal subject in a Kukuya matrix clause.

Once we look at the negative counterpart of an SVO sentence, we see that the negative prefix *ka-* is always attached to the verb preceding the SM, as shown in (19).

- (19) a. Taará á-mún-i me.
 1.father ISM.PST-see-PST I SG.PRO
 'Father saw me.'
- b. Taará ka-ká-mún-i me ni.
 1.father NEG-ISM.PST-see-PST I SG.PRO NEG
 'Father did not see me.'

Assuming that SM is a spell-out of ϕ -features on T, the NegP that hosts the negative prefix should then be situated above TP. The specTP would intervene between the negative and subject agreement prefixes, and therefore specTP cannot be the subject position but the grammatical subject should occupy a high A-position which I suppose to be specFinP. This correctly accounts for the linear order of the prefixes and Subject-Verb order, as illustrated in (20). Since there is no interpretational difference between the preverbal subjects in affirmative and negative SVO sentences (the subject

in a negative SVO can also be topical or non-topical), I assume that there is a unique structural position for the preverbal subject in SVO, which is specFinP. Rizzi (1997) claims that FinP participates both in processes related to the CP domain for discourse manipulations as well as in the inflectional domain, which renders the mixed A'/A-properties of specFinP. Here I suppose that the preverbal subject's being placed in specFinP is also consistent with the fact that it is available for both topical andthetic interpretations.



In (20) I show that the subject DP moves from specTP to specFinP, which at first glance should be ruled out for affirmative sentences by anti-locality since this Spec-to-Spec movement crosses only FinP without the NegP being sandwiched in between, thus is too local (Grohmann 2003, 2011; Bošković 2016; Erlewine 2016). This violation of anti-locality can be circumvented if we consider that there is a Polarity Phrase (Holmberg 2015) between FinP and TP which is phonologically null in an affirmative sentence, therefore the movement from specTP to specFinP is not computed as too close. I also assume that the φ -features are located on T where they get spelled out after agreement, and T is also equipped with the EPP feature that raises the subject from its thematic position to specTP first. The motivation of the movement from specTP to specFinP may be discourse-related or just a formal requirement.

In this section, I have provided the syntactic representation of SVO order in Kukuuya. By investigating what the FV intrinsically encodes, I discussed the structural position of the verb in SVO, arguing that the verb can stay in

its base position or head-move to T or Mood. I then showed that the SM in Kukuya functions as an agreement marker and the subject in SVO is clause-internal. Based on the relative position of the negative prefix, I proposed that the preverbal subject in SVO should be in a high A-position, which is specFinP. Taking the SVO syntax as a starting point, in the next section I look into the structural representation of the IBV focus construction.

5.2 Where is the IBV focus position?

In chapter 4, I provided evidence for the IBV focus construction to be in a transitional stage between biclausal and monoclausal. The monoclausal properties it displays are: 1) vowel coalescence and H tone plateauing shows that the IBV focused element is always prosodically phrased together with the following verb; 2) the fact that IBV focus construction can also be used to express VP focus indicates that focus can extend from the IBV element up to the whole VP, which is not a character of a biclausal cleft. The biclausal properties it retains include the H tone on the SM, the occurrence of class 1 SM *ka-* and the verb-final H tone in non-subject focus, as well as the ban on the use of the negative prefix *ka-* on the verb. In this section, I discuss the structural position of the IBV focused element. I first demonstrate that the IBV focused element is not base-generated but has undergone A'-movement to a structurally defined position. Then I show that focused subject and focused non-subject are placed in the same structural position. Next I illustrate which structural position in the derivation maps onto the IBV focus position. Since I suppose that to understand the derivation of the IBV focus construction, the subject marking alternation should be explained first, I leave the full structural representation of the IBV focus construction until the end of the chapter, but only discuss the height of the Focus Projection (FocP) in this section.

5.2.1 IBV focus as A-bar syntax

In this section I use several examples to show that the focused element is not base-generated in a high FocP in the C-domain but has raised from a vP-internal position through A'-movement.

A'-movement typically shows clause-boundedness and reconstruction effects. Clause-boundedness indicates that an A'-moved element can cross multiple clause boundaries, and reconstruction effects mean that the moved element is interpretable in its original position, thus being subject to various binding and scope relations. The sentences in (21) show the fact

that a focused element can be interpreted in the lower embedded clause while showing up in the IBV position of the matrix clause.

- (21) a. We **kí-má** á-dzií wurí taará ká-fíuum-a?
 2SG.PRO 7-what 2SG.SM-like 2SG.COMP 1.father 1SM.SBJV-buy-FV
 ‘What do you want father to buy?’
- b. Me **wúna mwáana** n-dzií ká-yók-o ma-dzá.
 1SG.PRO only 1.child 1SG.SM-like 1SM.SBJV-bath-FV 6-water
 ‘I only want the CHILD to take a bath.’

Since the focused elements are obviously interpreted in the lower clause and assuming that base-generation cannot reconstruct, there must be some syntactic movement taking place. The sentences in (21) involve the matrix verb “want/like” which is known as a “raising to object” verb, and we see the focused phrases that occur in the matrix IBV position are interpreted in the lower clause.

There are also constructions in (22) in which the matrix verb is not a typical raising verb, and we also observe that the *wh*-words occur in the IBV position of the matrix verb. Since *wh*-elements need to take scope, they have to be interpreted in the matrix clause, while their non-*wh*-part is interpreted in the embedded clause.

- (22) a. We **ku-ní** ká-tsuom-ó wurí líi-ye
 2SG.PRO 17-which 2SG.IMPF-think-FV 2SG.COMP 1PL.SM.FUT-go
 mbhíi?
 9.hunting
 ‘Where do you think that we will go hunting?’
- b. %Ndé ma-sáani **ku-ní** ká-lak-í ndíri taará
 1.PRO 6-plate 17-which 1.SM.PST-say-PST 1.COMP 1.father
 á/(*ká)-tí bví?
 1SM.PST-launch.PST 9.falling
Int: ‘Where did he say that father threw the plates?’

- c. Ma-sáani **ku-ní** ndé ká-lak-í ndíri taará
 6-plate 17-which 1.PRO 1.SM.PST-say-PST 1.COMP 1.father
 á/(***ká**)-tí bví?
 1SM.PST-launch.PST 9.falling
 ‘Where did he say that father threw the plates?’ [cleft]

Example (22a) is taken from spontaneous speech, while the elicited sentence (22b) meets some intra- and inter-speaker inconsistencies on its felicity. As for the cleft construction (22c) in which the *wh*-element is placed before the subject instead of being in the matrix IBV position, all the consultants judged it as grammatical but quite marginal. In the examples in both (21) and (22), we see that the focused element can cross clause boundaries, which is characteristic of A'-movement.

If these examples above show that there is discourse-related movement, in this case focus movement, under the standard successive cyclic movement assumptions the focused element should move to specCP of the lower clause first and then continue to move to the matrix clause. Assuming that specCP is an A'-position, and that mixed A-A'-A chains of movement as typical improper movement should be ruled out (Chomsky 1973, 1981; May 1979; Lasnik and Saito 1988; Williams 2003; Safir 2019), it suggests that the IBV position must be an A'-position².

Examples of quantifier float and split-off numerals are also suggestive of movement, or at least can be accounted for by movement. In (23) we see a postverbal quantifier “all” agreeing in noun class with the IBV focused object “beans”, which shows that the NP *báana* should have been in the

²Interestingly, the speakers insisted that for both (22b) and (22c), the class 1 subject marking in the embedded clause can only take the canonical shape *a-* and cannot be *ka-*, while the subject marking on the matrix verb must be *ka-*. What may be problematic here is the unavailability of the class 1 SM *ka-* in (21b-c). As I will argue in the next section, the use of class 1 SM *ka* always signals the extraction of a non-subject constituent, the unavailability of *ka-* here may imply that the focused element has never moved to the edge of the embedded CP but may have moved directly from its base verbal complement position to the matrix IBV position without any intermediate landing site, which can violate the Subjacency Condition, thus obscuring the evidence for the IBV to be an A'-position. I will come back to discuss this later in section 5.4.

postverbal domain at some point, where it agrees with the quantifier; then the NP undergoes focus movement to the IBV position, and the quantifier is stranded in its base position. In (24) which is a corrective response to a yes-no question, the focused NP is in IBV while the modifying numeral is postverbal, which again shows that the IBV element may have moved from the postverbal domain and the numeral modifier is stranded. It should be noted that quantifier stranding by itself does not show A-bar movement per se, as quantifier stranding can also happen with A-movement.

- (23) Ngaaka báana má-désu káá-wí mhoì.
 1.grandmother 2.children 6-bean 1SM.PST-give.PST 6.all
 ‘The grandmother gave the children all the beans.’

- (24) (“Did mother buy three oranges?”)
 Ndé má-ko káá-fúúm-í má-tíri.
 1.PRO 6-banana 1SM.PST-buy-PST 6-three
 ‘She bought three BANANAS.’

There are also potential counterexamples against the A'-movement account. In the canonical word order (25a), the pronoun *ndé* in the possessive phrase can be bound by the patient object when both objects are postverbal; while in (25b) the possessive phrase in IBV cannot co-index with the postverbal object “dog” but can only refer to another kind of animal such as a cat, and the IBV possessive phrase cannot contain a reflexive pronoun that is co-referential with the postverbal object. These facts indicate that the IBV element cannot be interpreted in the postverbal domain, thus there is no reconstruction effect that A'-movement is characteristic of. However, in the elicitation I found these tests on reconstruction to be quite contrived and sometimes it was difficult for the speakers to give very reliable judgements, so I leave them for further work.

- (25) a. Me á-m-wí ná mvá [bvi-kídzá bvíi
 1SG.PRO PST-1SG.SM-give.PST every 1.dog 8-food 8.CONN
 ndé].
 1.PRO
 ‘I gave each dog_i its_{i/j} food.’

- b. Me bvi-kídzá bvíí ndé/?ndé-wunkúlu
 1SG.PRO 8-food 8.CONN 1.PRO/1.PRO-RFL
 á-m-wí ná mvá.
 PST-1SG.SM-give.PST every 1.dog
 ‘I gave its_{?i/j} food to every dog_i.’

Focus movement as analogous to *wh*-movement is sensitive to syntactic islands such as complex NP and relative constructions (Ross 1967) under the general principle of subjacency (Chomsky 1986). If there is indeed some A'-movement taking place, we would expect the extraction of a focused constituent from a syntactic island to be impossible. In (26) we see the presence of a complementiser followed by a free relative clause, if the *wh*-phrases in the matrix IBV position have their origin in the lower clause, it means that unexpectedly there is *wh*-movement from an embedded question which is a syntactic island. In other words, the lack of sensitivity to a syntactic island in (26) runs counter to the idea that focused element has A'-moved to the IBV position. The tricky point here is whether we should treat the elements such as *ndíri* and *bóri*, which very possibly originate from the verb “say”, as fully grammaticalised complementisers, since their distribution is quite optional and restricted (see more in chapter 2 section 2.5.1). If it does not function as a true complementiser, then it needs not to block the movement of *wh*-words. There is also recent research showing that some classic “island” configurations are fully transparent for A'-extraction in many African languages, which indicates that syntactic islands may not work in African languages as assumed on the basis of European languages (see Schurr et al. 2023; Kandybowicz 2023, *to appear*). I leave this for future research.

- (26) a. Ndé ná ká-mún-í ndíri [wũ-túr-i
 1.PRO 1.who 1SM.PST-see-PST 1.COMP 1REL.PST-steal-PST
 mi-pará]?
 4-money
 ‘Who did s/he see that stole the money?’

- b. We **ku-ní** **á-yúk-í** **bóri**
 2SG.PRO I7-which 2SG.PST-hear-PST 2.COMP
 [ku-ká-kwí ndé]?
 I7REL-ISM.PST-die.PST 1.PRO
 ‘Where (the place of death) did you hear that s/he died?’

Based on the discussion above, though some counterarguments exist, I propose that the IBV focused element is not base-generated but has undergone A'-movement to a structurally defined position. Next I will identify which position it is.

5.2.2 Structural position of the IBV element

In this subsection, I explore which structural position the IBV focused element occupies. I first show that there should be one unique structural focus position for both subject and non-subject. Then I present several possible approaches to structurally representing the IBV focus position and propose that the IBV element is in a high FocP.

5.2.2.1 Unifying subject and non-subject focus

In chapter 3, I have shown that the IBV position is available for focus on different syntactic roles such as subject, object and adjunct. Considering the asymmetry between subject and non-subject on their canonical position relative to the verb, we may wonder whether the subject and non-subject are placed in the same structural position when they are IBV focused, or there are distinct projections hosting them. Here I provide some evidence in favour of a unique focus position analysis.

The most apparent evidence is the strict verb adjacency requirement on the focused element (except for the focused element in a cleft). In (27a) and (27b) we see that if a preverbal element is focal, whether it is a subject or non-subject, it cannot be followed by any other preverbal element. If we

assume the same structural height of the verb in these sentences, it shows that the focal subject and non-subject compete for the same hierarchical position.

- (27) a. Ndé ma-lí ku dzáandu ká-fúum-i.
 1.PRO 6-wine 17.LOC 5.market 1SM.PST-buy-PST
 'He bought some wine at the market.'
 ✗'What did father buy at the market?'
 ✓'Where did father buy the wine?'
- b. Mvá mpúkú ká-siib-i.
 1.dog 1.rat 1SM.PST-catch-PST
 'The dog caught the rat.'
 ✗'Who caught the rat?'
 ✓'What did the dog catch?'

A focal subject and a focal non-subject cannot co-occur in the preverbal domain, but only the IBV element can be interpreted as focused. In (28) we see that there cannot be multiple foci in the preverbal domain. In (28a) the subject and object *wh*-phrases cannot co-occur preverbally, and in (28b) the subject question is not compatible with predicate doubling which is considered to be a special use of the IBV focus strategy (see more in chapter 3 on predicate doubling). The ban on multiple foci no longer exists when one focal element is in IBV and the other is postverbal, as seen from the contrast in (28c) and (28d). These examples again show that the focal subject and non-subject compete for a unique structural position.

- (28) a. *Ná kí-má ká-fúum-i?
 1.who 7-what 1SM.PST-buy-PST
Int: 'Who bought what?'
- b. *Ná kí-lila kâ-lil-a?
 1.who INF-cry 1SM.IMPF-cry-FV
Int: 'Who is crying?'
- c. *Ná wúna bi-ko ká-swaak-í?
 1.who only 8-clothes 1SM.PST-wash-PST
Int: 'Who washed only the clothes?'

- d. Ná á-swaak-í wúna bi-ko?
 1.who 1SM.PST-wash-PST only 8-clothes
 ‘Who washed only the clothes?’

When negating the IBV focused element, the negative marker *ka-* always precedes it, whether it is a subject or a non-subject phrase, as shown in (29). Assuming that the position of the negative marker is fixed in the hierarchy, the fact that it precedes both the focal subject and non-subject can show that they are placed in the same structural position.

- (29) a. Taará ka ngo ká-dwí ni.
 1.father NEG 1.leopard 1SM.PST-kill.PST NEG
 ‘Father did not kill the PANTHER.’ [SOV]
- b. Ki-wáli ka ndzulí á-dzí ni.
 7-duck NEG 1.cat 1SM.PST-eat.PST NEG
 ‘The duck was not eaten by the CAT.’ [OSV]

The tense auxiliary *âli* which is used to express remote past tense and can in some circumstances bear inflection, is also seen to precede the IBV focused element. In (30a) the auxiliary precedes the focused subject, while in (30b-c) it occurs after the topical elements and precedes the focused adjunct. Here the relative position to the auxiliary can also show that the focal subject and non-subject occupy the same structural position below Aux.

- (30) a. Nzó âli [kí-fúlá]_{FOC} kí-bólik-i.
 9.house RPST 7-wind 7SM.PST-destruct-PST
 ‘The WIND destructed the house.’
- b. Ndé Bibulu âli [ku-ní]_{FOC} ká-táal-i?
 1.PRO Bible RPST 17-which 1SM.PST-look-PST
 ‘Where did s/he read the Bible?’
- c. Me ma-dzá âli [ŋa kalá mbali]_{FOC}
 1SG.PRO 6-water RPST 16.LOC inside 9.yard
 á-n-yók-i.
 PST-1SG.SM-bath-PST
 ‘I took the bath IN THE YARD.’

There is one counterexample as in (31) in which the focused object precedes both the auxiliary and the verb. Though the speakers judged this positioning of the auxiliary to be quite marginal but it is said to be grammatical. This example also challenges the notion of “IBV focus” on whether the focused element is actually adjacent to verb or to finiteness. At this point I leave aside this counterexample for further research.

- (31) Mu-káli aa mu-kokó ná ká-li ká-béer-i?
1-wife 1.CONN 1-king 1.who ISM-RPST ISM.PST-beat-PST
‘Who had the queen (wife of the king) beaten?’

Focused subjects and focused non-subjects do not show interpretational differences in Kukuya. In some studies, notably in Belletti (2008, 2012), a distinction on the interpretation of the focused subject and non-subject in the French and Italian *cleft* constructions has been noted. She noticed that in a subject cleft, the focused subject can be interpreted as the focus of new information, while a clefted object only allows for a contrastive/corrective focus interpretation, based on which she proposed different derivation for subject and non-subject clefts. If this distinction also holds in Kukuya, we may expect the same interpretational differences in the IBV subject and non-subject focus strategies since they both originate from clefts. However, as shown in chapter 3, I don’t see a similar distinction on their interpretation, both IBV subject and non-subject focus can be the answer to a *wh*-question expressing new information focus, and they can be used to express corrective/contrastive and identificational focus, which again suggests that they should occupy the same structural position.

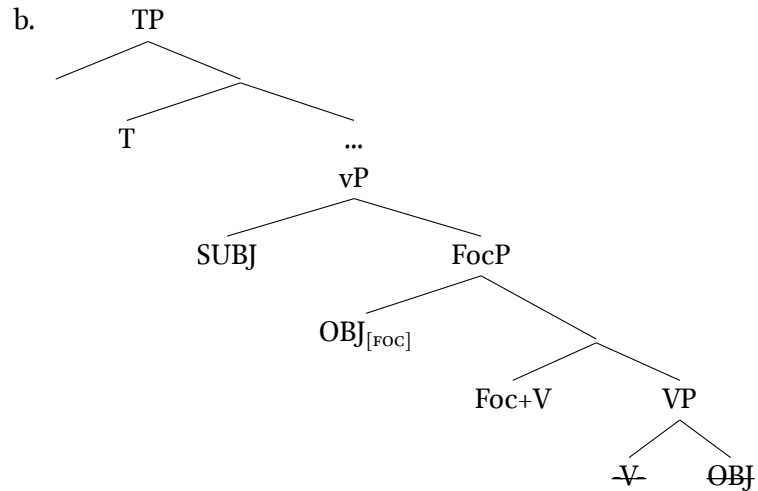
In summary, I have suggested that there is one unique structural position for the IBV focused subject and non-subject, which also means that the focused subject is placed *ex-situ* in a structurally defined focus position. Next I consider whether this Focus position is low or high in the hierarchy.

5.2.2.2 Low or high FocP?

First I exclude the possibility for the IBV element to be in specTP. Since the IBV focused element is always linearly adjacent to the verb, we may wonder if it is just located in specTP, as long as we postulate a [FOC] feature on T that can attract the focused element to its specifier. However, we cannot explain why in an SOV sentence T cannot agree with the intervening focused object if it is in specTP, since there seems to be nothing preventing this agreement, so apparently the focused element cannot be in specTP.

In section 5.1 I proposed that the subject in SVO order is in a high A-position which is specFinP, and it has moved through specTP. Supposing for the moment that the grammatical subjects in SOV and SVO occupy the same structural position, we may wonder if the focused element is located in a position lower than the TP. Assuming that the IBV element is in an A'-position, which can be the specifier of a FocP, here I discuss Belletti (2004) and Aboh's (2007) proposal that the edge of vP can be divided into a split A'-domain, along the lines of the split CP-domain advocated by Rizzi (1997). We first consider Belletti's (2004) proposal on a vP-internal FocP as schematised in (32a), and a corresponding derivation in the context of object focus is illustrated in (32b). Belletti (2004) proposes that there is a FocP within the vP periphery whose specifier hosts the focused subject, and she further suggests that this vP-internal low FocP is specialised in the expression of new information focus.

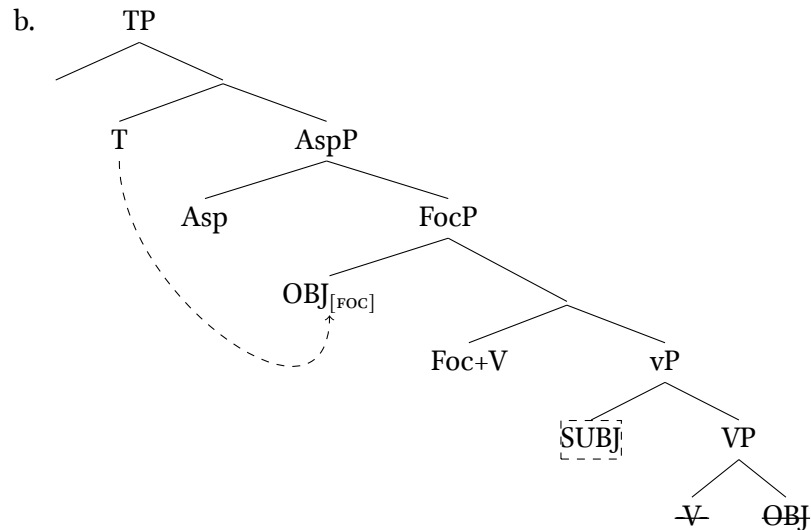
(32) a. vP<TopicP<FocusP<VP (Belletti 2004)



An obvious problem that this approach brings is that since the focused subject is considered to be located in a FocP, it should be attracted by some [Foc] feature from its base position in specvP, which stipulates that the Foc head must c-command the vP, otherwise the subject can never be visible to the [Foc] feature thus cannot be attracted to specFocP, which renders subject focus in the IBV position impossible. Since the FocP should be an A'-position, it is not possible that the focused subject is base-generated in the low specFocP within the vP and then raised by T. So Belletti's (2004) approach cannot be adopted for Kukuuya.

The proposed hierarchy in Aboh (2007) is shown in (33a), which is slightly different from Belletti (2004) in that the FocP projects at the edge and dominates vP. A corresponding derivation in the context of IBV object focus is illustrated in (33b).

(33) a. AspP<FocusP<(TopicP)<vP<VP (Aboh 2007)



Under this approach the Foc head can attract the subject from specvP in the context of subject focus, but it also runs into some severe problems. As illustrated in (33b), at the point in the derivation when T is merged, it probes downwards into its c-command domain to search for a matching goal and raise it to its specifier, since T has an EPP feature in Kukuya. However, at this point the FocP is merged above vP and would attract the focused object to move to its specifier, thus the closest goal that T can find is the focused object in specFocP but not the subject DP, which comes up with a locality problem. In this case the subject in specvP cannot agree with T because of the intervening focused object, and the agreement pattern is wrongly predicted. The locality problem can be overcome by assuming relativised probing, for example a nominative Case feature on T can only find the subject as a matching goal (although it remains to be seen whether Case plays a role in Kukuya). In addition, in section 5.1 I proposed that the verb head-moves to T or some head between v and T to incorporate the *rv*, by the end of the derivation in (33b) the focused object terminates in specFocP and the verb should have head-moved above it, after linearisation the word order should be SVO rather than SOV. Even if the verb ultimately moves to a head lower than the FocP, we should assume the SM on T to be lowered onto the verb, jumping over the focused element in specFocP. If

this were to happen, we would expect affix hopping to be also possible in non-subject relatives to allow a preverbal subject, which turns out to be wrong (see section 5.3). Due to these deficiencies, Aboh's (2007) approach cannot work either for the IBV focus construction in Kukuya.

The low FocP approaches also predict that there are some A-positions such as specTP and specFinP above the IBV focused element, which means that the subject in an SOV sentence can be indefinite non-specific. Consider example (34) in which the theme object is focused, although it is difficult for the speakers to judge whether the initial subject in SOV is definite or not, they accept the word *mbuurú* "person" to occur in the sentence-initial position, but rejected it to be modified by *nguumó* "one" which renders this initial subject DP indefinite. So we see that in the SOV order, the initial subject cannot be non-topical. Also recall that there is often a secondary topic between the initial subject and IBV focus position as in (35), it is plausible to deduce that the initial subject should also be topical.

- (34) **Mbuurú** (*nguumó) baa-ntsúú má-désu ká-búnum-i.
1.person 1.one 2-chicken 6-bean 1SM.PST-feed-PST
'The person fed the chicken the BEANS.'

- (35) We **mbuurú** ku-ní á-mún-i?
2SG.PRO 1.person I7-which 2SG.SM-see-PST
'Where did you see the person/*anyone?'

In an OSV construction where an object is fronted and the subject is IBV focused, as shown in (36) and (37), we see that when the words *mbuurú* "person/someone" and *kilóko* "thing/something" occur sentence-initially, they can only have a definite reading but cannot be indefinite, which can also show that the initial topical element in an IBV focus construction cannot be in an A-position but must correspond to some higher projection in the C-domain. If we consider the SOV and OSV orders as parallel IBV focus constructions, this also suggests that there is no A-position preceding the IBV element and the IBV element should be in a higher A'-position.

- (36) *(In a catching game, father was asked to catch a group of people, but finally he caught nobody.)*

#**Mbuuru** taará ka-ká-siib-i ni.
 1.person 1.father NEG-ISM.PST-catch-PST NEG

Int: 'Nobody was caught by father.'
 'The person was not caught by father.'

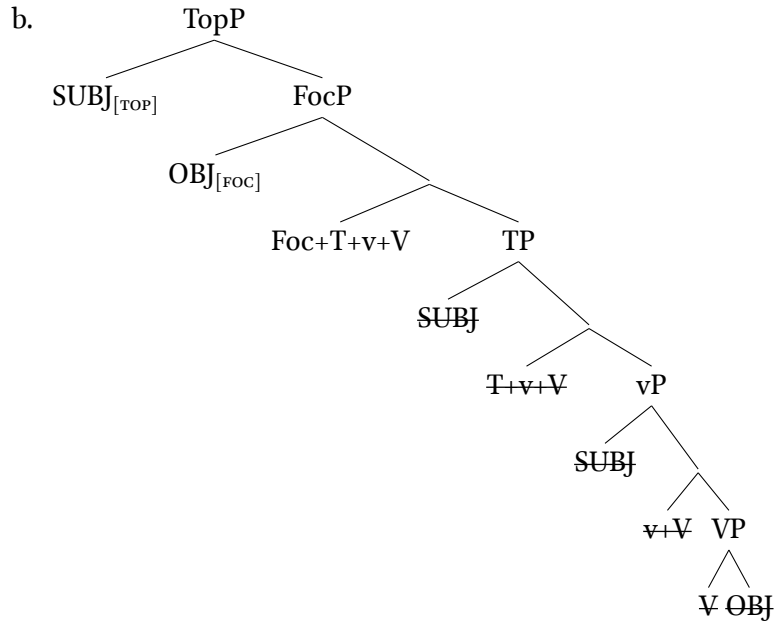
- (37) *(You had prepared many dishes for mother and you went out, when you came back, the food remained untouched.)*

#**Ki-lóko** ngúku ka-ká-dzí ni.
 7-thing 1.mother NEG-ISM.PST-eat.PST NEG

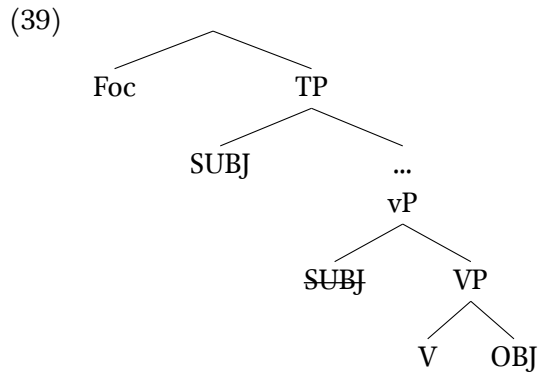
Int: 'Nothing was eaten by mother.'
 'The thing was not eaten by mother.'

As an alternative to the discussed analysis, I propose that the IBV focused element in Kukuya is in a high FocP which is above TP as in (38a), following the cartographic approach in which the CP domain is articulated into a series of functional heads and projections associated with discourse/interface effects (Rizzi 1997). A corresponding derivation is given in (38b), in which the intermediate landing site of the focused object on the vP phase edge is omitted, and the verb is assumed to terminate at Foc.

(38) a. TopP < FocP < TP < AspP < vP

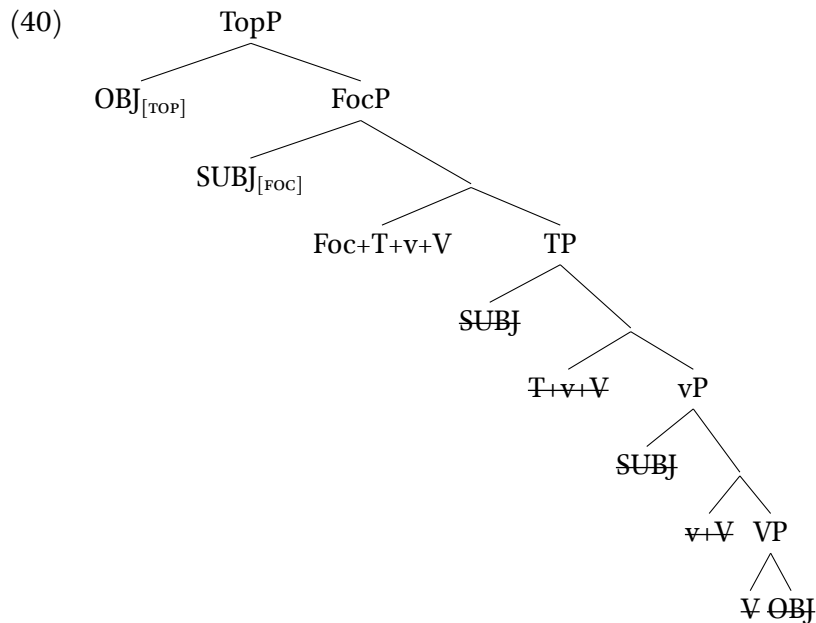


This high FocP approach is consistent with the topic interpretation of the initial element, and the high FocP is available for both subject and non-subject to get focused in its specifier, as both of them can move there. Under this approach, the SOV order in object focus is correctly predicted, since even the verb ultimately moves to T and then to Foc, the verb complex is still structurally lower than the focused object in the high FocP, as shown in (38b). The locality problem when subject agreement takes place no longer exists, since T is merged earlier than Foc, the subject in specvP is always the closest goal of T for the agreement relation to be established, and then the Foc can attract any element within the TP that bears the [FOC] feature, as illustrated in (39). The verb adjacency to the focused element is also accounted for in the high FocP proposal, assuming that the FocP is in the extended domain of the TP and always immediately dominates it. This high FocP approach is also superior to the low FocP in that the cleft origin of the IBV focus construction predicts that the reanalysed focused phrase should be situated in a higher domain than the vP. Based on these arguments, I propose that the high FocP proposal thus is the most suitable analysis for the IBV focus strategy.



One potential problem it may bring is the position of the negative prefix *ka-*, since I have shown that in SVO it is between Fin and T and linearly it precedes the IBV focused element, which is not compatible with (39). Another potential problem is whether the movement from specTP/specFinP to specFocP violates anti-locality. I will come back to discuss these problems in section 5.4.

The derivations shown above all illustrate object focus. Since I have proposed that a focal subject occupies the same structural position as a focal object, the derivation of IBV subject focus construction is provided in (39), in which the OSV order is also correctly predicted. In (39) I show that the Foc head attracts the focal subject in specTP and moves it to its specifier; the object that bears a [TOP] feature is raised by the Top head directly from its base position. In (38b) and (39) I also assume that the initial topical element has moved from the lower clause and is not base-generated, which also needs further explanation. These questions are discussed in section 5.4.



In the above discussion, I compared the pros and cons of the low and high FocP approaches and concluded that the high FocP analysis better accounts for the morphosyntactic properties of the IBV focus construction, so the IBV focused element occupies a high FocP above TP. However, the derivation in (38b) is not yet the full picture of the IBV focus construction, and there are still some questions remaining to be answered. Under the structure in (38b) and (39), the class 1 *a-* versus *ka-* SM alternation in SVO and SOV is not expected, since the subject agreement mechanism would be the same in the situations of subject and non-subject focus, as the subject can always agree with T and be raised to specTP. Recall that the class 1 SM alternation *a-/ka-* is attested both in the IBV focus and in non-subject relatives which I consider to be closely connected. By first looking into the agreement patterns in relative clauses we may get some hints for an explanation of the SM alternation in the IBV focus strategy. In the next section, I investigate the subject marking asymmetry in subject and non-subject relative clauses in Kukuuya, in order to provide an account of the subject marking alternation in the IBV focus strategy.

5.3 Subject marking asymmetry and IBV focus

In this section I explore another intriguing morphosyntactic property in the IBV focus construction, namely the class 1 *a-* versus *ka-* SM alternation in subject/non-subject focus. As I have demonstrated in chapter 4 that IBV focus has a cleft origin and its verb form has retained some relative properties, I first present and analyse the subject agreement asymmetry between preverbal and postverbal subjects in Kukuya relative constructions, and then I extend the analysis onto the IBV focus construction in section 5.4. In section 5.3.1, I give a detailed description on the agreement relations in Kukuya subject and non-subject relatives, showing that while full subject agreement is always available in subject relatives, agreement with the postverbal subject in non-subject relatives is determined by the presence of the [Person] feature on the subject. In section 5.3.2, I provide an account of the subject marking asymmetry, seeking for explanation from the featural inventory and internal structure of the subject DPs, as well as specific SM spell-out rules.

5.3.1 Agreement in subject and non-subject relatives

Crosslinguistically, languages differ in their agreement inventories, and how much agreement is expressed under different structural configurations can vary. As for Bantu languages, much previous research has shown that subject agreement is typically associated with the vP-external subject, and the subject's being raised to specTP according to many researchers is a necessary condition for its agreement with T (Demuth and Harford 1999; Buell 2005; Carstens 2005; Zerbian 2006; Baker 2008; van der Wal 2009; Zeller 2013).

In chapter 2, I have introduced how different types of relative constructions are expressed in Kukuya, in this subsection I discuss the subject agreement asymmetry in subject and non-subject relatives in Kukuya. I look in more detail at the structural position of the subject and the agreement relations within relative clauses. I show that subject marking can differ with regard

to whether the subject is preverbal or postverbal, as well as whether the subject is pronominal or lexical. We start from the agreement pattern in subject relatives.

5.3.1.1 Agreement in subject relatives

In a subject relative in Kukuya, the relativised subject is indexed twice on the verb: by the relative marker and the subject marker, as shown in (41) and (42). Here I treat the relative marker as a relative pronoun rather than the overt complementiser in the clausal spine (Cheng, *forthcoming*). I take the REL marker to represent φ on C and the SM φ on T.

- (41) [Li-meé **li-líi-súruk-i**] líi-búl-i ndzulí mu-tswê.
 5-stone 5REL-5SM.PST-fall-PST 5SM.PST-hurt-PST 1.cat 3-head
 ‘The stone that fell hurt the cat’s head.’

- (42) Mu-kái wu-kíele á-luon-i [ma-mbaá **ma-má-ye** ndé
 1-woman 1-young 1SM-follow-PST 6-light 6REL-6SM-go 1.PRO
 tíi ku ma-mee].
 until 17.LOC 6-waterfall
 ‘The young woman followed the light which directed her to the waterfall.’

The agreeing relative marker takes the same shape as the proximal demonstratives (see chapter 2 section 2.3); the subject markers appear in their canonical forms. The REL-SM sequence always has a rising tone pattern LH, which is also the same tone scheme of the NEG-SM sequence, thus it shows that the Pre-initial and the Initial slots form a rising tone pattern, which is also typologically characteristic of many Bantu languages in the Congo basin (Van de Velde 2021). When the subject marker has a CV shape that starts with a bilabial consonant, namely the class 2 *ba-*, class 4 *mi-*, class 6 *ma-* and class 8 *bvi-*, the subject marker is often phonologically merged with the preceding relative pronoun, which results in a single relative prefix with

a rising tone on a long vowel, as shown in (43). This phonological merge corresponds to the intermediate step in the agreement cycle in Bantu relatives proposed in Van de Velde (2021), as schematised in (44).

- (43) Báana **baá**-yi-sá mí-táami ku ntsá mvúla má-baá
 2.children 2REL.SM-IMPF-do 4-joy 17.LOC inside 3.rain 6-heat
 máá-sî bó.
 6SM.PST-do.PST 2.PRO
 ‘The children who were playing in the rain got fever.’

- (44) $HEAD_i REL_i-AGR_i-V (...)\implies HEAD_i REL_i-V (...)$

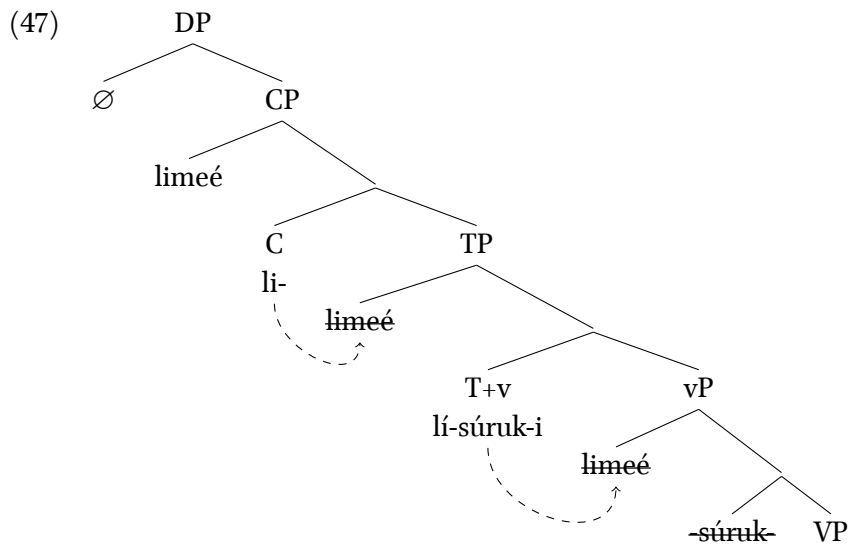
As reported in some literature on Bantu relatives (Kinyalolo 1991, Henderson 2013; Van de Velde 2018, 2021), it is often impossible to relativise the first and second person pronouns with a fully agreeing relative marker, and this is also the case in Kukuya as shown in (45). This can be understood by the fact that relative markers originate from demonstratives which usually cannot modify the first and second person pronouns. In (45) we see that the relative marker agrees with the 1st person plural subject in [NUMBER] and [GENDER], while only the subject marker is specified for [PERSON]. This shows that while T can show agreement with full φ -features, C is impoverished for [PERSON], as represented in (46).

- (45) Bhií **ba-lí**-ká-sál-á máa-ntséke
 1PL.PRO 2REL-1PL.SM-IMPF-work-FV 6-field
 lí-ká-sílik-a bú-su.
 1PL.SM-IMPF-wake-FV 14-front
 ‘We who work in the fields wake up early.’

- (46) a. φ -features in C: [GENDER], [NUMBER]
 b. φ -features in T: [PERSON],[GENDER], [NUMBER]

Since we see that in subject relatives the preverbal subject agrees with both C and T in their [u φ] features and the two heads show distinct feature specification, I assume that they bear separate sets of [u φ] features (Henderson

2011; Carstens and Diercks 2013), thus I am not following the Feature Inheritance account (Chomsky 2005, 2008; Miyagawa 2010) in which the [$u\phi$] features on T are inherited from C. There seems to be no C-T agreement in Kukuya relatives, since we see that [Person] is not suppressed for the 1st and 2nd person SMS. This assumption is necessary in order to account for the simultaneous relative and subject marking on the verb. The derivation of (41) above is illustrated in (47), in which the dashed arrows indicate agreement. The relativised subject starts out from its thematic position specvP, and T probes down and agrees with it, then the subject is raised to specTP; when C is merged with a different set of [$u\phi$] and a [REL] feature, it probes and finds the subject in specTP, the subject values the [$u\phi$] features on C and finally moves to specCP.



Some other points to be considered in the above derivation are 1) the height of the verb complex, as one may wonder if the verb ultimately head-moves to C by the end of the derivation. Here I suppose that the verb does not necessarily move to a higher head, but can just stay in v or T which depends on the TAM. We will see later that this is distinct from non-subject relatives. 2) Another question is whether the subject has also moved through specFinP or just specTP, recalling that in SVO it is in a high A-position. We have seen that in relative clause there is no NegP to host the negative prefix

ka-, so I suppose that there is no independent motivation for the external argument to pass through specFinP.³

For subject relatives, we see that the agreement pattern in Kukuya does not deviate from previous analyses on Bantu subject agreement that only a preverbal subject DP can show full agreement with T. Next we look at non-subject relatives.

5.3.1.2 Agreement in non-subject relatives

Now we are going to consider the agreement patterns in non-subject relatives which include object relatives and adjunct relatives (e.g. temporal and locative clauses that lack a relativised head NP). Similar to subject relatives, in a non-subject relative construction the head NP is indexed by a relative pronoun which is of the same set as that in subject relatives. What distinguishes non-subject relatives from subject relatives is the subject agreement pattern and the presence of the subject in a postverbal position. Interestingly, unlike many other Bantu languages in which the reduction of REL-SM succession is attested and subject marking is largely suppressed in non-subject relatives (Van de Velde 2018), in Kukuya and other Teke varieties, in non-subject relatives the SM slot on the verb is always available and filled.

There is an overt postverbal subject in non-subject relatives in Kukuya, but the subject marking on the verb differs with regard to the type of subject. When the postverbal subject is a speech participant, namely the first and second person pronouns, the subject marker takes the same shape as in SVO and subject relatives, as illustrated in (48) and (49). In (48) T agrees with the 1st person singular pronoun and is spelled out as the nasal prefix *n-*; in (49) both the [uφ] on the auxiliary and the lexical T agree with the 1st person plural pronoun and the subject marker is also spelled out in its canonical form *li-* with an additional grammatical H tone for relative

³Here I assume that in Kukuya subject relatives the subject moves from specvP to specCP *through* specTP, which runs counter to some proposals that subject *wh*-movement from specTP(specIP) to specCP is crosslinguistically banned (Bošković *to appear*).

marking.⁴

- (48) Yǎ nkú má ngámo yi-má-n-tá me.
 with 9.story 9.one 9REL-AUX.FUT-1SG.SM-tell 1SG.PRO
 ‘There is a story that I will tell.’
- (49) ma-meé ma-lii-li líi-tí bhii bvi
 6-stone 6REL-1PL.RPST-AUX 1PL.RPST-launch.PST 1PL.PRO 9.falling
 ‘the stones that we had thrown away’

Examples with 2nd person subjects are given in (50), in both sentences T agrees with the overt postverbal 2nd person subject pronouns. In (50a) we see that the 2nd person singular SM *a-* converges with the preceding class 5 relative marker *li-* and triggers vowel coalescence; and in (50a) the 2nd person plural SM takes the shape *li-*, which is the same as in the matrix clause and is identical to the 1st person plural SM.

- (50) a. li-meé leé-tí we bvi
 5-stone 5REL.2SG.SM.PST-throw.PST 2SG.PRO 9.falling
 ‘the stone that you(pl.) threw away’
- b. mfú yi-líi-ték-í bé
 10.hair 10REL-2PL.SM.RPST-sell-PST 2PL.PRO
 ‘the hair that you(pl.) sold’

When the postverbal subject is a lexical DP, or a pronoun of noun classes other than class 1/2, only a default class 7 subject marker *ki-* can be used, and full ϕ -feature agreement is never allowed, as shown in (51) and (52). In these examples the relative marking is always present and the agreement is with the head noun. In example (51a) the postverbal subject is a class 2 lexical DP and in (51b) the subject is in class 1; in both cases we see that the canonical agreement morphology cannot be used, but only the default SM *ki-* can and must appear on the verb.

⁴Here I will not be concerned with the derivation of complex tenses or hyperagreement (Carstens and Diercks 2009; Carstens 2011) but the point here is to show that there is always subject agreement for speech participants.

- (51) a. Ki-pfúo ki-kíi/*báá-fúúm-í báana baa-mvá
 7-bread 7REL-7/*2SM.PST-buy-PST 2.children 2-dog
 báá-dzí.
 2SM.PST-eat.PST
 ‘The bread that the children bought was eaten by the dogs.’
- b. mfú yi-kíi/*káa-li kíi/*káa-ték-í taará
 10.hair 10REL-7/*1SM.RPST-COP 7/*1SM.RPST-sell-PST 1.father
 ‘the hair that father had sold’

In (52), the postverbal subject is a class 4 pronoun which refers to some pigs, in this example still only the *ki-* marker can be used, and the class 4 SM *mi-* cannot occur on the verb.

- (52) Bviila bvi-kíi/*míi-dz-í njé
 8.food 8REL-7/*4SM.PST-eat-PST 4.PRO
 ‘the food that they (the pigs) ate’

Intriguingly, when the postverbal subject is the class 1 pronoun *ndé*, one of two possible SMS can appear on the verb. In examples (53) and (54), we see that the SM that indexes the postverbal pronoun *ndé* can be either realised as the form *ka-* or as *ki-*. In (55), both the auxiliary and the lexical verb can take the SM *ka-*. For all the speakers I consulted, the two subject markers are in free alternation and it seems difficult to tell any difference in use. Here the *ki-* is the default class 7 SM and I treat *ka-* as an allomorphic class 1 agreement marker. It is noteworthy that *ka-* is only used when the postverbal subject is the pronoun *ndé* but not for class 1 lexical DPs.

- (53) Ki-sáli ki-káá/kíi-lil-í ndé ka-kí-li tsítse ni.
 7-reason 7REL-1/7SM.PST-cry-PST 1.PRO NEG-7SM-COP clear NEG
 ‘The reason why s/he cried is not clear.’

- (54) Bhií líi-tsuk-í mu ku-káá/kíi-túr-í ndé
 1PL.PRO 1PL.SM.PST-talk-PST 18.LOC 17REL-1/7SM.PST-steal-PST 1.PRO
 mi-pará.
 4-money
 ‘We talked about (the fact) that s/he stole the money.’

- (55) bi-ko bi-káa-lí káa-ték-i ndé
 8-clothes 8REL-1SM.RPST-COP 1SM.RPST-sell-PST 1.PRO
 ‘the clothes which s/he had sold’

Similarly, when the postverbal subject is the class 2 pronoun *bó*, the subject agreement marker on the verb either shows up as the canonical class 2 SM *ba-* or as the default marker *ki-*, as shown in (56) and (57).

- (56) Mi-féme mi-báa/kíi-fúúm-í bó míi-bár-i.
 4-pig 4REL-2/7SM.PST-buy-PST 2.PRO 4SM.PST-escape-PST
 ‘The pigs which they bought escaped.’

- (57) bi-ko bi-báa/kíi-lí báa/kíi-ték-i bó
 8-clothes 8REL-2/7SM.RPST-COP 2/7SM.RPST-sell-PST 2.PRO
 ‘the clothes which they had sold’

To summarise, in non-subject relatives the SM slot on the verb can in some circumstances show agreement with the postverbal subject and is filled by a default SM elsewhere. The agreement pattern with different kinds of postverbal subjects is primarily summarised in Table 5.1.

Before accounting for the differential subject marking strategies, I first provide more evidence to support the idea that the *ka-* and *ki-* markers in the above examples are indeed *agreement* markers instead of other functional categories. The careful reader may notice that in the introduction on the IBV focus and cleft constructions (see chapter 3 and 4), I have shown that the same subject marker *ka-* is used when a non-subject is focused preverbally. Since focus movement and relativisation both involve

SM	1st	2nd	class 1/2 pronouns	others
SG	-N-	-Ø-	-ka-/-ki-	-ki-
PL	-li-	-li-	-ba-/-ki-	

Table 5.1: SM of postverbal subjects in Kukuya non-subject relatives (*to be revised*)

wh-movement, we may wonder whether the *ka*- actually marks extraction of the non-subject constituent. However, if it does, we cannot explain why *ka*- is only attested when a class 1 non-subject NP is extracted, but does not occur with NPs in other classes, so apparently the *ka*- marker is not directly associated with *wh*-extraction. Since it is in complementary distribution (and free variation) with *ki*-, which is the same for 1st and 2nd person SMS (see example (59) below), and 1st and 2nd person SMS have been shown to be agreement markers, I suppose *ka*- to be a class 1 subject agreement marker.

The *ka*- as a class 1 SM allomorph is actually attested in many other West-Coastal Bantu languages such as in the B80 group and the Kikongo cluster. In Bostoen and Mundeke's (2012: 149) discussion on the origin of *ka*- in SOV in Mbuun (B87), they consider that the use of *ka*- could be a result of "a historical conflation of two distinct sets of markers, a *ka*- originally linked with past tense/perfective aspect and one having a separate origin that became associated with information structure." They also conjecture that the *ka*- ultimately originates from an identificational copula which may have grammaticalised as a verbal marker of object focus, and the fronting of the focused object to IBV can also relate to how *ka*- became integrated to the verb. Bostoen and Mundeke (2012)'s account cannot be applied straightforwardly to Kukuya. If we consider the fact that the identificational copula always precedes the IBV element in Kukuya (see (30) above), it is not likely that the *ka*- also originates from a copula. I agree with Bostoen and Mundeke (2012)'s conjecture that historically the *ka*- should have a separate origin, however in this chapter I would like to provide a synchronic perspective on analysing it as an agreement marker. Since I have proposed

that the IBV focus strategy is innovated from the cleft, a basic assumption in this chapter is that the class 1 SM *ka-* in both constructions are the same agreement marker.

As for the prefix *ki-*, it is elsewhere attested as the infinitive marker and it can also mark clause dependency (possibly a separate homophonic morpheme), sometimes with a conditional interpretation, as shown in example (58) (also see chapter 2 section 2.5.4).

- (58) Me **ki-m-bvúruk-a** nzó, bu me
 ISG.PRO DEP-1SG.SM-return-FV 9.house CONJ 1SG.PRO
 m-bvi nsiina.
 1SG.SM-fall.PST 9.ground
 ‘Returning home, I sat down on the ground.’

Here I don’t consider the *ki-* in non-subject relatives to be either an infinitive or dependency marker. As shown in the above examples, we see that the relative verb can be inflected, as the FV can take the suffix *-i*, which is not typical of an infinitive verb. We also notice that in (59) the 1SG subject cannot co-occur with the *ki-* prefix, and this cannot be explained if the latter is an infinitive marker which usually does not put a constraint on what kind of subject it is associated with. The contrast in (58) above and (59) also shows that *ki-* does not mark dependency of a relative clause, otherwise we would expect that it can co-exist with the 1SG subject marker in (59). Therefore, I treat the *ki-* prefix in non-subject relatives to be a default morpheme that is indicative of agreement impoverishment, so it does not display any agreement relation with the postverbal subject. In fact, *ki-* is also the common default SM in matrix clauses, which functions as an expletive SM, as shown in (60) (also see chapter 2 section 2.4.1).

- (59) ntaba wu-***ki/m-fúúm-í** me
 1.goat 1REL-^{*}7SM/1SG.SM-buy-PST 1SG.PRO
 ‘the goat which I bought’

- (60) Ka-kí-li ntáli taará káá-sí me ni.
 NEG-7SM-COP 9.bed 1.father 1SM.PST-make.PST 1SG.PRO NEG
 ‘It was not a bed that father made for me.’

Based on the examples above, it seems that in non-subject relatives full φ -agreement is only possible with [+PARTICIPANT] subjects, namely 1st and 2nd person pronouns. This is in line with the Person Licensing Condition proposed in Béjar and Rezac (2009) that an interpretable [+PART] feature must be licensed by entering into an Agree relation with a φ -probe, which is adapted and revised in Preminger (2011, 2019) and Coon and Keine (2020) as in (61).

- (61) *Person Licensing Condition (PLC)* (Béjar and Rezac 2009; Preminger 2011; Coon and Keine 2020)
 A [PARTICIPANT] feature on a DP in the same clause as a person φ -probe must be agreed with by that φ -probe;

However, in the above examples we also find φ -agreement with class 1/2 (3rd person) pronominal subjects (even if the default morphology can be an alternative). While *ka-* is arguably a class 1 subject marker, agreement with the postverbal class 2 pronoun in (56) clearly appears as the canonical form *ba-*, and this is also what we have seen in the *ba*-passive constructions (see chapter 3 section 3.3.2). Now the question comes to why only 1st and 2nd person pronouns can show full agreement but lexical DPs cannot, and why class 1/2 pronouns are special to other class pronouns? Since the 1st and 2nd person pronouns are specified for the [+PARTICIPANT] feature that sets them apart from other subject DPs, we may want to investigate if class 1/2 pronouns are associated with any similar feature that enables them to agree with T. Here I propose that the cut-off point on the availability of subject agreement for postverbal subjects is the presence of a [+PERSON] feature, within which different values are specified. The [PARTICIPANT] pronouns have the 1st and 2nd person values, while the class 1/2 pronouns *ndé* and *bó* can be equipped with a 3rd [Person] feature that is responsible for their *in situ* agreement with T.

Considering that postverbal class 1/2 pronouns can either take the de-

fault SM morphology or an agreement marker, and [Person] features may play a specific role in subject marking, I hypothesise in (62) that the pronouns *ndé* and *bó* actually express two underlying sets of pronouns with different feature specifications. One is a pair of the conventional class 1/2 pronouns without the [PERSON] feature, which results in the default *ki*-morphology; the other is a pair of SG/PL pronouns that contain the [PERSON] feature, which are *featurally* equal to the 3rd person pronouns and can agree with T. Thus I suppose that it is the different values of [PERSON] that are responsible for the differential SM morphology of *ndé* and *bó*, as shown in Table 5.2. The φ -feature specification of different types of subject DPs is illustrated in Table 5.3. If T agrees with a subject that has [Person], full agreement results, and if it agrees with just [Number]/[Gender], the default SM is inserted.

- (62) *3rd person pronoun hypothesis:*
 There is a dichotomy within “*ndé*” and “*bó*” between 3rd person and class 1/2 pronouns with regard to the presence of [PERSON] feature.

SM	[+PERSON]			[-PERSON]	
	1st	2nd	3rd	class 1/2	other classes
singular	-N-	-Ø-	-ka-	-ki-	-ki-
plural	-li-	-li-	-ba-	-ki-	-ki-

Table 5.2: SM of postverbal subjects with different values of [PERSON] feature

Subject type	Feature specification
1st/2nd/3rd person pronouns	[PERSON], [NUMBER] and [GENDER]
lexical DPs, class pronouns	[NUMBER] and [GENDER]

Table 5.3: Feature specification of different types of subject DPs

To illustrate the structural representation of non-subject relative clauses, we should also know the position of the postverbal subject and the height

of the verb. Here I suppose that the postverbal subject is *in situ* in specvP, based on prosodic evidence and the agreement morphology. Prosodically, the assumption is that if a language usually phrases the verb and object in the VP together, the verb and the following logical subject are expected to show conjoint phonological phrasing as well, if the latter is in a vP-internal position. Recall that in a V-O structure if there is a metatonic H tone (see chapter 2 section 2.1.1.3), it always spreads from the FV of the verb onto the following prefix, which indicates that the verb and the object are phrased together. In (63) we see that the verb-final H tone also carries over onto the prefix of the postverbal subject NP, so the verb and the postverbal subject NP form one prosodic phrase and the subject should be vP-internal. The lack of full subject agreement also suggests that the subject is in a low position, since T has a separate set of ϕ -features which must be spelled out upon valuation if the subject DP has ever moved through specTP as in SVO and subject relatives.

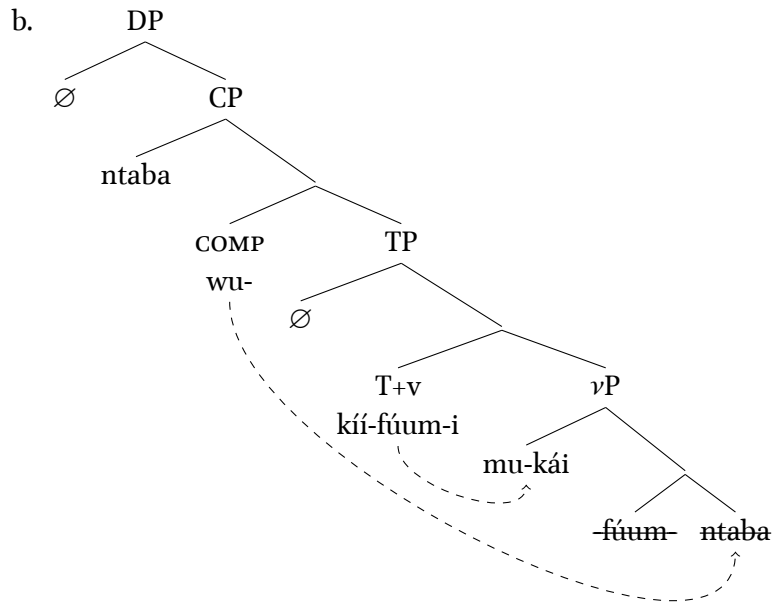
We should notice that the verb is in different height in subject and non-subject relatives. In a subject relative, the verb has the option to just stay in v, but in a non-subject relative it has to head-move to a higher head such as T to derive the OVS word order. I suppose that the linear necessity of a postverbal subject in the non-subject relative is due to the morphophonological requirement for the relative prefix to be prosodically attached to the verb complex, which means that a preverbal subject cannot intervene.

One may wonder why T cannot agree with the subject and raise it to specTP first, and then the verb just moves from T to C to satisfy the linear order requirement, since the specTP position is active in this language, and T has separate ϕ -features relevant for subject-verb agreement. I suppose that T to C movement is not possible in Kukuya, since there is no independent motivation for T to adjoin to C. According to the assumed head-movement approach to the formation of the verb complex as in (3) above, the verb head-moves out of the vP in order to incorporate the verbal suffixes, but the inflectional prefixes are just spelled out in their base positions, including the relative prefix, so the verb does not have to move to C to incorporate the relative prefix. Therefore, in a non-subject relative the subject remains *in situ* (but see the Shona non-subject relatives in section 5.3.2.1). The EPP feature of T could still be motivated in this circumstance,

but the lower copy is spelled out *in situ*, probably due to morphological requirements, which remains to be investigated in future research.

A structural representation of a non-subject relative clause is illustrated in (63). Note that specTP is empty in this derivation.

- (63) a. ntaba wu-kíí-fúúm-í mú-kái
 1.goat 1REL-7SM.PST-buy-PST 1-woman
 ‘the goat that the woman bought’



In Table 5.3 above I propose that all types of subject DPs, including the person pronouns, contain the [GENDER] and [NUMBER] feature, but only the 1st/2nd/3rd person pronouns are equipped with the [PERSON] feature. The next question would be how this distinction can help us understand the subject marking in Kukuya relatives. In the next subsection I provide an analysis on the subject marking strategies with preverbal and postverbal subjects.

5.3.2 Featural account of the subject marking asymmetry

In this subsection, I give an account of the subject marking asymmetry in Kukuya relatives. I show that the agreement asymmetry with preverbal and postverbal subjects should be attributed to the structural height of the subject, the internal structure and feature inventory of the subject DP, and some specific SM spell-out rules. I first briefly introduce similar phenomena on the subject marking asymmetry in some other Bantu languages. In the rest of the subsection, I mainly answer the following questions:

- a What causes the impoverished agreement with a postverbal subject?
- b How do person pronouns differ from other pronouns and lexical DPs?
- c What are the lexical insertion rules of SMS in Kukuya?

5.3.2.1 Subject marking asymmetry in Bantu

Looking across languages, one finds finite verbal agreement with both higher and lower arguments. Agreement with structurally lower arguments is often referred to as long-distance agreement (LDA) (Boeckx 2009; Polinsky and Potsdam 2001; Bhatt 2005), even when the agreed-with argument is linearly and structurally quite local. It often reveals an asymmetry that favours agreement with higher arguments as the default, while downward agreement is often defective in the sense that it tracks only a subset of the ϕ -features. Since for LDA there is no movement triggered by the EPP feature which links movement and feature checking, it may lead to an expectation that LDA should be less marked than agreement with structurally higher arguments (Bjorkman and Zeijlstra 2019). For example in Standard Arabic, in SV word order (64a) the verb agrees with the subject in gender and number, while in VS word order (64b) the verb is invariantly singular, agreeing with the subject only in gender.

- (64) a. ?al-bint-aani qadim-ataa.
the-girl-NOM.DL came-3.F.DL
'The two girls came.' [SV Standard Arabic]
- b. qadim-at al-bint-aani.
came-3.F.SG the-girl-NOM.DL
'The two girls came.' [VS Standard Arabic]
(Harbert and Bahloul 2002: 45)

The agreement asymmetry between preverbal and postverbal subject has also drawn attention in previous studies in Bantu syntax, and plenty of microvariation on the agreement patterns has been reported. In many Bantu languages, agreement can only target preverbal DPs, independent of the argument's grammatical role, while in some languages downward agreement is also attested. Different analyses have been put forward to account for the agreement asymmetry in Bantu (Collins 2004; Carstens 2005; Zeller 2006, 2008; Marten 2007, 2011; van der Wal 2008, 2012; Halpert 2012; Carstens and Mletshe 2015; Ngoboko 2016; and many others). Some subject inversion constructions in Bantu are illustrated in (65). We see that in Matengo the verb agrees with the postverbal subject in noun class, while in Northern Sotho only a default class 17 SM can be used when the subject is inverted.

- (65) a. Gu-hábwiki nko:ngo.
3SM-fall.PERF 3.tree
'A tree has fallen down.' [Matengo N13] (Yoneda 2011: 756)
- b. Go-binne ba-sadi fela.
17SM-dance.PST 2-woman only
'Only women danced.' [Northern Sotho S32] (Zerbian 2006: 70)

In Bantu non-subject relatives, when the subject occurs postverbally, some languages exhibit full agreement with the subject such as in Shona (66a), while in some languages a postverbal lexical subject DP cannot be indexed even by a default SM on the verb, for example in Mbuun (66b).

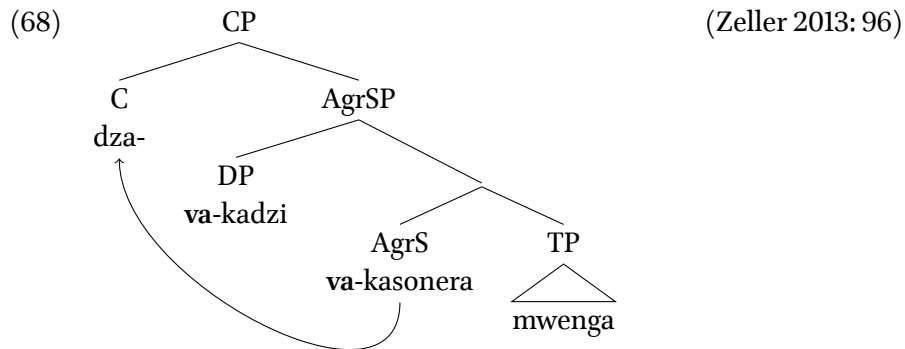
- (66) a. Mbatya dza-va-ka-son-er-a va-kadzi mw-engu.
 10.clothes 10REL-2SM-PST-sew-APPL-FV 2-women 1-bride
 ‘Clothes which the women sewed for the bride ...’ [Shona S10]
 (Demuth and Harford 1999: 42)
- b. e-nkáán e-wó-súm táar e-dzimmi
 4-book 4REL-PST-buy 1.father 4SM-disappear.PERF
 ‘The book that father bought has disappeared.’ [Mbuun B87]
 (Bostoen and Mundeke 2011: 93)

The agreement asymmetry between pronominal and lexical subjects, and the associated quirky class 1 SM alternation are also attested in non-subject relatives in some other West-Coastal Bantu (WCB) languages which Teke belongs to. For example in Hungan, when the postverbal subject is a class 1 lexical DP (67a), the SM on the verb can only take its canonical form *a-*; while the postverbal subject is a class 1 pronoun (67b), the SM can only occur as *ka-* which is interestingly also used in the context of non-subject focus in the matrix clause. It could be relevant to investigate more examples on this differential SM strategy with pronominal and lexical subjects in West-Coastal Bantu, which I leave for future research.

- (67) a. kit ki-a/(**ka*)-swiim-in Kipes
 7.chair 7REL-1SM-buy-PST 1.Kipes
 ‘the chair that Kipes bought’
- b. kit ki-ka/(**a*)-swiim-in yaan
 7.chair 7REL-1SM-buy-PST 1.PRO
 ‘the chair that s/he bought’ [Hungan H42] (Takizala 1974: 21)

After illustrating the agreement asymmetry phenomena in different (Bantu) languages, now I turn to investigate how postverbal subject agreement is realised in Kukuya. The subject agreement pattern in Kukuya non-subject relatives seems to be situated in between the Bantu languages above in (66), since it neither displays full agreement with all types of postverbal subjects, nor does it completely ban agreement with postverbal subjects.

If we compare the Shona example in (66a) above with Kukuya, we may wonder why agreement with lexical subject DPs cannot be realised in Kukuya non-subject relatives just like in Shona. The Shona type of agreement is accounted for by assuming that the subject has moved to the specifier of an agreement category such as AgrS or T and agrees with it, then the verb moves to C and precedes the subject (Demuth and Harford 1999; Zeller 2013), so the subject is still postverbal but shows full φ -agreement, as illustrated in (68) which is the derivation of (66a). The underlying structure of Kukuya non-subject relatives is different, since the subject remains *in situ*. I leave the explanation of these different agreement patterns and underlying positions of the postverbal subject in Bantu non-subject relatives for future research.



The derivation in (68) is consistent with the stipulation by Baker (2003), Collins (2004) and Carstens (2005) that an EPP feature on T in Bantu goes hand in hand with φ -features on T, and movement of the subjects to the preverbal position, namely specTP, is a part of the operation Agree and is a result of an EPP feature on T. This EPP feature is valued when T probes its c-commanding domain and finds a matching DP which can value T's [$u\varphi$] features, then T raises the agreeing DP to its specifier. Supposing that in Bantu, a subject which is placed in or has ever passed through specTP (specAgrSP, specFinP) must always shows A-agreement, which is also the case in SVO and subject relatives in Kukuya, the postverbal subject's not being able to (fully) agree with T in Kukuya (and Mbuun) should be attributed to the fact

that it is not raised but remains *in situ*, as I have shown previously. Now the question is how agreement with an *in situ* subject is realised in Kukuya.

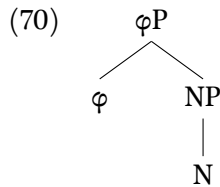
5.3.2.2 How does T agree with a postverbal subject?

Now let us consider why a postverbal subject shows less agreement (except for pronouns with [Person]) than a preverbal subject, and how featural distinction can help understand this asymmetry. Recall that when occurring postverbally, only pronouns with [Person] feature can show agreement with T, while lexical DPs and class pronouns can only take the default SM. If we compare a lexical DP with a person pronoun, we may intuitively conjecture that their different abilities in agreement may lie in their distinct size, since a lexical DP seems to be more complex and structurally “bigger” than a pronoun. This intuition brings me to seek an explanation in the defective goal framework by Roberts (2010) who proposes that a goal is “defective” if it has a *subset* of the features of the probe (Roberts 2010: 62) (see Iorio 2014, Van der Wal 2015, 2022 and Perry *to appear* for this account applied to subject and object agreement in Bantu languages and Irish), as defined in (69).

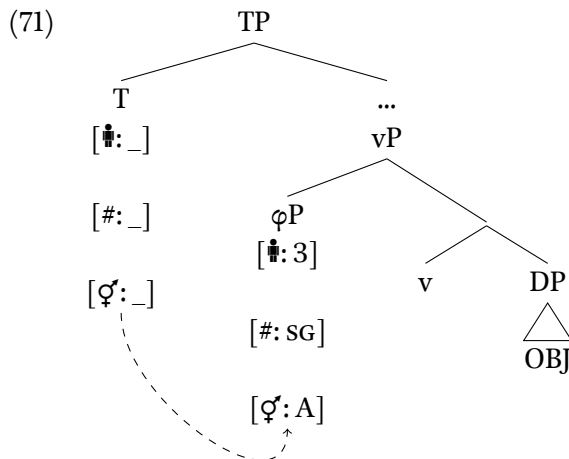
(69) Defective goal: (Roberts 2010: 62)

A goal G is defective iff G's formal features are a proper subset of those of G's Probe P.

To illustrate, a pronoun can be a φ P which is structurally smaller than a pronominal DP (Déchaine and Wiltschko 2002), as shown in (70), which only contains ϕ -features and no D-feature. When it agrees with a probe that contains [u ϕ] features and values them, the probe will contain the same valued features as the φ P and possibly more than those, while the φ P itself does not contain any features that are not present on the probe, thereby counting as a defective goal of the probe.



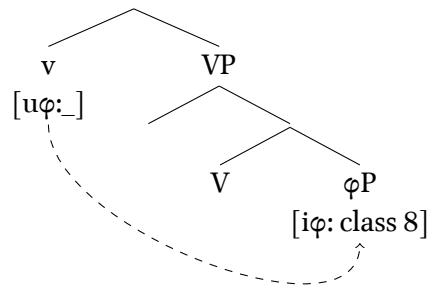
Consider (71), at the point when T is merged during the derivation with unvalued [uPerson], [uNumber] and [uGender] features, it probes downwards to search for matching interpretable features on a goal in its c-command domain, the subject in specvP is the closest goal and since in this case it is a ϕ P pronoun that contains the [Person], [Number] and [Gender] features which form an (improper) subset of the features on T, this subject ϕ P counts as a defective goal of T.



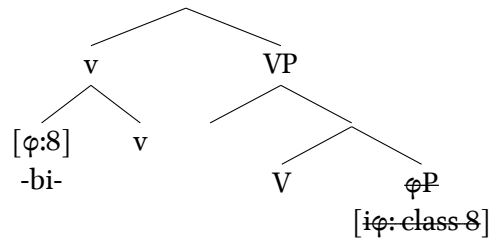
Upon agreement, there are two indistinguishable copies of the same set of features shared on the probe and the goal which form a chain, this leads them to undergo the PF-operation of chain reduction (Nunes 2004). The chain reduction takes place and the ϕ -features are only spelled out on the *higher* probe, copies other than the highest copy in a chain are deleted at PF. This process is comparable to the deletion of the lower copy after phrasal movement. Taking object marking in Bantu languages as an example to illustrate (Van der Wal 2022: 39), in (72a) the little v which is equipped with

[$u\phi$] features, probes down to find the internal argument with [$i\phi$]. If the object is a ϕ P whose nominal features are a subset of the probe's, it counts as a defective goal and after Agree the ϕ -probe on v will be spelled out as an object marker as in (72b).

(72) a.



b.



(van der Wal 2022: 39)

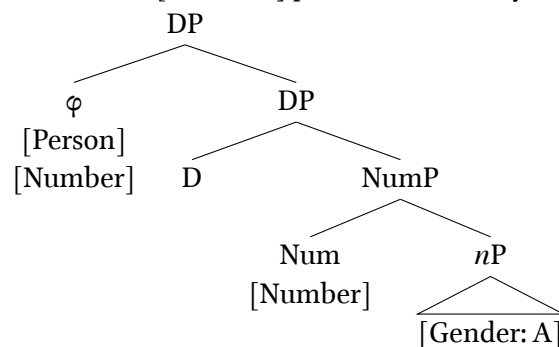
As mentioned, Iorio (2014) has applied the defective goal approach to account for subject marking in Bembe. In (73), we see that in Bembe object relative clauses the SM on the verb and the postverbal lexical subject DP are in complementary distribution. Iorio proposes that in (73a) what T agrees with is a defective subject ϕ P in *specvP* whose features are spelled out on T as the SM. If the subject is a full DP as in (73b), only the DP is spelled out and subject marking is impossible, as full DPs have features that are not represented in the probe, most remarkably a *lexical root* (Iorio 2014: 320).

- (73) a. *bi-lewa bi-ba-a-kol-á* (**batu*)
 8-food 8REL-2SM-N.PST-buy-FV 2.person
 “the food that they have bought”

- b. bi-lewa bi-(*ba)-a-kol-á batu
 8-food 8REL-2SM-N.PST-buy-FV 2.person
 “the food that the people have bought” [Bembe D54] (Iorio
 2015: 274)

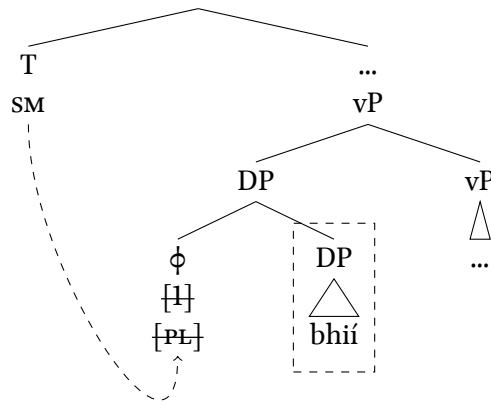
Applying this approach to Kukuuya, in which subject marking doubling is always attested with [+PERSON] pronouns in matrix and relative clauses, I propose that these pronouns including speech participants and *ndé/bó* are DPs that have a separate ϕ -layer in addition to the DP, as shown in (74). Here I suppose that an important difference between agreement with preverbal and postverbal subjects is that, in non-subject relatives T agrees with only the ϕ -features on the *outermost* head of the *in situ* subject DP instead of the whole DP, namely the extra ϕ -layer in (74). This is consistent with the Minimal Agree approach proposed by Kobayashi (2022) in the spirit of the labelling algorithm of Chomsky (2013, 2015). Under this approach, ϕ -featural agreement can apply in two ways: Minimal Agree targets only the outermost head of the phrase for reasons of minimal computation, while Full Agree matches all the ϕ -features of the phrase, probably for the labelling of Spec-Head configurations. For the agreement with a preverbal subject, I will return to discuss this later.

(74) Structure of [+PERSON] pronouns in Kukuuya



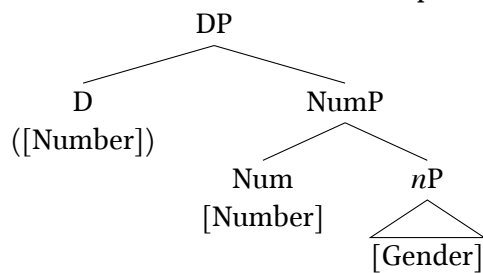
The DP with an extra ϕ -layer in (74) has been analysed as a “big DP” in different languages (Uriagereka 1995; Belletti 1999; Bleam 1999; Cechetto 2000) and Bax and Diercks (2012: 196) adopted this big DP hypothesis in accounting for Manyika object marking, in which the extra layer is itself

- (76) T agrees with the extra ϕ -layer of a [+PERSON] subject DP



As for lexical DPs and class pronouns, they lack the extra ϕ -layer ([Person]-layer) and the outermost D head cannot serve as a defective goal of T, since it should contain other D-features such as [Def] which do not have uninterpretable reflexes on T, no matter whether [Number] percolates to D, as illustrated in (77).

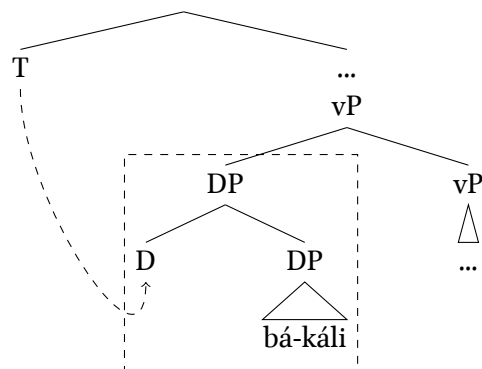
- (77) Structure of lexical DPs and class pronouns in Kukuya



In example (78), the *in situ* subject is a lexical DP *bá-káli* “wives” which does not have a separate [Person]-layer. In this case, agreement can still be established between T and the subject DP, but the ϕ -features are not spelled out on T since the goal is not deficient, and the whole DP is spelled out, as illustrated in (79).

- (78) ki-pfúo ki-kíí-télék-í bá-káli
 7-bread 7REL-7SM.PST-prepare-PST 2-wife
 ‘the bread that wives prepared’

- (79) T agrees with a postverbal lexical subject DP



However, in Kukuya there is a morphophonological requirement that the SM slot always be spelled out in some way. This may be due to the LH tone scheme of the REL-SM sequence which requires there to be two TBUs, or the reduced inflectional morphology of this language has to make use of the SM slot to avoid ambiguity (between subject and non-subject relatives). Actually this requirement is attested in many other varieties of Teke, so it may be extended to the whole Teke group (80). For example in Teke-Eboo (B74), when the postverbal subject is a lexical DP as in (81a), the SM on the verb can only take the class 15 prefix *u-* which is of the same shape as the infinitive prefix and is also used in expletive constructions as in (81b). The infinitive class prefix as a default SM in non-subject relatives has to my knowledge only reported in Teke within Bantu. It also supports the idea that a ϕ -probe that fails to agree spells out as default (Preminger 2009, 2014). The *ki-* prefix here can be viewed as an elsewhere exponent for the T probe.

- (80) The SM slot cannot be phonologically null in Teke relative clauses.

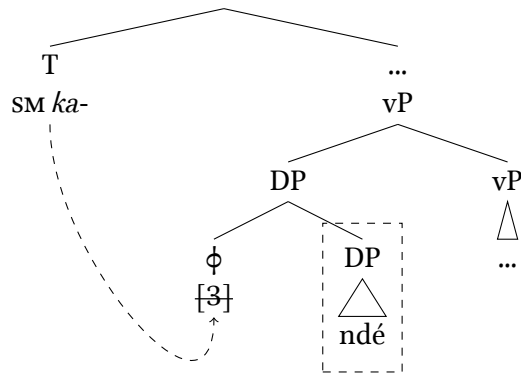
- (81) a. a-nzwo ma-ú-twii a-tara a bhi
 6-house 6REL-15SM-build.PST 2-father 2.CONN 1PL.PRO
 ‘the houses that our fathers built’
- b. U-faan-i we sal-i.
 15SM-should-PST 2SG.PRO work-SBJV
 ‘You have to work.’ [Teke-Eboo B74, Raharimanantsoa 2012]

To summarise here, for an *in situ* subject DP, T agrees only with the outermost head of the DP, and only spells out its φ -features when the goal is defective. Since the outermost layer of the [+Person] pronouns is defective, [+Person] pronouns show agreement with T. As for lexical DPs and other pronouns that do not have the extra layer, since the features on the D head do not form a subset of those on the T probe, it cannot count as a defective goal of T, and therefore the ϕ -features on T cannot get spelled out. The morphophonological requirement on a non-empty SM slot then instructs PF to insert a default morphology, which is the class 7 SM *ki-*.

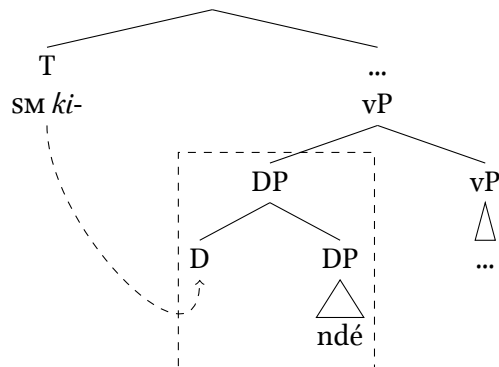
As for the proposed [+/-Person] differentiation between the two sets of pronouns which have the forms *ndé* and *bó*, the choice of SMS can be similarly explained as follows. In example (82), when the postverbal subject *ndé* is equipped with the extra ϕ -layer that contains a 3rd [Person] feature and agrees with T, the SM *ka-* is spelled out (see the lexical insertion rules in section 5.3.2.3) as shown in (83a); when it does not contain the extra layer, it serves as the canonical class 1 pronoun and does not provide any defective goal for T, thus taking the default SM *ki-* as in (83b). In both cases the pronominal DP *ndé* is spelled out. The alternation between the SMS *ba-/ki-* for the postverbal pronoun *bó* (see examples (55) and (56) above) can be accounted for in the same way.

- (82) Bhií líi-tsuk-í mu ku-káá/kíi-túr-í ndé
 1PL.PRO 1PL.SM.PST-talk-PST 18.LOC 17REL-1/7SM.PST-steal-PST 1.PRO
 mi-pará.
 4-money
 ‘We talked about (the fact) that s/he stole the money.’

- (83) a. T agrees with the 3rd [Person] pronoun *ndé*



- b. T agrees with the class 1 pronoun *ndé*



There are also examples like (84) and (85) below, in which the object relative clause lacks an overt subject, while the SM can surface as the default prefix *ki-*, or as an agreement marker that shows [Number] and [Person] specification but not [Gender]. I suppose that the SM *ki-* is just spelled out as the default marker as an elsewhere strategy when there is nothing to agree. The agreement marker appears when there is a postverbal covert *pro* subject, which may be structurally equal to a ϕ P, thus counting as defective with respect to the probe, but it can never reflect [Gender], namely the noun class of the antecedent subject in the matrix clause, as in example (85).

- (84) Ngúku áá-télek-i [ntsúú wu-káá/kíí-ték-i]?
 1.mother 1SM.PST-prepare-PST 1.chicken 1REL-1/7SM.PST-sell-PST
 ‘Did mother prepare the chicken to sell?’
- (85) Mi-féme míí-túr-i [ma-ko ma-kíí/báá/*míí-dzí].
 4-pig 4SM.PST-steal-PST 6-banana 6REL-7/2/*4SM.PST-eat.PST
 ‘The pigs stole the bananas to eat.’

So far I have provided an analysis of the subject agreement patterns with different types of postverbal subjects in Kukuya, adopting the defective goal approach. One crucial assumption under this analysis is the disassociation of EPP and Agree, which means that it is not just because the subject DP is *in situ* that renders it unable to agree, but the ϕ -features on T cannot be spelled out when it has no defective goal. This naturally brings out another question: if the agreement with a postverbal element indeed depends on whether the goal is defective, then how can we account for the full agreement with a *preverbal* subject?

Here I adopt the defective goal approach again as an alternative to the EPP-associated agreement, I follow van der Wal (2022: 204) and Perry (*to appear*)’s suggestion in proposing that the *movement* of the goal over the probe triggers spell-out of the features on the probe as well. That is to say, T has an EPP feature that is satisfied by raising the agreeing goal to its specifier, and the movement of the subject DP leaves behind a copy/trace which includes full ϕ -features that the moved DP contains, as shown in (86). This copy/trace is possibly comparable to a ϕ P pronoun and can count as defective, as shown in (87), so the ϕ -features including [Person], [Number] as well as [Gender] are spelled out on T. I suppose that here the agreement relation is not between the T probe and the trace but still with the subject DP, but it is the movement of the DP over the probe that triggers the spell-out of the features on T. In this way, the subject marking in SVO and subject/non-subject relatives can be unified.

are shown in (88). In the spirit of Distributed Morphology (Halle and Marantz 1993, 1994), lexical insertion rules may be underspecified such that for the first person singular SM /N-/, we only need to specify the [Person] feature but keep the singular value of the [Number] feature as default. The Subset Principle ensures that the most highly specific rule determines the spell-out form, so the SM is spelled out as /li-/ in the presence of a more specified [Number] feature [pl]. Note that in these two cases [Gender] is underspecified and does not affect the spell-out of the SMS. For SMS of different noun classes, I suppose that only the necessary feature values for [Gender] and [Number] are relevant, for example specified [Gender: C] and the [Number: pl] features trigger the spell-out of the class 6 SM *ma-*.

- (88) *Spell-out rules of subject markers in Kukuya*
- /N-/ \iff [1]
 - /li-/ \iff [1], [pl]
 - /Ø-/ \iff [2]
 - /ma-/ \iff [C], [pl]
 - /ki-/ \iff [D]
 - /bvi-/ \iff [D], [pl]

Now we consider the spell-out rules of the class 1/2 (and 3rd person) SMS. I suppose that *ka-* is spelled out in the context of [Person: 3], which is the only feature that needs to be specified. When the postverbal subject in a non-subject relative is a 3rd person pronoun *ndé* that has a separate ϕ -layer to host the interpretable [Person: 3] feature, T agrees with this deficient layer and the valued [uPerson] on T is spelled out as the SM *ka-*. Since I have proposed that [Gender] is not visible to T when the subject DP is postverbal, because it never percolates to the outermost head of that DP, I claim that *ka-* is not specified for [Gender]. Instead, the presence of [Gender: A] is responsible for the spell-out of the class 1 SM *a-*, which is a more specified feature than [Person: 3]. Assuming that class 1 DPs and person pronouns both contain [Gender: A], and that T can always agree with a preverbal DP in full ϕ -features, a class 1 lexical DP and the two sets of *ndé* always take the SM *a-* when occurring preverbally. The insertion of the SM *ba-* is rather special, since it only needs to reference the [Number: pl] feature, which I suppose to be the cases in example (85) above and in the impersonal *ba-* passives (see

chapter 3 section 3.3.2.2). The presence of a [Person: 3] feature in the context of agreement with a postverbal 3rd person *bó*, and that of the [Gender: A] when T agrees with a preverbal class 2 DP in SVO and subject relatives do not change the shape of the SM, which always gets spelled out as *ba-*.

- (89) *Spell-out rules of class 1/2 (and 3rd person) SMs in Kukuya*
 /ka-/ \iff [3]
 /a-/ \iff [A], ([3])
 /ba-/ \iff [pl], ([3]), ([A])

In this section, I have proposed an analysis of the position-dependent subject agreement in Kukuya relatives, mainly based upon the defective goal approach. Recall that one of the main aims of the chapter is to investigate the syntax of class 1 SM alternation in the IBV focus strategy; in the next section I look into the mechanism of the differential subject marking in the context of IBV focus, for the sake of better understanding the nature of this construction.

5.4 Subject marking alternation and IBV focus

Of interest for the current section is how to account for the class 1 SM alternation in the IBV focus strategy, which is connected to the agreement asymmetry in relatives that I just discussed. In this section I propose that the [Person] feature in Kukuya is in fact a discourse-related feature and is associated with both givenness and animacy. In an IBV non-subject focus construction with an initial class 1/2 subject, I suggest that in the derivation of the vP part, a discourse-linked ϕ P that contains a [Person: 3] feature is merged as the external argument of the verb in specvP and is agreed by T as a defective goal; and a topical element that controls this ϕ P is externally merged in a later stage of the derivation. I start the analysis by illustrating the structural representations of the cleft constructions and extend the analysis to the IBV focus construction. I also compare the SM alternation with the postverbal pronoun in Nzadi non-subject relatives to investigate their connections. To account for the mixture of monoclausal and biclausal properties of the IBV focus construction, I propose two underlying structures that synchronically co-exist. Some remaining questions for future research conclude the section.

Recall that the class 1 subject marking alternation happens in the circumstances below. As seen in (90), the *ka-* morphology is attested in the reduced cleft (90a) when a non-subject is focused sentence-initially, in the IBV non-subject strategy (90b), as well as in the negative sentence (90d); in SVO and IBV subject focus (90c), the SM appears as *a-*. The *a-* versus *ka-* class 1 SM alternation is also summarised in Table 5.4. From this table, the only apparent generalisation we can make is that *ka-* occurs when a non-subject is focused *ex situ*, and whether the subject is linearly adjacent to the verb does not correlate to the SM allomorphy, since we see that in the reduced cleft the class 1 subject is linearly adjacent to the verb. The occurrence of *ka-* in negative SVO may have separate motivations.

- (90) a. [Mí-féme]_{FOC} [mu-kái]_{TOP} káá-fúum-i.
 4-pig 1-woman ISM.PST-buy-PST
 ‘(It was) some PIGS that the woman bought.’ [reduced cleft]

- b. [Mu-kái]_{TOP} [mí-fémé]_{FOC} káá-fúum-i.
 1-woman 4-pig 1SM.PST-buy-PST
 ‘The woman bought the PIGS’ [IBV object focus]
- c. [Mi-féme]_{TOP} [mú-kái]_{FOC} áá-fúum-i.
 4-pig 1-woman 1SM.PST-buy-PST
 ‘The WOMAN bought the pigs.’ [IBV subject focus]
- d. Mu-kái ka-káá-fúum-i mi-féme ni.
 1-woman NEG-1SM.PST-buy-PST 4-pig NEG
 ‘The woman did not buy the pigs.’ [negative]

Context	word order	class 1 SM
canonical	SVO	a-
IBV non-subject focus	S(O)OV/S(O)XV	ka-
IBV subject focus	O(O)SV	a-
reduced cleft non-subject focus	OSV	ka-
negation	SVO	ka-

Table 5.4: The *a-* versus *ka-* class 1 SM alternation in different constructions

Since the class 1 SM *ka-* also occurs in the reduced cleft and the IBV focus strategy has a cleft origin, I start the discussion by looking into the structural representation of the basic and reduced clefts, which can lend support to the investigation on the SM alternation and the syntax of the IBV focus construction.

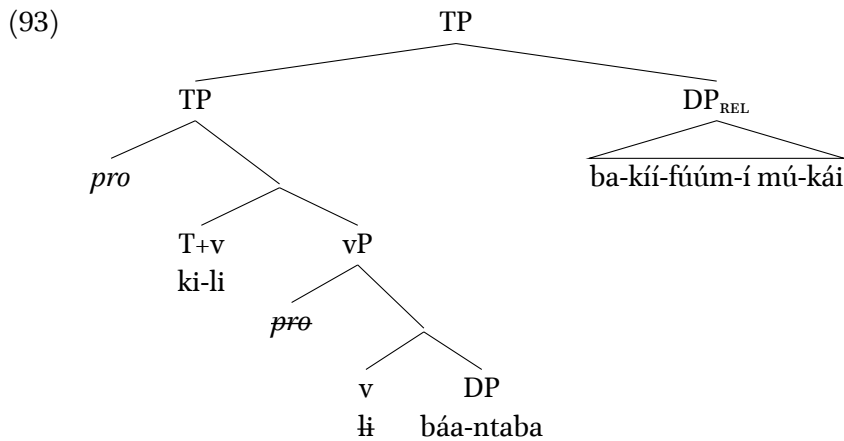
5.4.1 The derivation of cleft constructions

I start from investigating the syntactic derivation of clefts. A construction that I have labeled as a basic cleft is repeated in (91), which looks like a lot an *it*-cleft construction in English as shown in the translation. However, if we look at the construction carefully, we find that it differs from an English *it*-cleft in that the relative part of the sentence is actually not a modifying

relative clause but there is an overt relative pronoun *ba-* on the verb, and if we look at the pseudo-cleft (in which the copula is omitted) in (92), we see that the part *ba-kíí-fúúm-í mú-kái* is a free relative, which is translationally equal to “what mother bought” and is intrinsically a referring DP rather than a modifying CP, since it can refer to some entity by itself. This suggests that we are not dealing with a basic cleft in (91) but an identificational copular clause with an adjoined DP. The structural representation of (91) is thus illustrated in (93).

- (91) (Kí-li) báa-ntaba ba-kíí-fúúm-í mú-kái.
 7SM-COP 2-goat 2REL-7SM.PST-buy-PST 1-woman
 ‘It was some goats that the woman bought.’

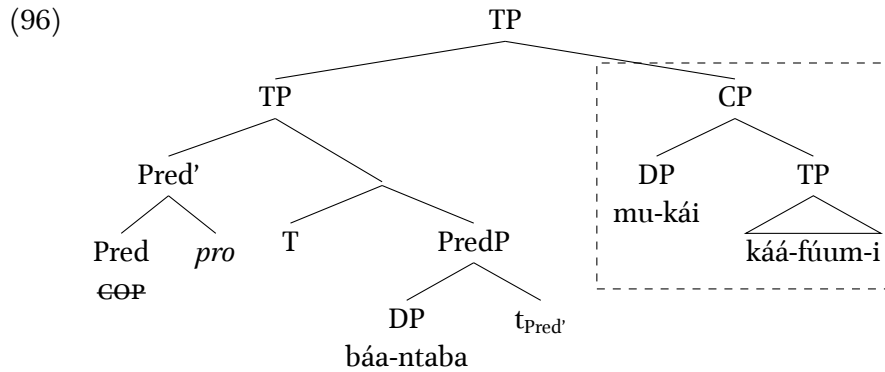
- (92) Ba-kíí-fúúm-í mú-kái báa-ntaba.
 2REL-7SM.PST-buy-PST 1-woman 2-goat
 ‘What the woman bought were some goats.’



Now let us consider the derivation of what I have labeled as the reduced cleft in (94) and (95). In these two examples there is no overt copula preceding the initial predicative NP, and no relative pronoun on the verb.

- (94) **Báa-ntaba mu-kái káá-fúum-i.**
 2-goat 1-woman LSM.PST-buy-PST
 ‘It was some goats that the woman bought.’
- (95) **Ngáŋwa taará káá-fúum-í báa-ntaba?**
 9.truth 1.father LSM.PST-buy-PST 2-goat
 ‘Is it true that father bought some goats?’

Here it is interesting to investigate how the reduced relative verb form, namely the deletion of the relative marker, can cancel the necessity of subject inversion as in the non-subject relatives. An obvious motivation of the subject fronting is that the deletion of the relative marker voids the morphophonological requirement on the adjacency of the relative marker and the verb, thus a preverbal subject becomes possible. If so, the subjects in (94) and (95) seem to be in specTP where they should have been (see the discussion on the *in situ* subject in section 5.3.1.2), we may wonder why the class 1 SM does not occur as the canonical shape *a-* but as *ka-*. Here I suppose that the preverbal subject must occupy a position higher than specTP, and the constructions in (94) and (95) actually correspond to a reduced basic cleft, in which the relative pronoun is also covert, therefore the relative part of the sentence (94) *mu-kái káá-fúum-i* is translationally equal to “that the woman bought” and the subject *mu-kái* “woman” is not the subject of a TP but of a CP which is adjoined to the predicative part, as illustrated in (96). Here I assume that the copula is a manifestation of the Predicate head (Bowers 1993; Adger and Ramchand 2003; den Dikken 2006; Cheng and Downing 2013).



Now it is time to magnify the adjoined CP part in (96) to see what its internal structure is like and how the SM *ka-* is spelled out. We first need to know which functional categories are suppressed in the reduced cleft. In an articulated CP system, I follow Belletti (2008, 2009, 2012) to assume that the CP in a cleft is a reduced CP that lacks at least the topmost ForceP, and possibly more projections are deleted. I suppose that in an embedded CP of a cleft in Kukuya, the ForceP, the higher TopP and the FocP are deleted, and the COMP should reside in Fin. Now the questions are what the T in the reduced CP agrees with and where the preverbal subject is placed in the articulated reduced CP. We first consider the subject agreement within the CP.

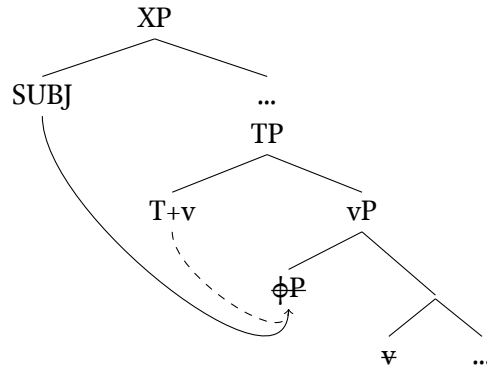
Recall that in the discussion of the grammaticalisation of the IBV focus strategy (chapter 4 section 4.4), to account for the class 1 SM alternation in the reduced cleft, I hypothesised an intermediate stage in which there was a VS/SVs alternation in Kukuya (reduced) clefts. This proposal was made by analogy to the non-subject relatives of Nzadi, in which the subject is commonly placed postverbally (VS), while a lexical subject NP can also appear in a preverbal position only if there is a co-indexing pronoun occurring immediately after the verb (SVs), as seen from the contrast in (97a) and (97b). Nzadi has also developed an initial focus strategy from the non-subject relatives (see chapter 4 section 4.4.2), as shown in (98), in which the presence of a preverbal subject does not require the postverbal pronoun.

- (97) a. èsúú nà ò món òkáár bàân
 day that PST see woman children
 ‘the day that the woman saw the children’
- b. èsúú nà òkáár ò món ndé bàân
 day that woman PST see she children
 ‘the day that the woman saw the children’
- [Nzadi B865] (Hyman 2012: 113)
- (98) ně bààr ó môn?
 who people PST see
 ‘Who did the people see?’ [Nzadi B865] (Hyman 2012: 107)

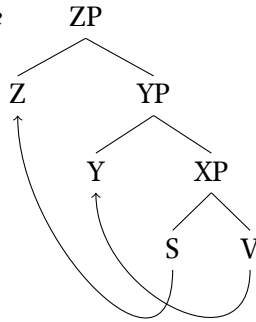
I suppose that the Nzadi focus construction in (98) is comparable to the Kukuya reduced cleft, and in both constructions the postverbal anaphor does not appear. I conjecture that like in Nzadi, at some historical point in a Kukuya reduced cleft, the presence of a preverbal subject also triggered a pronominal copy after the verb, which might have been an overt pronominal DP. In a later stage, this pronominal element is reduced to a φ P which is structurally smaller (see (70) above) than a pronominal DP (see (74) above). This φ P counts as a defective goal of T and agrees with the corresponding [u φ] on T, as illustrated in (99). The valued φ -features on T are spelled out as SM, while the φ P as the lower end of the chain has no component left to be spelled out thus becoming phonologically null.⁵

⁵As for Nzadi, the lack of subject agreement (Hyman 2012) in this language suggests that its T head has completely lost the [u φ] features, so the absence of the postverbal anaphor in (98) suggests that the pronominal element is simply deleted, rather than being reduced to a φ P and agreeing with T. Alternatively, in Nzadi non-subject relatives there was once also a postverbal φ P that is connected to the preverbal subject, and due to the loss of [u φ] on T, the unagreed φ P can only be spelled out as an overt postverbal anaphor.

(99) Agreement with a postverbal ϕ P in Kukuya



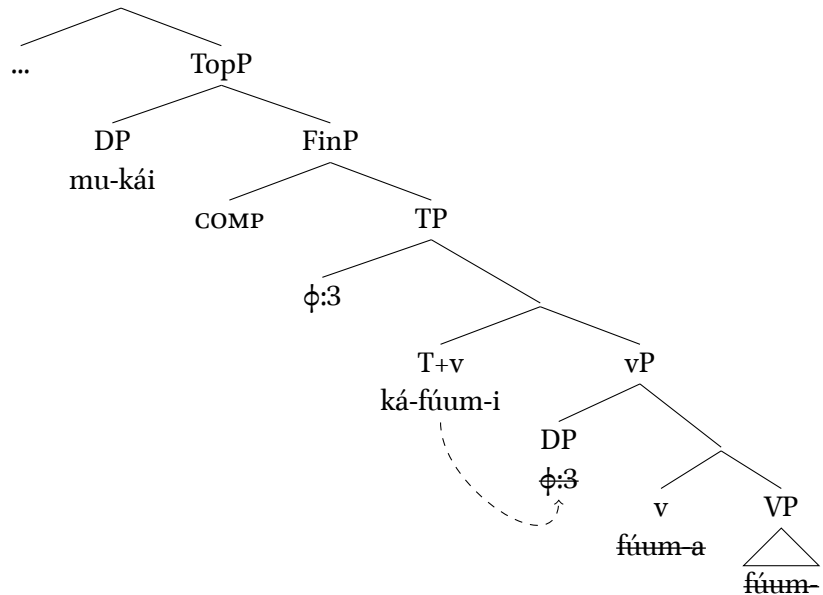
The Subject-Verb- ϕ P hypothesis being proposed, to present the full derivation of the Kukuya reduced cleft, we need to know the position of the preverbal subject, which in turn depends on its relation with the postverbal ϕ P. In Hyman (2012)’s discussion on the SVs structure in Nzadi non-subject relatives, the postverbal *s* appears superficially to be a copy/trace that the subject leaves when it moves to a higher position, as shown in (100). Hyman posed the question where the preverbal subject is located, namely what the ZP in (100) actually represents, and he found no pragmatic distinction between the SVs/VS structures, so it is unlikely that the movement of subject in Nzadi SVs is triggered by topic fronting. In example (101), the sequence *na ŋg* was analysed as a pronominal+*wh*-element, literally translatable as “that which, that whom” (Hyman 2012: 101) and is always optional. When there is a preverbal subject in an initial focus or a non-subject relative construction, it must occur between the *na* and *ŋg*, which suggests that the preverbal subject should occupy a high position in the CP domain, which I suppose to be between COMP head and the operator. This may also provide some hints on the structural position of the subject in the reduced cleft in Kukuya.

(100) *SVs Structure* (Hyman 2012: 104)

(101) *ibaa na mbvá_i ŋgo té nǔ_i*
 man DET dog WH.PST bite it
 ‘the man that the dog bit’ [Nzadi B865] (Crane et al. 2011: 10.20)

Here I propose that in a Kukuya reduced cleft what T agrees with is a φ P that functions as the external argument of the verb, and the preverbal DP is located in a position above TP, which is a base-generated topic in the C-domain. The preverbal topic is always co-referential with the φ P and the φ P can reflect the features on the topical DP. To illustrate, the derivation of example (94) above is partly shown in (102). We see that a φ P that contains the 3rd [Person] feature is merged in specvP and is probed by T. Since it is a defective goal for T, T agrees with it and the SM *ka-* is spelled out, while the φ P itself is deleted. The Fin head is merged next with the COMP feature. Then the TopP is merged and the topic *mukái* ‘woman’ is merged in specTopP. This approach is plausible, since it would explain why the class 1 SM does not take the canonical shape *a-* under the linear Subject-Verb adjacency. If the preverbal element originates from specvP and ultimately moves to a higher position above TP, it should have moved through specTP and triggered agreement on T. The dispreference for a preverbal indefinite non-specific subject in the reduced cleft also suggests that the preverbal DP should be distinguished from a grammatical subject in specTP. However, the base-generation analysis of the preverbal DP can bring the question on its exact relation with the co-referring φ P, considering that the φ P cannot be the trace/copy of a base-generated DP.

(102)



I propose that the ϕ P contains the 3rd [Person] feature possibly because it refers to a topical and animate element in the left periphery, and this is why the SM alternation is restricted to class 1 DPs. This proposal is reminiscent of the hypothesis that is introduced in Richards (2008, 2014), in which he proposes that DPs referring to referents on the high end of the animacy, definiteness have a [Person] feature. Van der Wal (2015) further links [Person] to givenness scales. Therefore, 1st and 2nd person are always animate and definite so always have the [Person] feature, and a 3rd person referent can be added to the [Person] feature when it is given and animate (also see Ormazabal and Romero 2007; Adger and Harbour 2007; van der Wal 2015, 2022). In Kukuya, only the ϕ P that refers to a topical class 1/2 DP bears this [Person] feature, possibly because the reconstructed Bantu gender A typically contained humans which are high on the animacy hierarchy (Denny and Creider 1976; Claudi 1997). This ϕ P corresponds to the extra [Person]-layer that I proposed for the person pronouns in the analysis of non-subject relatives (see section 5.3.2), which means that it mentions only [Person] and [Number] but not [Gender]. This is consistent with the idea that the ϕ P is not equal to a copy/trace. If it were a trace, presumably it should contain all the ϕ -features that the subject DP has,

including [Gender], but under the lexical insertion rules in section 5.3.3.3 the spell-out of *ka-* is not expected. The relation between the preverbal topic and the ϕ P may be a reflection on the historical association between a preverbal subject and a postverbal anaphor as in the SVs constructions introduced above, which still needs further research.

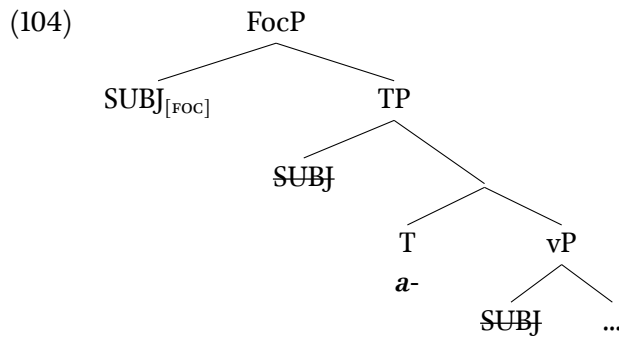
In summary, in a reduced cleft in which a non-subject constituent is focused, the preverbal subject-like DP is a base-generated topic in the left periphery but not the true argument of the verb. A ϕ P that is co-referential with the preverbal topic is merged in the external thematic position in *specvP* and it is equal to the 3rd [Person] feature when the preverbal topical DP is class 1/2. The SM *ka-* is the spell-out of the agreement relation between T and the ϕ P. In the next subsection I extend the analysis to account for the subject marking alternation in the IBV focus construction.

5.4.2 The derivation of the IBV focus construction

Now I discuss the class 1 subject marking alternation in the IBV focus construction, as repeated in (103).

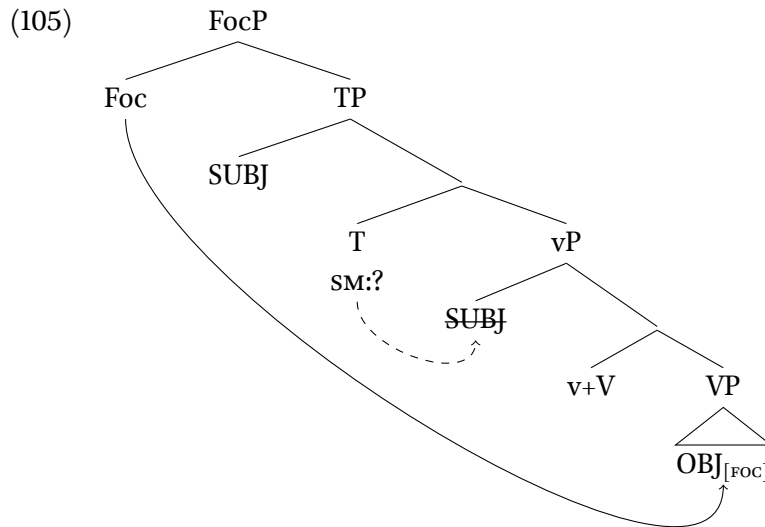
- (103) a. [Mu-kái]_{TOP} [mí-fémé]_{FOC} káá-fúum-i.
 1-woman 4-pig 1SM.PST-buy-PST
 ‘The woman bought the PIGS.’ [IBV object focus]
- b. [Mi-féme]_{TOP} [mú-kái]_{FOC} áá-fúum-i.
 4-pig 1-woman 1SM.PST-buy-PST
 ‘The WOMAN bought the pigs.’ [IBV subject focus]

We first consider the IBV subject focus construction. In section 5.2, I proposed that the IBV focused element is placed in the specifier of a FocP in the extended TP domain above the TP, which allows the Foc head to attract the subject with the [Foc] feature which has already agreed with T and has been raised to specTP, as illustrated in (104). Under this circumstance, the class 1 subject marker *a-* appears just as expected.



Next we think about the situation when a non-subject is focused in IBV. If the derivation were like in (105), when the Foc head is merged and attracting the object that bears the [Foc] feature, nothing seems to prevent the SM *a-* from occurring, since subject agreement has already taken place in the canonical way and the Foc is above TP so the focused object would not count as an intervenor for the agreement. However, since we know that the SM

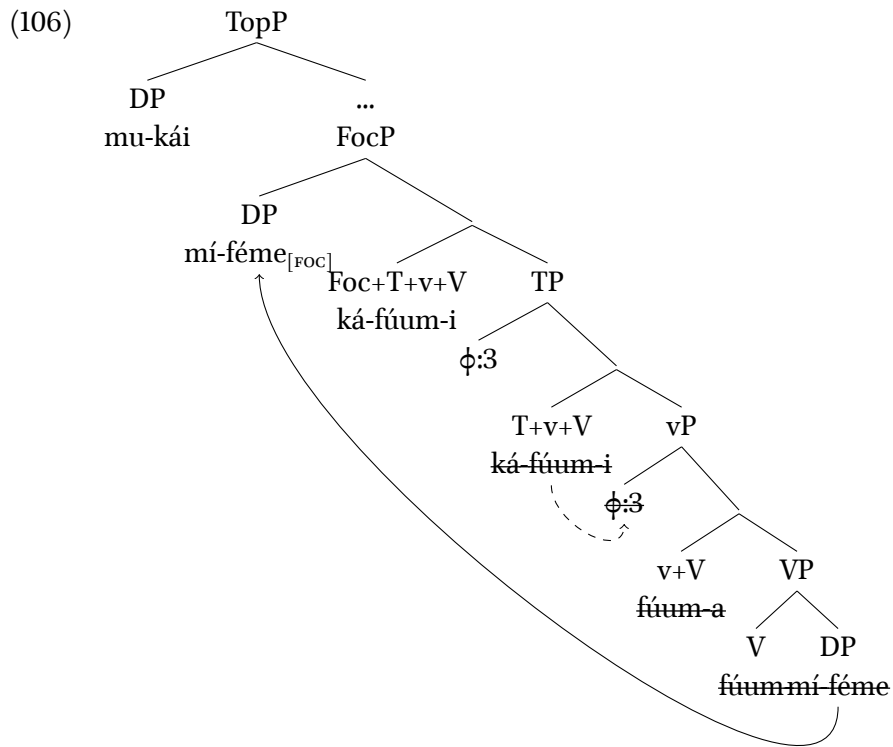
here must be *ka-*, so the derivation of the non-subject IBV focus construction must be different,



Recall that in the previous subsection I proposed that in a reduced cleft, the *SM ka-* represents the agreement between T and a ϕ P that equals to the 3rd [Person] feature. Since diachronically the reduced cleft is supposed to be the precursor of the IBV focus strategy, I propose that the subject agreement pattern in the latter may be explained by a similar mechanism. When the subject/agent is given or accessible in the context and the non-subject is focused in IBV, the subject agreement relation is established between T and a ϕ P which is co-referential with the clause-external initial topical agent. We consider the derivation of the IBV object focus example (103a) above, for which I provide the derivation in (106) and explain the procedures step by step next.

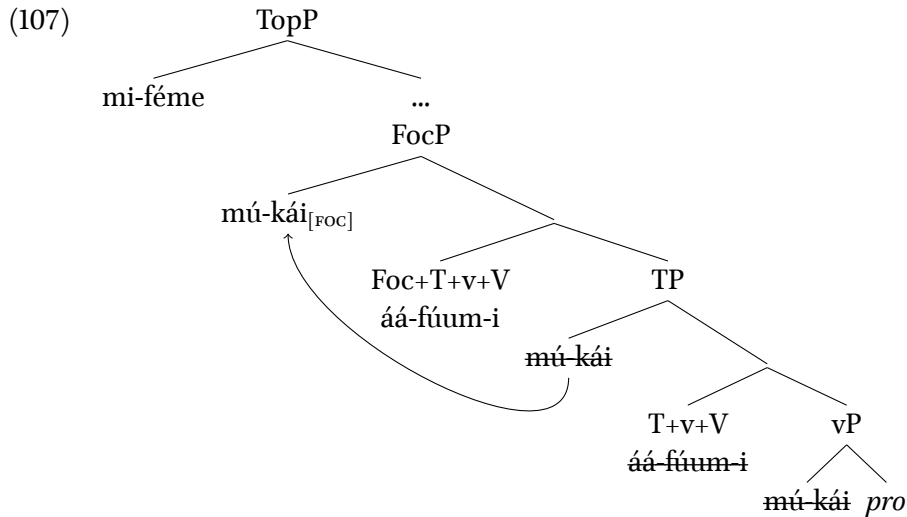
As (106) shows, in the derivation of the vP part which expresses the event, the external argument is a ϕ P that equals to 3rd person [Person] feature and encodes the animacy and givenness of the agent in the context. This ϕ P is agreed with and gets phonologically deleted when T is merged,

the SM *ka-* is spelled out on T and the verb head-moves to T to incorporate the final vowel. Then the FocP as an extended projection in the TP domain enters into the derivation and attracts the focused object *mi-féme* “pigs” to its specifier. The verb head-moves one step further to Foc, which may account for the strict Foc-Verb adjacency. Lastly, the topic DP *mukái* “woman” is base-generated in the specifier of TopP in the left periphery for discourse purposes and is co-referential with the [+Person] ϕ P.



In section 5.2.2 above I have shown that the IBV focused subject and non-subject are placed in one unique structural position, thus the structural analysis in (106) should also be applicable to IBV subject focus construction. The derivation of (103b) is given in (107), in which the initial topical “object” is a base-generated topic. In this way the structural representation of

IBV subject and non-subject focus constructions are unified, since both involve an initial base-generated topic and a clause-internal focused element. A crucial distinction between IBV subject and non-subject focus is in the vP part: when a non-subject is focused, the initial topic controls a clause-internal ϕ P that in turn controls subject marking on the verb, since subject marking is always required on the Kukuya verb; when the subject is focused in IBV, the initial topical element does not control any overt index on the verb, since object marking has been completely lost in this language, so I just put a *pro* in the object position in the derivation.



The separation of the initial topic from the rest of the sentence and the use of the ϕ P may be probably a repair strategy for the IBV focus construction in which there are more than one preverbal elements, since a preverbal focused and a preverbal topical element, for example when they are both in class 1, would be hard to be distinguished if subject marking keeps the same form. I will discuss this more in the next section.

5.5 Summary and remaining questions

This chapter started from investigating the structural representation of the canonical SVO word order. I showed that the subject is in an A-position, namely specFinP and the verb can head-move to T or stay in its base position. Then I provided an analysis on the structural position of the IBV focused element. I first showed that the IBV element has A'-moved from the lower part of the clause, and there is one unique structural position that is available for both IBV focused subject and non-subject. Based on these I compared the pros and cons of the low and high FocP approaches and concluded that the IBV focused element is placed in a FocP in the extended TP domain above TP. In the discussion on the class 1 subject marking alternation, I first deviated to discuss the subject agreement asymmetry in relative clauses. I discussed the different subject agreement patterns of postverbal person pronouns and the other DPs. Adopting the defective goal approach, I hypothesised that downward agreement only targets at the outermost layer of the goal and an extra [Person]-layer on the person pronouns plays the key role in agreement. I also proposed that the pronouns *ndé/bó* are underlyingly two sets of pronouns, namely class 1/2 and 3rd person singular/plural, which differ in feature specification. In the previous section I proposed that the structure of IBV subject and non-subject focus constructions is unified in that they can both involve an initial base-generated topic, and the rest of the sentence encodes the event. In an IBV non-subject focus construction, the class 1/2 initial topic controls a clause-internal ϕ P that has the 3rd person value and agrees with T, and this accounts for the spell-out of the SM *ka-*. In this section I mention and discuss some remaining questions that are relevant to the analyses in this chapter.

As for the 3rd [Person] and the ϕ P hypothesis, there is also language-internal evidence on the existence of a 3rd [Person] feature. Here I just mention the *ba*-passive construction, agreement pattern in the *say*-complementisers and the hortative imperative sentences. I have introduced the functional *ba*-passive construction in chapter 3, which I consider to be a case of a grammaticalised 3rd person feature. Some examples are given in (108)-(110). In this case the feature is always valued as plural and lacks an overt

referent.

(108) (*visual stimuli: what about the food?*)

Bviilá **báá-tél-i** bví ku mfúúlá.

8.food 2SM.PST-throw-PST 9.falling 17.LOC 9.road

'The food was thrown onto the road.'

(109) Kĩ-sí we kí-má ké ki-báá-túr-i?

7REL-make.PST 2SG.PRO 7-what 7.PRO 7REL-2SM.PST-steal-PST

'What that you made was stolen?'

(110) (*You found that the bananas on the table disappeared, and you asked father.*)

Ma-kó ná báá-wí?

6-banana 1.who 2SM.PST-give.PST

'The bananas were given to whom?'

The paradigm of *say*-complementisers (see chapter 2 section 2.7) is repeated in Table 5.6. We see that the complementisers do not show [Gender] agreement but only [Person] and [Number]. In example (111) the sentential subject is in class 4 and is animate, while the complementiser takes a 3rd plural form *bóri*. Here I suppose that the complementiser may agree with a covert *pro* which is co-referential with and controlled by the logophoric center of the embedded clause, but this *pro* only shows [Person] and [Number] distinctions, which is comparable to the ϕ P that is controlled by the initial topic in the IBV focus construction.

(111) Mi-fémé mí-kâ-tsuomó **bóri** me ɲé â-n-siiba.

4-pig 4SM-IMPf-think 2.COMP 1SG.PRO 4.PRO FUT-1SG.SM-attack

'The pigs think that I will attack them.'

Another possible piece of evidence on the 3rd person feature may be in the hortative imperative sentences. We see that in (112a) the SM on the verb

Person	singular	plural
1st	píri	píri
2nd	wurí	wurí
3rd	ndíri	bóri

Table 5.5: The agreeing pattern of the “say”-complementiser *-ri*

takes the *ka*- prefix with a H tone, with the verb-final L tone marking the imperative mood and blocking the metatony. Here it is unclear whether this *ka*- is the same as that in the IBV focus construction or it belongs to other categories. This prefix is not likely to be the present *kâ*-, not only because of the tone on it, but also because if it is the present aspect marker, we should expect that it also occur in (112) when the SM is plural. Here I conjecture that the SMS in (112) may also show the existence of the 3rd person feature, since in the hortative imperative there is never an overt subject but the implicit agent is a 3rd person. So I conjecture here that this *ka*- SM prefix may be the same as that in the IBV focus strategy. Since I have proposed that the *ka*- only occurs within an embedded CP, it is also interesting to see that French translation “qu’il ferme la porte” also formally corresponds to a modifying CP.

- (112) a. Ká-kúra ki-dzulibi!
 1SM.IMP-close 7-door
 ‘(Let) him close the door!’
- b. Bá-fúuma ma-li!
 2SM.IMP-buy 6-wine
 ‘(Let) them buy the wine!’

As for the base-generation analysis of the initial topic in the IBV focus construction, it seems that the construction should be treated as a “topic-event” structure which places a lexical dislocated topic in the initial position and the comment forms a complete clause which may contain a co-indexing ϕ P. So a sentence with IBV focus like “Father bought some GOATS” is underlyingly interpreted as “As for father, *he* bought some GOATS.”

There is prosodic evidence showing the contrast between the initial topic and the rest of the topical domain. In a construction such as (113), the speakers noted that in natural speech there should be a prosodic break after *taará*, but *mwáana* and *kímá* are phrased together, which can be detected from the H tone plateauing effect. This can be taken as evidence that the initial topic is dislocated but the secondary topic may be more integrated to the rest of the sentence.

- (113) Taará, mwáána kímá ká-sî?
 1.father 1.child 7-what 1SM.PST-do.PST
 ‘What had the father done to the child?’

Since there is no object marking on the verb in Kukuya, the occurrence of object marker cannot be used to diagnose the syntactic status of the fronted objects. In Kukuya only resumptive pronouns can be used to co-refer to a topicalised object, as shown in (114). But resumption is always optional in Kukuya according to most speakers. For some speakers, the resumption is obligatory or preferred when the fronted patient is a speech participant as in (115), but there are both inter- and intra-speaker inconsistencies on this judgement, while there are also speakers who rejected the speech participants to be fronted. The resumptive marking is therefore not a very reliable diagnostic to investigate the syntactic status of the fronted objects for Kukuya, but it should be an interesting area for further research.

- (114) a. Mu-tí ma-dzá taará á-mwáal-i (mó).
 3-tree 6-water 1.father 1SM-water-PST 6.PRO
 ‘FATHER watered the tree.’
 b. Mu-tí ma-dzá taará á-mwáal-i (ndé).
 3-tree 6-water 1.father 1SM-water-PST 3.PRO
 ‘FATHER watered the tree.’
- (115) Me mpfúúmú á-sak-í (me).
 1SG.PRO 1.chief 1SM.PST-search-PST 1SG.PRO
 ‘I was searched by the chief.’

In the IBV focus constructions like (116) and (117), the initial topical elements “father” and “child” are obviously not the argument of the verb but function as the primary topic of the sentence. The rest of the sentence is actually an event that the topic undergoes. For (116) the topic is “father” and a question on the event he experiences is “where did the hoe disappear?”; in (117) the topic is the child and the event s/he undergoes is the passing of the festival. I hypothesise that the initial element of the IBV focus construction does not make reference to the grammatical subject but the sentential topic, and the rest of the sentence of an event that is connected to the topic. It should be noted that here I only suggest a possible way of analysis, since I don’t have relevant data for all the syntactic tests, but I provide a direction of further research, which may turn out to be ultimately only apparent.

- (116) Taará **téme** ku-ní lí-dzinim-i?
1.father 5.hoe 17-which 5SM.PST-disappear-PST
‘Where did father lose the hoe? (*lit*: As for father, where did (his) hoe disappear?)’
- (117) Mwáana **ki-yinga** bu-ní kí-wir-i?
1.child 7-festival 14-which 7SM.PST-pass-PST
‘The child, how did the festival pass (for him/her)?’

There are also some remaining questions in this chapter. In the previous section I proposed that in the situation of IBV non-subject focus, a clause-internal ϕ P that equals to the 3rd [Person] feature is merged in specvP and refers to the agent that is given in the context, and a co-referential topical element is generated in the left periphery. Under this approach, a DP which does not belong to class 1/2 should not have a ϕ P counterpart in the derivation. Consider the IBV focus example in (118) in which the initial topic is not in class 1/2 but class 4, we see that the verb just take the class 4 SM *mú-*, if we keep the base-generation analysis of the initial element, how can we account for the insertion of the class 4 SM? Here I suppose that the class 4 topical DP may also control some ϕ -bundles that can address the [Gender] feature for T to agree, and [Person] is still restricted to class 1/2 DPs. Note that this is different from the situation in non-subject relatives in which the

postverbal DP's [Gender] feature is not shown on the outermost layer thus is not visible to T.

- (118) Mi-féme má-ko míi-túr-i.
 4-pig 6-banana 4SM.PST-steal-PST
 'The pigs stole some BANANAS.'

The occurrence of the class 1 SM *ka-* in negation, as repeated in (119), also remains to be explained. I suggest that it may also be accounted for by assuming that the verb in a negative sentence agrees with a ϕ P that has the 3rd [Person] feature instead of agreeing with a preverbal subject DP, and the preverbal class 1 DP in negation actually occupies a higher position. The prohibition on the use of the negative prefix *ka-* on a relative verb may be due to its competition with the relative marker for the same slot on the verb complex, which may also suggest that the preverbal DP is in a higher position in the C-domain. I leave this for future research,

- (119) a. Mwá wúa áá-kwî.
 1.dog 1.DEM 1SM.PST-die.PST
 'That dog died.'
- b. Mwá wúa ka-káá-kwî ni.
 1.dog 1.DEM NEG-1SM.PST-die.PST NEG
 'That dog did not die.'

Recall that in chapter 4, I showed that the IBV focus construction displays many monoclausal properties such as 1) deletion of the relative marker; 2) focus projection; 3) and conjoint prosodic phrasing. Nevertheless, it has still retained some characters of a biclausal cleft: 1) presence of the copula in negation; 2) H tone on the SM; 3) verb-final H tone in non-subject extraction; and 4) unavailability of the negative prefix *ka-*. I suppose that since the IBV focus strategy is still an ongoing innovation towards a monoclausal construction, different structures may be synchronically mixed up in the speakers linguistic knowledge. Monoclausal and biclausal constructions can co-exist. In the context of negation the biclausal structure is chosen, possibly because of the intervention of the negative prefix; while for example in interrogatives, the IBV focus construction is more monoclausal.

CHAPTER 6

Conclusion

This chapter first summarises some main findings on the word order and information structure in this language and indicates some larger theoretical questions that the thesis addresses. The second section mentions some remaining puzzles throughout the thesis that are left for future research.

6.1 Main findings in the thesis

In this section, I summarise the main conclusions on the syntax and information structure of Kukuya, and reflect upon some broader theoretical questions that the thesis addresses. I mainly discuss the following questions: the interface between syntax and information structure in Kukuya; the nature of “immediate-before-verb”; the diachronic and synchronic approaches; and the “economic” characteristics that Kukuya shows.

From the description in chapter 3 of the expression of information structure

in Kukuya, we see that there is considerable variation in its word order that is influenced by information packaging. The word order of this language seems to be better captured by reference to discourse roles such as topic and focus than to syntactic roles such as subject and object, which can be seen in two facts: the first is the use of a dedicated IBV focus position that is available for arguments including subject and object as well as for adjuncts; and the initial referent in the IBV focus construction always has a topic function in the sentence, regardless of whether a subject or a non-subject is focused in IBV. However, sentence configuration in Kukuya does not completely depend on information structure, since I also showed that focus can also be expressed *in situ* for objects and adjuncts, SVO order can also express theticity and the IBV position can be also used as pragmatically neutral. If there is indeed a continuum between grammatical role-oriented and discourse role-oriented that places individual languages in different points (Kerr et al. 2023), Kukuya should be located in a position that is more towards discourse-configurational, and information-structural notions should indeed be encoded in the syntax of Kukuya. In chapter 5, I proposed that while the class 1 SM *a-* is preserved under subject focus, the SM *ka-* is used for subject agreement with a 3rd [Person] ϕ P which is co-referential with a topical agent under non-subject focus, thus Kukuya employs word order in conjunction with subject agreement to encode argument relations and information structure.

In the thesis I provided both the diachronic and synchronic accounts of the IBV focus strategy, and we can see how the two approaches complement and inspire each other. I first hypothesised that the IBV focus strategy has a cleft origin based on the shared grammatical properties between the two constructions, which builds an important baseline for the synchronic analysis. I associated the SM *ka-* of the postverbal subject *ndé* ‘s/he’ in non-subject relatives and the SM *ka-* in IBV non-subject focus, also by analogy to the SVs construction in Nzadi non-subject relatives, from which the ϕ P hypothesis arose. I treated the *ka-* as an agreement prefix and analysed the mechanism of its spell-out from the generative point of view under the Minimalist framework. What I did not discuss is the historical source of the morpheme *ka-*, which is conjectured to historically originate from an identification copula in Bostoen and Mundeke (2012) for Mbuun. The synchronic representation of IBV focus also inspired diachronic analysis,

for example the mixed mono-clausal and bi-clausal properties on this construction which implies the ongoing grammaticalisation process.

One main goal of the research was to investigate how the strict Foc-Verb adjacency can be accounted for from the structural representation of the IBV focus construction. In chapter 5, I proposed that the IBV element is placed in the specifier of a high FocP above TP and the verb head-moves to the Foc head, which ensures the focused phrase to be always linearly left-adjacent to the verb. The verb-final H tone occurrence in the situation of non-subject focus, which is distinct from the tone pattern in canonical SVO order, may suggest a different position of the verb. Under this approach, the “immediate-before-verb” effect is just explained by the spec-head adjacency. It would be interesting to compare the IBV effect with the structural analyses on the IAV focus in Bantu, which in some studies is explained by the altruistic movement of non-focal elements out of the vP (Cheng and Downing 2012), and the right-branching FocP approach that accounts for the sentence-final focus (Ndayiragije 1999). However, in Kukuya there seems to be no other independent evidence for the T-to-Foc movement of the verb, on which further studies still need to be carried out.

Throughout the thesis, I also left some further research questions, in the next subsection I go through these questions.

6.2 Remaining research questions

In this section, I give a summary of some interesting phenomena in Kukuya that have been described or mentioned in the previous chapters, which are left unaddressed in the thesis and need more research in the future. I briefly (re)introduce some but all remaining research questions following the order of the chapters.

In the grammar sketch, I showed that the diminutives in Kukuya are formed by the partial republication of the noun stem, which is accompanied by vowel modification and tone change, as shown in the repeated example in (1). It is worthwhile investigating how the republication process and the tone assignment are realised, for example in a framework such as Distributed Morphology.

- (1) *kii-ku-ko* “small banana”
bii-bû-baana “small children”
ki-bĩ-bilí “small lola tree”
ki-sî-saka “small gourd”
ki-mbu-mbaa “little fire”
ki-nzû-nzo “small house”
ki-yũ-yũ “small peanut”

In the introduction of interrogative words, I mentioned another intriguing phenomenon which is the use of an agreeing pronoun immediately after the interrogative words *ná* “who” and *-má* “what kind/type”, as shown in examples (2) and (3). It is still not clear what the exact function of the pronoun is and what the relationship is between it and the interrogative element. In example (4), when an interrogative word occurs in a postverbal position in an embedded clause, it must be preceded by the marker *ka*, whose function is still not clear either.

- (2) Ndé ki-bhiimá kíí ngúku ya ná ndé
 1.PRO 7-corpse 7.CONN 1.mother with 1.who 1.PRO
 kâ-dziik-a?
 1SM.FUT-bury-FV
 ‘With whom will s/he will bury the corpse of mother?’
- (3) Maamá kí-má ké kâ-dzií kí-yáab-a?
 1.mother 7-what 7.PRO 1SM.IMPF-like INF-know-FV
 ‘What does mother want to know?’
- (4) Ba-ntsúú ba-kíí-ká-í *(ka) ná bá-yiká bá-bí?
 2-chicken 2REL-7SM.PST-grill-PST EMP 1.who 2SM-IMP 2-bad
 ‘The chicken that who grilled is getting bad?’

Some other questions left in the grammar sketch are for example the interpretative distinctions between the basic and emphatic forms of connectives, the functions of some occasionally occurring verbal affixes, and the use and agreement patterns of *say*-complementisers, which can only be further studied when new field data is available.

In the subject focus expressions introduced in chapter 3, I showed that preverbal focus is always allowed but a discourse-linked *wh*-subject such as “which” and “whose” cannot take the canonical SM like other *wh*-subjects. Instead, a cleft construction must be used, as shown in (5). The reason why preverbal D-linked *wh*-subjects are prohibited needs to be explained.

- (5) a. *Mwáana wu-ní á-mún-i Zacharie?
 1.child 1-which 1SM.PST-see-PST Zacharie
Int.: ‘Which child saw Zacharie?’
- b. Mwáana wu-ní wǔ-mún-i Zacharie?
 1.child 1-which 1REL-see-PST Zacharie
lit.: ‘Which child is the one who saw Zacharie?’

In chapter 3 section 3.2.2, I showed that verb focus can also be expressed via the SOV order, as shown in (6). The answer to a question like “what did

X do to Y?” can also involve the use of the IBV focus position. It is somehow surprising to see SOV order expressing verb focus, since the focus is neither placed on the IBV element nor projected to the whole VP, it would be interesting to investigate how the IBV position can express verb focus.

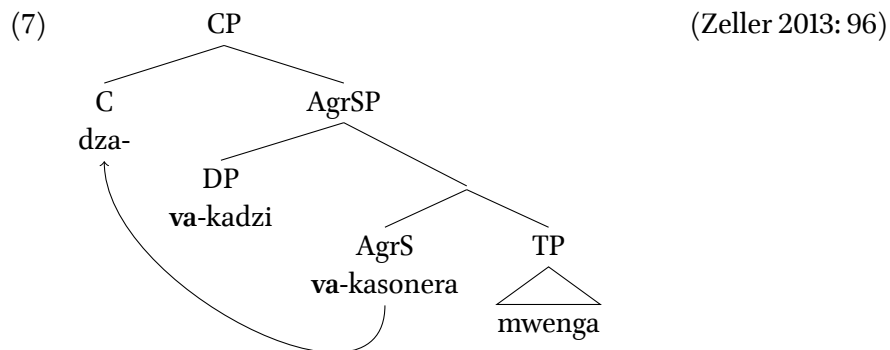
- (6) a. Ngolo Marie kí-má ká-sî?
 Ngolo Marie 7-what 1SM.PST-do.PST
 ‘What did Ngolo do to Marie?’
 b. Ngolo Marie ká-pfur-í.
 Ngolo Marie 1SM.PST-cheat-PST
 ‘Ngolo betrayed Marie.’

Some further questions on the information-structural expression that I did not discuss in detail in chapter 3 also include the relative order of multiple secondary topics, the expression of contrastive topics, and the pragmatically neutral use of the IBV position. All these questions need to be addressed in future studies.

At the end of chapter 4, I presented the Table 4.10 to show the micro-variation in the grammatical properties of Kukuya, Mbuun and Kisikongo that have all been reported to make use of the IBV focus strategy. It would be interesting to study if there are any dependencies or correlations among the features in the table. Compared to Kisikongo, it is intriguing to investigate why Kukuya lacks the subject inversion constructions. Inspired by the Subject-verb-pronoun construction in Nzadi non-subject relatives, I hypothesised an intermediate stage where there was one also a postverbal pronoun that is co-referential with the preverbal subject in the grammaticalisation process in order to account for the class 1 SM alternation. It would be worthwhile discovering more similar patterns on this SVs construction in neighbouring languages, and to study how this is related to the verb-final H tone in non-subject relatives.

To account for the syntax of subject agreement in non-subject relatives, I proposed that while a preverbal subject always shows full ϕ -agreement, agreement with an *in situ* subject only targets at the outermost head/layer of the subject DP but not the DP as a whole, and in this situation agreement

is impoverished for [Gender] because it is embedded in the *nP* and is never shown on the outermost head. The question is why the subject in Kukuya non-subject relatives cannot fully agree with T by being raised to specTP and T to C movement takes place to derive the linear Verb-Subject order, just like the non-subject relative in Shona as shown in (7). The parametric distinctions on the position of the subject and the consequent different agreement patterns between languages like Shona and Kukuya necessitate further studies.



The reason why there is the [+/-Person] distinction on the pronouns *ndé* and *bó* that I hypothesised in chapter 5 section 5.3.1.2 is not entirely clear at this point, and the speakers do not report any pragmatic difference between the choice of the SMS *ka-/ba-* versus *ki-*. A potential explanation would be that historically in Kukuya non-subject relatives there was once a postverbal pronoun that co-indexed with an animate and given preverbal subject and was reanalysed as bearing the 3rd [Person] feature; and diachronically the SVs construction is lost but the [Person] on the postverbal *ndé* and *bó* has for some reason maintained but is being weakened, so the *ki-* morphology becomes available. For the SM alternation under IBV non-subject focus, I proposed that the external argument of the verb is a φ P, and a co-referent DP is generated in the left-periphery. It would also be interesting to study the association between the historical SVs construction and the synchronic representation of the IBV non-subject focus construction. All these puzzles need to be explained by comparing Kukuya with more West-Coastal Bantu

language that display similar agreement patterns.

In chapter 5 I also proposed that the IBV focused element is placed in the specifier of a FocP that is above TP. Given the fact that postverbal focus is also available and can appear without any IBV element (see chapter 3 section 3.1), some relevant questions would be what factors motivated the language to innovate the IBV focus strategy; and if focus is syntactically present as the [Foc] feature, what does the postverbal focus tell us about the nature of that feature? Does it show that [uFoc] is located on FocP rather than the focused phrase?

The question mentioned above are all issues that have come to light because of the analyses of the information structure and syntax of the Kukuya language in this thesis. It would be interesting to see how these approaches can be applied to other West-Coastal Bantu languages to investigate the interactions among the IBV focus strategy, tone patterns and subject marking alternation, which are left for future studies.

APPENDIX A

Sample text

Conversation between wife and husband

The following text is an excerpt of a mock conversation in Kukuya between two male speakers who pretended to act as the wife and the husband and to have a quarrel. The “wife” is labelled as A and the “husband” as B. The glossing on some morphemes in the texts still need further investigation.

- (1) A. Maa-ndáka ma-kítĩ me ka-n-kukí mu
6-word 6-COMP 1SG.PRO NEG-1SG.SM-can 18.LOC
ki-kiele ni.
INF-tolerate NEG
‘The things that I cannot tolerate.’
- (2) B. We kâ-saká mí-pará nziimi!
2SG.PRO HAB-search 4-money much
‘You ask for much money!’

- (3) B. We yáab-i wurí me ka-n-lí ya
 2SG.PRO know-PST 2SG.COMP 1SG.PRO NEG-1SG.SM-COP with
 mi-pára ni!
 4-money NEG
 ‘You knew that I did not have money.’
- (4) B. Li-m-bák-í me, me m-wî me
 PST-1SG.SM-get-PST 1SG.PRO 1SG.PRO 1SG.SM-give.PST 1SG.PRO
 ki-m-wâ we, wũ me ka-ká-mínima ni.
 DEP-1SG.SM-give 2SG.PRO 1REL 1SG.PRO NEG-1SM-be.seen NEG
 ‘When I got (something), I did give you but (my giving) was not
 seen.’
- (5) A. We ka-maá-múna wurí baa-ndúku báá me
 2SG.PRO NEG-PERF-see 2SG.COMP 2-friend 2.CONN 1SG.PRO
 baa-mó, bá-láa kí-kó ki-kí-tsítse.
 2-other 2SM-wear 7-clothes 7REL-7AGR-clean
 ‘You have not seen that, my other friends, they wear clean
 clothes.’
- (6) A. Ki-ko ki kíí me wuna ki-fúum-i we
 7-clothes 7REL 7.CONN 1SG.PRO only INF-buy-PST 2SG.PRO
 ngakímó, ná tsúku ki-fá ntséke kâ-láa wúna
 7.one every 5.day DEP-come.from 9.field PRS-wear only
 ké kía.
 7.PRO 7.DEM.II
 ‘You bought only one piece of clothes for me, everyday I finish
 (working in) the fields I wear only that one.’
- (7) A. Ki-n-yá hé ki-m-fá ntséke
 DEP-1SG.SM-go also DEP-1SG.SM-come.from 9.field
 n-yok-í ma-dzá, m-bvúruk-i
 1SG.SM-take.bath-PST 6-water 1SG.SM-return-PST
 n-lá-i ké kía.
 1SG.SM-wear-PST 7.PRO 7.DEM.II
 ‘When I go and I come from the field, when I took bath, I always
 wore that piece of clothes.’

- (8) A. We *kâ-múna wurí* we *ki-báká mí-pará mii*
 2SG.PRO PRS-see 2SG.COMP 2SG.PRO DEP-get 4-money 4.CONN
we wúna ma-li ki-ɲwâ.
 2SG.PRO only 6-wine INF-drink
 ‘You see that if you get your money, you only drink wine.’
- (9) A. We *kâ-múna wurí* *búa* *kí-bvé li-kâ-yé*
 2SG.PRO PRS-see 2SG.COMP 14.DEM.II 7-good 1PL.SM-PRS-go
mu ntsá nzó?
 18.LOC inside 9.house
 ‘Do you think it’s good that we run the family like this?’
- (10) B. We *kâ-laká wurí* *me* *kâ-ɲwâ* *ma-li,*
 2SG.PRO PRS-say 2SG.COMP 1SG.PRO HAB-1SG.SM.drink 6-wine
maa-ndáka ma bu-kâ-n-súuk-i me
 6-problem 6.DEM.I 14REL-HAB-1SG.SM.solve-PST 1SG.PRO
ku nzó aa ngúku aa we.
 17.LOC 9.house 9.CONN 1.mother 1.CONN 2SG.PRO
 ‘You say that I drink wine, (but) I solved the problems of your maternal family.’
- (11) A. We *kâ-súuka wúna maá* *ku nzó aa ngúku*
 2SG.PRO PRS-solve only 6REL 17.LOC 9.house 9.CONN 1.mother
aa we, maá ku nzó aa ngúku aa
 1.CONN 2SG.PRO 6REL 17.LOC 9.house 9.CONN 1.mother 1.CONN
me, we ka-kâ-súuka ni.
 1SG.PRO 2SG.PRO NEG-HAB-solve NEG
 ‘You solve only the problems of your family, the problems in my family you do not solve.’
- (12) B. We *kâ-múna wúna maá* *kukí me,* *we* *kâ-laká*
 2SG.PRO PRS-see only 6REL PREP 1SG.PRO 2SG.PRO PRS-say
wurí me ka-kâ-m-wá we ni?
 2SG.COMP 1SG.PRO NEG-PRS-1SG.SM-give 2SG.PRO NEG
 ‘You only see what I do for my side, you say I give you nothing?’
- (13) A. *Bi-kâ-wá* *we me bí-ma?*
 8REL-PRS-give 2SG.PRO 1SG.PRO 8-what
 ‘What do you give me?’

- (14) B. Ki-bvé we ɲa nzó aa me li,
 7-good 2SG.PRO 16.LOC 9.house 9.CONN 1SG.PRO COP
 ka-kâ-m-wá ni, we a-bvĩ nsiina
 NEG-HAB-1SG.SM-give NEG 2SG.PRO PST-fall.PST 9.ground
 ɲa kíí me?
 16.LOC 7.CONN 1SG.PRO
 ‘It is good that you are always at my house, (if) I do not give you,
 (why) you settle in my home?’
- (15) A. Me ɲa m-bvĩ nsiina mu báana
 1SG.PRO 16.PRO 1SG.SM-fall.PST 9.ground 18.LOC 2.children
 ba-lí-maá-búra bhií.
 2REL-1PL.SM-PERF-give.birth 1PL.PRO
 ‘I stay here for the children that we have given birth to.’
- (16) A. Bu-n-yí me ɲa kíí we
 14REL-1SG.SM-come.PST 1SG.PRO 16.LOC 7.CONN 2SG.PRO
 ɲa, bu yiká baa-mbvĩ li-maá-búra
 16.DEM.I CONJ ?IMPF 2-white.hair 1PL.SM-PERF-give.birth
 báana.
 2.children
 ‘Since I came to your home, I now have white hair and we have
 had children.’
- (17) B. Bu bá-maá-tó bá-nna, káli nzaamí ka-ká-sí
 CONJ 2SM-PERF-arrive 2-four COND 1.god NEG-1SM-do.SBJV
 bú-bí ni, kéne báana bá-yiká ba-táani nké.na.nké
 14-bad NEG EMP 2.children 2SM-?IMPF 2-five CONJ
 ba-séneme?
 2-six
 ‘As we have four children, if the God does not do bad things,
 don’t we have five or six?’
- (18) B. Mu we, li-báká líí me, we wurí
 18.LOC 2SG.PRO 5-property 5.CONN 1SG.PRO 2SG.PRO 2SG.COMP
 me n-yé n-túr-á mu we!
 1SG.PRO 1SG.SM-go 1SG.SM-steal 18.LOC 2SG.PRO
 ‘Because of you, my possessions, you say I’m going to steal be-

cause of you!’

- (19) A. We sé bi-yé túra, we kukí mu ki-balá
 2SG.PRO even NEG-go steal 2SG.PRO can 18.LOC INF-smelt
 nkwî bu-kítî a-téke a-báká mú-pára
 9.firewood 14-COMP 2SG.SM-sell 2SG.SM-sell 3-money
 a-wá me, me m-fúum-i sé sabúnu.
 2SG.SM-give 1SG.PRO 1SG.PRO 1SG.SM-buy-SBJV even 5.soap
 ‘Even if you do not go to steal, you can smelt the firewood in
 order to sell and earn money and give the money to me, I could
 buy even the soap.’
- (20) B. Kǎn mu li-we li-n-yikí me lía,
 CONJ 18.LOC 5-strength 5REL-1SG.SM-?IMPF 1SG.PRO 5.DEM.II
 me m-bvúruka n-somó mu-súru mu kí-balá
 1SG.PRO 1SG.SM-return 1SG.SM-enter 3-forest 18.LOC INF-smelt
 nkwî mu li-we li-n-yikí me
 9.firewood 18.LOC 5-strength 5REL-1SG.SM-?IMPF 1SG.PRO
 lía.
 5.DEM.II
 ‘As my strength is now like this, I return to the forest to smelt the
 firewood?’
- (21) A. Me yiká sabúnu ni, me kâ-m-bvuoká
 1SG.PRO ?IMPF 5.soap NEG 1SG.PRO HAB-1SG.SM-massage
 sabúnu li-bá-kâ-swaaká ma-sáani.
 5.soap 5REL-2SM-HAB-wash 6-plate
 ‘I don’t have soap, I have the massage with the soap with which
 people wash plates.’
- (22) A. We wî me mi-pará, me m-bák-i
 1SG.PRO give.SBJV 1SG.PRO 4-money 1SG.PRO 1SG.SM-get-SBJV
 mu m-fúúmá kí-ko yă sabúnu.
 18.LOC 1SG.SM-buy 7-clothes with 5.soap
 ‘You give me money, I get it to buy the clothes and the soap.’
- (23) A. Ba-kái baa-mó wúna bu ba-kǎ-kála.
 2-woman 2-other only 14.PRO 2SM-PRS-stay
 ‘The other women only stay (in their families) in this way.’

- (24) A. Ma-tála, se mfú yii baa-mó,
 IMP-look even 10.hair 10.CONN 2-other
 bu-n-kíni me bu, mfú ki-tsúu
 14REL-1SG.SM-?be.attractive 1SG.PRO 14.DEM.I 10.hair INF-plait
 ni.
 NEG
 ‘Look at others’ hair, (though) I am still presentable, (I could not) plait the hair.’
- (25) A. Bá-tsúa yii taata, yii me yiká kâ-kalá wúna
 2SM-plait 9.CONN well 9.CONN 1SG.PRO ?IMPF PRS-stay only
 yii bu yii.
 9.CONN CONJ 9.CONN
 ‘They plait the hair well, my hair always stays like this.’
- (26) B. Nyáka nzaamí kâ-nyanima bi-dí-desu
 let.IMP 1.God 1SM.PRS-recompense 8-DIM-bean
 bvi-n-sí me bvía.
 8REL-1SG.SM-do.PST 1SG.PRO 8.DEM.II
 ‘If the God recompense all the beans that I worked for.’
- (27) B. Taará aa we á-sí-i bu-ta, lí-téke
 1.father 1.CONN 2SG.PRO 1SM.PST-leave-PST 14.gun 1PL.SM-sell
 bu-ta bu-kítĩ li-baká mi-pára mi-kítĩ lí-kâ-fúuma
 14-gun 14-COMP 1PL.SM-get 4-money 4-COMP 1PL.SM-PRS-buy
 bi-lóko.
 8-thing
 ‘Your father left a gun, we sell the gun to get money in order to buy things.’
- (28) A. ɲaa we yi-ká yii ndáka kâ-sénema ku
 now 2SG.PRO ?IMPF 9.CONN 9.problem PRS-meet 17.LOC
 ntsá maa-nzó máá me, wurí yi-ká
 inside 6-house 6.CONN 1SG.PRO 2SG.COMP ?IMPF
 ba-bá-kâ-dzwá me.
 2REL-2SM-PRS-kill 1SG.PRO
 ‘Then you are meeting a big problem with my family members, they will kill me!’

- (29) B. Bu-tá búú nzó ki-taará, wurí n-téke,
 14-gun 14.CONN 9.house 7-father 2SG.COMP 1SG.SM-sell
 n-téke bu-ní?
 1SG.SM-sell 14-which
 ‘The gun of (your) father’s family, (if) I sell, how do I sell?’
- (30) A. We ki-yé kuki yă nzáli, pfúr-i wurí bu-tá
 2SG.PRO DEP-go PREP with 1.river lie-SBJV 2SG.COMP 14-gun
 bú-bvĩ mu ma-dzá.
 14SM-fall.SBJV 18.LOC 6-water
 ‘You go to where there is a river, you may lie that the gun falls
 into the water.’
- (31) A. ŋa ku-kâ-tsuomó we wurí me
 16.LOC 17REL-PRS-think 2SG.PRO 2SG.COMP 1SG.PRO
 m-bvúruka n-kalá ŋa kíí we wó me
 1SG.SM-return 1SG.SM-stay 16.LOC 7.CONN 2SG.PRO or 1SG.PRO
 n-yé?
 1SG.SM-go
 ‘Now you think that I stay here or I go?’
- (32) B. Búa bu-kâ-múna we, we ka mbuurú
 14.DEM.II 14REL-PRS-see 2SG.PRO 2SG.PRO EMP 1.person
 á-maá-múna ku ntáli, ka bu ní?
 2SG.SM-PERF-see 17.LOC 9.other.place NEG 14.PRO NEG
 ‘So as you see, you have met with a person in other place, isn’t
 it?’
- (33) A. Me ka ku-kítĩ me n-yé ku
 1SG.PRO NEG 17-COMP 1SG.PRO 1SG.SM-go 17.LOC
 maá-múna mbuurú mu kí-ma ni.
 PERF-1SG.SM.see 1.person 18.LOC 7-what NEG
 ‘If I leave, that is not because I met with someone else.’

- (34) A. Me li ma-n-yúka bóri tsúku ngalímo
 1SG.PRO RPST PERF-1SG.SM-hear 2.COMP 5.day 5.one
 mu-kái wu-téneme we ku balabála, we tá
 1-woman 1REL-meet 2SG.PRO 17.LOC ?9.street 2SG.PRO ?EMP
 maá-fúuma ndé ki-mbúuni.
 PERF-buy 1.PRO 7-?cloth
 ‘I heard that one day you met a woman on the street, and you
 bought her a piece of cloth.’
- (35) B. Wũ-lak-í we ndáka yía, wúa wũ-ka-dzií
 1REL-say-PST 2SG.PRO 9.parole 9.DEM.II 1.DEM.II 1REL-PRS-like
 we wó wũ-bel-í bu li we ŋa kí
 2SG.PRO or 1REL-dislike-PST 14REL COP 2SG.PRO 16.LOC 7.CONN
 me?
 1SG.PRO
 ‘The one who said this to you, he likes you or he does not like
 that you are in my house?’
- (36) A. Mu-káli aa we, ndé-nkúlu ya baa-ndúku
 1-wife 1.CONN 2SG.PRO 1.PRO-RFL with 2-friend
 á-lak-í ndíri ki-ko ki ki-kí-fúúm-í
 1SM.PST-say-PST 1.COMP 7-clothes 7.DEM.I 7REL-7SM-buy-PST
 má-kángu máá me.
 6-?fiance 6.CONN 1SG.PRO
 ‘(It was) your ‘woman’, herself with her friends, she said that the
 clothes were bought by my husband.’
- (37) B. Me tá hé a-maá-n-yuká kúa líí
 1SG.PRO ?EMP also PST-PERF-1SG.SM-hear 17.DEM.II 5.CONN
 we hé nziimí nziimí, me a-m-bí m-khúula.
 2SG.PRO also much much PST-1SG.PRO 1SG.SM-NEG 1SG.SM-ask
 ‘I also heard a lot (what was said) about you, but I did not ask.’

- (38) A. Me ka-n-yáab-i ba-balaka ba-kima ni ba-kítí
 1SG.PRO NEG-1SG.SM-know-PST 2-male 2-other NEG 2-COMP
 we lak-í me ni, bu-n-yá me
 2SG.PRO say-PST 1SG.PRO NEG 14REL-1SG.SM-come 1SG.PRO
 ɲa kíí we, we a-maá-múna me
 16.LOC 7.CONN 2SG.PRO 2SG.PRO 2SG.SM-PERF-see 1SG.PRO
 n-yení ki-mokó ku nzó aa mbuurú?
 1SG.SM-go.away INF-?date 17.LOC 9.house 9.CONN 1.person
 ‘I did not know the other man you said, since I came to your house, you have seen me go to meet others in anyone’s house?’
- (39) A. Wúna ɲa-kâ-ya ndé ɲa bu, me
 only 16REL-HAB-come 1.PRO 16.PRO 14.PRO 1SG.PRO
 ki-n-túmá ndé ki-balá me li-kwî, ndé
 DEP-1SG.SM-use 1.PRO INF-smelt 1SG.PRO 5.firewood 1.PRO
 bal-í me li-kwî, bhíí yă ndé
 smelt-PST 1SG.PRO 5-firewood 1PL.PRO with 1.PRO
 ka-lí-li ya li-bála ni, kéné we
 NEG-1PL.SM-COP with 5-connection NEG EMP 2SG.PRO
 maá-múna?
 PERF-see
 ‘Only when he comes here, if I ask him to smelt the firewood for me, he smelts the firewood for me, we don’t have relationship, have you seen that?’
- (40) B. Ka me hé a-n-yuk-í, li-tála mhoî ku
 EMP 1SG.PRO also PST-1SG.SM-hear-PST 1PL.SM-look 6.all 17.LOC
 ma-tsúku máá bú-su.
 6-day 6.CONN 14-front
 ‘But I did hear (that), let us both see the days to come.’
- (41) B. Ma-tála ku-lí-fí, ma-tála
 IMP-look 17REL-1PL.SM.PST-come.from.PST IMP-look
 ku-lí-yé.
 17REL-1PL.SM.FUT-go
 ‘(Let us) see where we came from and where we will go.’

- (42) A. Me n-dzii kéné me n-kal-í ɲa
1SG.PRO 1SG.SM-like EMP 1SG.PRO 1SG.SM-stay-SBJV 16.PRO
mu kâ-sá mú-sálá yǎ bé.
18.LOC HAB-do 3-work with 2PL.PRO
'(Then) I like to stay here to work with you.'

Bibliography

Abandzounou, Prince. 2012. Description phonologique de l'Iboõ, parler bantu de la zone B de la République du Congo. MA thesis, Université de Marian Ngouabi, Brazzaville.

Abels, Klaus and Peter Muriungi. 2008. The focus marker in Kĩtharaka: Syntax and semantics. *Lingua* 118 (5): 687-731.

Aboh, Enoch O. 2007. Leftward focus versus rightward focus: The Kwa-Bantu conspiracy. *SOAS Working Papers in Linguistics* Vol. 15.

Adger, David and Gillian Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34 (3): 325-359.

Andrews, Avery D. 2007. The Major Functions of the Noun Phrase. *Clause Structure, Language Typology and Syntactic Description, Vol. 1: Clause Structure* (Second Edition), Timothy Shopen (ed.), 132–223. Cambridge: Cambridge University Press.

Baka, Jean. 2000. L'adjectif en bantu. Bruxelles: Université libre de Bruxelles, thèse de doctorat.

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford Univer-

sity Press.

Baker, Mark. 2003. Agreement, dislocation, and partial configurationality. In A. Carnie, H. Harley, and M. Willie (eds.), *Formal Approaches to Function in Grammar*, 107–32. Amsterdam: John Benjamins.

Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

Bastin, Yvonne. 1989. El prefijo locativo de la clase 18 y la expresión del progresivo presente en Bantu. *Estudios africanos* 4: 35-55; 61-86.

Bax, Anna and Michael Diercks. 2012. Information structure constraints on object marking in Manyika. *Southern African Linguistics and Applied Language Studies* 30 (2): 185–202.

Bearth, Thomas. 2003. Syntax. In Derek Nurse and Gérard Philippson (eds.), *The Bantu Languages*, 121-142. London, New York: Routledge.

Béjar, Susana. 2003. *Phi-syntax: A theory of agreement*. Toronto, ON: University of Toronto.

Béjar, Susana and Milan Rezac. 2009. Cyclic agree. *Linguistic Inquiry* 40 (1): 35-73.

Belletti, Adriana. 1999. Italian/Romance clitics: Structure and derivation. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe*, 543–579. Berlin: Mouton de Gruyter.

Belletti, Adriana. 2004. Aspects of the low IP area. In L. Rizzi (ed.), *The structure of CP and IP*, 16–51. Oxford: Oxford University Press.

Belletti, Adriana. 2008. The CP of clefts. *Studies in Linguistics*: 7.

Belletti, Adriana. 2012. Revisiting the CP of Clefts. *Discourse and grammar. From sentence types to lexical categories*: 91-114.

Bennett, Patrick R., Ann Biersteker, Withira Gikonyo, Susan Hershberg, Joel Kamande, Caroyne Perez and Martha Swearingen. 1985. *Gikuyu ni Kioigire: A first course in Kikuyu*. Published by the African Studies Program of the University of Wisconsin-Madison.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23 (4): 757-807.

Bjorkman, Bronwyn M. and Hedde Zeijlstra. 2019. Checking up on (ϕ -) Agree. *Linguistic Inquiry* 50 (3): 527-569.

Blanchon, Jean A. 1998. Semantic/pragmatic conditions on the tonology of the Kongo noun phrase: a diachronic analysis. In Larry Hyman and Charles Kisseberth (eds.), *Theoretical Aspects of Bantu Tone*. Stanford, California: CSLI Publications.

Boeckx, Cedric. 2009. On Long-Distance Agree. *An International Journal of Theoretical Linguistics* 111 (1): 1-32.

Bonnafé, Pierre. 1973. Une grande Fête de la vie et de la mort: le miyali, cérémonie funéraire d'un seigneur du ciel kukuya (Congo-Brazzaville). *L'homme*: 97-166.

Bowers, John. 1993. The syntax of predication. *Linguistic inquiry* 24 (4): 591-656.

Bonnafé, Pierre. 1979. *Nzo lipfu, le lignage de la mort: la sorcellerie, idéologie de la lutte sociale sur le plateau kukuya*. Société d'éthnologie Vol. 5.

Bonnafé, Pierre. 1987. *Histoire sociale d'un peuple congolais Livre I, La terre et le ciel*. Paris: ORSTOM.

Bostoen, Koen and Léon Mundeke. 2011. Passiveness and inversion in Mbuun (Bantu B87, DRC). *Studies in Language* 35 (1): 72-111.

Bostoen, Koen and Léon Mundeke. 2012. Subject marking, object-verb order and focus in Mbuun (Bantu, B87), *Southern African Linguistics and*

Applied Language Studies, 30 (2): 139-154

Bostoen, Koen and Joseph Koni Muluwa. 2019. Nsong (B85d). In M. Van de Velde et al. (eds.), *The Bantu Languages* (Second Edition), 414–448. London: Routledge.

Bostoen, Koen, Bernard Clist, Charles Doumenge, Rebecca Grollemund, Jean-Marie Hombert, Joseph Koni Muluwa and Jean Maley. 2015. Middle to late Holocene Paleoclimatic change and the early Bantu expansion in the rain forests of Western Central Africa. *Current Anthropology*, 56 (3): 354-384

Bostoen, Koen and Joseph Koni Muluwa. 2021. The Immediate Before Verb Focus Position in West- Coastal Bantu: Some Comparative Data. Workshop presentation presented at the *Typological perspectives on focus marking in African languages*, Online workshop, May 27.

Botne, Robert. 1995. The pronominal origin of an evidential. *Diachronica* 12 (2): 201-221.

Botne, Robert. 2020. Evidentiality in African Languages. *Evidentials and Modals*. Brill, 460-501.

Bouka, Léonce Yembi. 1994. L'accord des adjectifs du protobantou en zone H. *Africana linguistica XI*, Tervuren, pp. 13-18.

Buell, Leston C. 2005. Issues in Zulu morphosyntax. PhD dissertation, University of California Los Angeles.

Buell, Leston. 2006. The Zulu conjoint/disjunct verb alternation: Focus or constituency?. *ZAS Papers in Linguistics* 43: 9-30.

Buell, Leston. 2009. Evaluating the immediate postverbal position as a focus position in Zulu. In *Selected proceedings of the 38th annual conference on African linguistics*, 166-172. Cascadilla Proceedings Project Somerville, MA.

Calloc'h, Jean-René. 1911. *Vocabulaire français-ifumu (batéké) précédé*

d'éléments de grammaire. Paris: Greuthner.

Carstens, Vicki. 2008. DP in Bantu and Romance. In Cécile de Cat and Katherine Demuth (eds.), *The Bantu–Romance Connection: A Comparative Investigation of Verbal Agreement, DPs, and Information Structure*, 131-165. Amsterdam: John Benjamins.

Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23 (2): 219-279.

Carstens, Vicki. 2011. Hyperactivity and Hyperagreement in Bantu. *Lingua* 121 (5): 721-741.

Carstens, Vicki and Michael Diercks. 2013. Agreeing how? Implications for theories of agreement and locality. *Linguistic inquiry* 44 (2): 179-237.

Carstens, Vicki and Loyiso Mletshe. 2015. Implications of Xhosa expletive constructions. *Linguistic Inquiry* 46 (2): 187-242.

Cecchetto, Carlo. 2000. Doubling structures and reconstruction. *Probus* 12: 93-126.

Chafe, William. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (ed.), *Subject and Topic*, 27-55. New York: Academic Press.

Cheng, Lisa L.-S. and Laura Downing. 2007. The prosody and syntax of Zulu relative clauses. In N. C. Kula and L. Marten (eds.), *Bantu in Bloomsbury: SOAS Working Papers in Linguistics* 15, 51–63. Available at <https://www.soas.ac.uk/linguistics/research/workingpapers/volume-15/>.

Cheng, Lisa L.-S. and Laura J. Downing. 2012. Against FocusP: arguments from Zulu. *Contrasts and positions in information structure*: 247.

Cheng, Lisa L.-S. and Laura Downing. 2013. Clefts in Durban Zulu. In K. Hartmann and T. Veenstra (eds.), *Cleft Structures*, 141-164. Amsterdam: John Benjamins.

Chomsky, Noam. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers*. MIT, Cambridge, MA.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In K. Hale and S. J. Kayser (eds.), *The View from Building 20*, 1–52. Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Chomsky, Noam. 2005. Three factors in Language Design. *Linguistic Inquiry* 36: 1–22.

Chomsky, Noam. 2008. On phases. In R. Freidin, C. P. Otero, and M. L. Zubizarreta (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133–66. Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In E. D. Domenico, C. Hamann, and S. Matteini (eds.), *Structures, strategies and beyond*. Amsterdam: John Benjamins.

Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, Massachusetts: MIT Press.

Claudi, Ulrike. 1997. Some thoughts on the origin of gender marking. In R. K. Herbert (ed.), *African Linguistics at the Crossroads*, 63–74. Cologne: Rüdiger Köppe.

Cobbinah, Alexander and Friederike Lüpke. 2009. No cut to fit—zero coded passives in African languages. In Mathias Brenzinger and Anna-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics*, 153–165.

Cologne: Köppe.

Collins, Chris. 2004. The agreement parameter. *Triggers* 75: 115-136.

Coon, Jessica and Stefan Keine. 2022. Feature gluttony. *Linguistic Inquiry* 52 (4): 655-710.

Costa, João and Nancy C. Kula. 2008. Focus at the interface: evidence from Romance and Bantu. In C. De Cat and K. Demuth (eds.), *The Bantu-Romance connection*, 293-322. Amsterdam: John Benjamins.

Crane, Thera Marie, Larry M. Hyman and Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi [B865]: A Bantu language of Democratic Republic of Congo*. Vol. 147. University of California Press.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press.

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow, Essex, UK: Pearson Education.

Cruschina, Silvio. 2021. The greater the contrast, the greater the potential: On the effects of focus in syntax. *Glossa: A Journal of General Linguistics* 6 (1). 1–30.

Daeleman, Jan. 1983. Tone-groups and tone-cases in a Bantu tone-language. *International Journal of Applied Linguistics* 60(1): 131-141.

Dalrymple, Mary and Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana linguistica* 4 (1): 85-165.

Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33 (3): 409–42.

De Kind, Jasper. 2014. Pre-verbal focus in Kisikongo (H16a, Bantu). *ZAS Papers in Linguistics*, 57, 95-122.

De Kind, Jasper, Sebastian Dom, Gilles-Maurice de Schryver and Koen Bostoen. 2015. Event-Centrality and the Pragmatics-Semantics Interface in Kikongo: From Predication Focus to Progressive Aspect and Vice Versa. *Folia Linguistica Historica* 36: 113-163.

Demuth, Katherine. 1987. Discourse functions of word order in Sesotho acquisition. In R.S. Tomlin (ed.), *Coherence and Grounding in Discourse*, 98-108. Amsterdam: John Benjamins.

Demuth, Katherine and Carolyn Harford. 1999. Verb raising and subject inversion in comparative Bantu. *Journal of African Languages and Linguistics* 20 (1): 41-61.

den Dikken, Marcel. 2006. Specificational copular sentences and pseudoclefts. *The Blackwell companion to syntax*: 292-409.

Denny, J. Peter and Chet Creider. 1976. The semantics of noun classes in Proto-Bantu. *Studies in African Linguistics* 7: 1-30.

Devos, Maud and Koen Bostoen. Bantu do/say polysemy and the origins of a quotative in Shangaci. *Africana Linguistica* 18 (1): 97-132.

Dimmendaal, Gerrit. 1995. Metatony in Benue-Congo: Some further evidence for an original augment. In E. N. Emenanjo and O. Ndimele (eds.), *Issues in African languages and linguistics*, 30-38. Aba: NINLAN.

Downing, Laura J. 2019. Disentangling tone, intonation and register in selected Bantu tone languages. うそ: 13.

Downing, Laura J. and Lutz Marten. 2019. Clausal morphosyntax and information structure. In M. Van de Velde et al. (eds.), *The Bantu Languages* (Second Edition), 270-307. London: Routledge.

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2024.

Ethnologue: Languages of the World. Twenty-seventh edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74: 245–73.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34: 429–479.

Etsio, Edouard. 1999. *Parlons Teke: langage et culture*. Harmattan Press, Paris.

Fuchs, Zuzanna and Jenneke van der Wal. 2022. The locus of parametric variation in Bantu gender and nominal derivation. *Linguistic Variation* 22 (2): 268–324.

Hadermann, Pascale. 1996. Grammaticalisation de la structure infinitif + verbe conjugué dans quelques langues bantoues. *Studies in African Linguistics* 25 (2): 155–169.

Frajzyngier, Zygmunt. 1982. Indefinite agent, passive and impersonal passive: A functional study. *Lingua* 58. 267–290.

Gibson, Hannah, Andriana Koumbarou, Lutz Marten and Jenneke van der Wal. 2017. Locating the Bantu conjoint/disjoint alteration in a typology of focus marking. In J. van der Wal and L. M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*: 61–100.

Gibson, Hannah, Rozenn Guérois and Lutz Marten. 2017. Patterns and Developments in the Marking of Diminutives in Bantu. *Nordic Journal of African Studies* 26: 344–383.

Givón, Talmy. 1975. Focus and the scope of assertion: Some Bantu evidence. *Studies in African linguistics* 6 (2): 185.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In C. N. Li (ed.), *Subject and Topic*, 149–88. New York: Academic Press.

Givón, Talmy (ed.). 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study*. Amsterdam: John Benjamins.

Givón, Talmy. 1984. Direct object and dative shifting: semantic and pragmatic case. In Frans Plank (ed.), *Objects: Towards a Theory of Grammatical Relations*, 151–182. London and Orlando: Academic Press.

Givón, Talmy. 1990. *Syntax. A Functional Typological Introduction*, vol. 2. Amsterdam and Philadelphia: Benjamins.

Givón, Talmy. 2001. *Syntax*, Vol I. Philadelphia, PA: John Benjamins.

Gluckman, John. 2023. Towards a Typology of Bantu Complementizers. Presentation at *the 54th Annual Conference of African Linguistics* at the University of Connecticut. June 13.

Goes, Heidi and Koen Bostoen. 2021. Typology and Evolution of Diminutives in the Kikongo Language Cluster. *Africana Linguistica* 27: 59-102.

Grégoire, Claire. 1993. L'ordre SOV? Du Nen: Exception Ou Généralisation D'une Tendance Répandue En Bantoue Du Nord-Ouest. Paper presented at *23rd Colloquium on African Languages and Linguistics*, Leiden.

Grohmann, Kleanthes K. 2003. *Prolific Domains, on the anti-locality of movement dependencies*. Amsterdam: John Benjamins.

Grohmann, Kleanthes K. 2010. Anti-locality: Too-close relations in grammar. Ms Thesis, University of Cyprus.

Guérois, Rozenn. 2015. A Grammar of Cuwabo (Mozambique, Bantu P34). PhD dissertation, University of Lyon 2.

Güldemann Tom. 1996. *Verbmorphologie and Nebenprädikationen im Bantu. Eine Studie zur funktional motivierten Genese eines konjugationalen Subsystem*. Bochum: Universitätsverlag Dr. N. Brockmeyer

Güldemann, Tom. 2002. When 'say' is not say: the functional versatility

of the Bantu quotative marker *ti* with special reference to Shona. In: Güldemann, Tom, von Roncador, M. (eds.), *Reported Discourse. A Meeting Ground for Different Linguistic Domains*. Typological Studies in Language 52. John Benjamins, Amsterdam, 253–288.

Güldemann, Tom. 2003. Present progressive vis-à-vis predication focus in Bantu: A verbal category between semantics and pragmatics. *Studies in Language* 27 (2), 323-360.

Güldemann, Tom. 2008. Quotative Indexes in African Languages: A Synchronic and Diachronic Survey, volume 34 of *Empirical Approaches to Language Typology*. Mouton de Gruyter.

Güldemann, Tom, Ines Fiedler, Yukiko Morimoto and Kirill Prokhorov. 2010. Preposed verb doubling and predicate-centered focus. In *International Conference of the SFB*, Vol. 632: 8-10.

Güldemann, Tom, Ines Fiedler and Yukiko Morimoto. 2014. The verb in the preverbal domain across Bantu: infinitive “fronting” and predicate-centered focus. Unpublished manuscript. Humboldt University Berlin.

Güldemann, Tom, Sabine Zerbian and Malte Zimmermann. 2015. Variation in information structure with special reference to Africa. *Annu. Rev. Linguist.* 1(1), 155-178.

Güldemann, Tom and Ines Fiedler. 2022. Predicate partition for predicate-centred focus and Meeussen’s Proto-Bantu “advance verb construction”. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois and Sara Pachiariotti (eds.), *On Reconstructing Proto-Bantu Grammar*. Berlin: Language Science Press.

Gundel, Jeanette. 1988. Universals of topic-comment structure. In M. Hammond, E. A. Moravcsik and J. R. Wirth (eds.), *Studies in syntactic typology*, 209-242. Amsterdam: John Benjamins.

Guthrie, Malcolm. 1948. *The Classification of the Bantu Languages*. London: Oxford University Press.

Guthrie, Malcolm. 1960. Teke radical structure and common Bantu. *African language studies* 1: 1-15.

Guthrie, Malcom. 1971. The western Bantu languages. *Current trends in linguistics* 7: 357-366.

Hajek, John. 2006. On doubly articulated labial-velar stops and nasals in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 29 (2): 127-135.

Hajek, John. 2009. Labiodental Nasal 'ŋ' in the Angami Area. *Linguistics of the Tibeto-Burman Area* 32 (2): 107-117.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 111-176, MIT Press, Cambridge, MA.

Hamlaoui, Fatima and Emmanuel-Moselly Makasso. 2013. Bàsàa object left dislocation, topicalization and the syntax-phonology mapping of intonation phrases. Paper presented at *the 5th International Conference on Bantu Languages*, Paris, 12-15 June. ZAS MA thesis.

Hamlaoui, Fatima and Emmanuel-Moselly Makasso. 2015. Focus marking and the unavailability of inversion structures in the Bantu language Bàsàá (A43). *Lingua* 154: 35-64.

Harris, Alice C. and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Harbert, Wayne and Maher Bahloul. 2002. Postverbal subjects in Arabic and the theory of agreement. *Themes in Arabic and Hebrew syntax*. Dordrecht: Springer Netherlands. 45-70.

Henderson, Brent. 2006. *The Syntax and Typology of Bantu Relative Clauses*. PhD. dissertation, University of Illinois, Urbana-Champaign.

Henderson, Brent. 2011. Agreement, locality, and OVS in Bantu. *Lingua* 121 (5): 742-753.

Holmberg, Anders. 2015. *The syntax of yes and no*. Oxford University Press.

Hombert, Jean-Marie. 1986. The Development of Nasalized Vowels in the Teke Language Group (Bantu). In K. Bogers, H. Van der Hulst and M. Mous (eds.). *The Phonological Representation of Suprasegmentals*. Foris Publications. 359-376.

Hombert, Jean-Marie. 1987. Phonetic conditioning for the development of nasalization in Teke. *Proceedings of the 11th International Congress of Phonetic Sciences*. Vol. 2.

Hyman, Larry M. and Ernest Rugwa Byarushengo. 1984. A model of Haya tonology. In G. N. Clements and J. Goldsmith (eds.), *Autosegmental studies in Bantu tone*, 53-103. Dordrecht: Foris Publications.

Hyman, Larry M. and John R. Watters. 1984. Auxiliary focus. *Studies in African linguistics*, 15 (3), 233-274.

Hyman, Larry M. 1987. Prosodic domains in Kukuya. *Natural Language and Linguistic Theory*, 311-333.

Hyman, Larry M. 2001. Privative tone in Bantu. *Cross-linguistic studies of tonal phenomena*: 237-257.

Hyman, Larry M. 2007. Reconstructing the Proto-Bantu verbal unit: internal evidence. Bantu in Bloomsbury: Special issue on Bantu linguistics. *SOAS Working Papers in Linguistics* 15: 201-211.

Hyman, Larry M. and Maria Polinsky. 2010. Focus in Aghem. *Information structure: Theoretical, typological, and experimental perspectives*, 206-233.

Hyman, Larry M. and Florian Lionnet. 2011. Metatony in Abo (Bankon), A42. *UC Berkeley PhonLab Annual Report* 7.7.

Hyman, Larry M. 2012. Post-Verbal subject in the Nzadi relative clause. *Journal of African languages and linguistics* 33 (1): 97-117.

Hyman, Larry M. 2013. Penultimate lengthening in Bantu. *Language typology and historical contingency: In honor of Johanna Nichols*: 309-330.

Hyman, Larry M. 2015. Positional prominence vs. word accent: Is there a difference?. UC Berkeley PhonLab Annual Report 11.11.

Hyman, Larry M. 2017. Disentangling conjoint, disjoint, metatony, tone cases, augments, prosody and focus in Bantu. In J. van der Wal and L. M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*.

Hyman, Larry M. 2019. Segmental phonology. In M. O. Van de Velde, K. Bostoen, D. Nurse and G. Philippson (eds.), *The Bantu Languages* (Second Edition), 128–49. London: Routledge.

Hyman, Larry M., Florian Lionnet and Christophère Ngolele. 2019. Number and animacy in the Teke noun class system. *African linguistics across the disciplines*: 89.

Idiatov, Dmitry and Mark Van de Velde. 2016. Stem-initial accent and C-emphasis prosody in North-Western Bantu. *6th International Conference on Bantu Languages*.

Iorio, David. 2014. Subject and object marking in Bembe. PhD dissertation, Newcastle University.

Jendraschek, Gerd. 2009. Origin and development of the Iatmul focus construction: Subordination, desubordination, resubordination. *Folia Linguistica* 43 (2). 345–390.

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.

Kavari, Jekura U., Lutz Marten and Jenneke Van der Wal. 2012. Tone cases

in Otjiherero: head-complement relations, linear order, and information structure. *Africana Linguistica* 18 (1): 315-353.

Kawasha, Boniface. 2007. Passivization in Lunda. *Journal of African Languages and Linguistics* 28: 37-56.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Keenan, Edward L. and Matthew S. Dryer. 2007. Passive in the World's Languages. *Clause Structure, Language Typology and Syntactic Description*, 1, 325-361.

Kerr, Elizabeth J., Allen Asiimwe, Patrick Kanampiu, Zhen Li, Ernest Nshemezimana and Jenneke van der Wal. 2023. Bantu word order between discourse and syntactic relations, *Linguistique et Langues Africaines* 9 (1). DOI: <https://doi.org/10.4000/lla.9496>.

Kinyalolo, Kasangati K. W. 1991. Syntactic dependencies and the Spec-head agreement hypothesis in Kilega. University of California, Los Angeles.

Kinyalolo, Kasangati K. W. 2003. Limiting the scope of V-movement in Kilega. (Handout of a talk given at Stony Brook, New York).

Kobayashi, Akiko. 2022. Past participle agreement in French and Italian: A two-agree analysis. *Glossa* 7: 1-41.

Koni Muluwa, Joseph and Koen Bostoen. 2014. The immediate before the verb focus position in Nsong (Bantu B85d, DR Congo): a corpus-based exploration. *ZAS Papers in Linguistics* 57: 123-135.

Kouarata, Guy, Sara Pacchiarotti and Koen Bostoen. 2023. Nouvelle contribution à l'inventaire et la description des parlers teke (bantou, B70-80). *Linguistique et langues africaines*, 9(2), 1-33

Krifka, Manfred. 2007. The semantics of questions and the focusation of answers. In Chungmin Lee, Matthew Gordon and Daniel Büring (eds.),

Topic and focus, crosslinguistic perspectives on meaning and intonation, 139–150. Dordrecht: Springer

Kristensen, Ole-Bjørn and Kristensen, Anne-Lise. 1986. Esquisse préliminaire de la phonologie de l'EBOO, parler teke des Plateaux. SIL-Congo. (Unpublished document).

Kula, Nancy C. and Lutz Marten. 2010. Argument structure and agency in Bemba passives. In Karsten Legère and Christina Thornell (eds.), *Bantu languages: Analyses, description and theory*, 115-130. Cologne: Rüdiger Köppe.

Kuroda, Sige-Yuki. 1972. The categorical and the thetic judgment: Evidence from Japanese syntax. *Foundations of language*, 153-185.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, England: Blackwell.

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

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press [Cambridge Studies in Linguistics].

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8: 339–359

Li, Charles N. and Sandra A. Thompson. 1976. In C. Li (ed.), *Subject and topic: a new typology of language*, 457-490. New York: Academic Press.

Li, Zhen. 2020. A note on the functional passives in Teke-Kukuya (Bantu B77, Congo). In Van der Wal, Jenneke, Heleen Smits, Sara Petrollino, Victoria Nyst, and Maarten Kossmann (eds.), *Essays in African languages and linguistics: in honour of Maarten Mous*, 267-286. Leiden: ASCL Occasional Publication 41.

Linton, Pauline. 2016. *Le système verbale de la langue latege* (Deuxième révision). Léconi: Global Outreach Mission.

Linton, Pauline. 2018. *Grammaire de la langue latege pour les locateurs natifs*. Léconi: Global Outreach Mission.

Lionnet, Florian. 2017. Stem-initial prominence in West and Central Africa: Niger-Congo, areal, or both. 48th Annual Conference on African Linguistics.

Makouta-Mbougou, Jean-Pierre. 1976. Études descriptive du fumu, dialecte teke de Nagamaba, Brazzaville. PhD thesis, Paris III, Sorbonne Nouvelle, Paris.

Marlo, Michael R. and David Odden. 2019. Tone. In M. O. Van de Velde, K. Bostoen, D. Nurse and G. Philippson (eds.), *The Bantu Languages* (Second Edition). London: Routledge. 150-171.

Marten, Lutz and Jenneke van der Wal. 2014. A typology of Bantu inversion constructions. *Linguistic Variation* 14 (2): 318–68.

May, Robert. 1979. Must comp-to-comp movement be stipulated? *Linguistic Inquiry* 10: 719–725.

Meeussen, Achille Emile. 1967. Bantu grammatical reconstructions. *Annales du Musée Royal de l'Afrique Centrale* 61.

Meinhof, Carl. 1906. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Reimer.

Meinhof, Carl and Nicolaus J. van Warmelo (eds.). 1932. *Introduction to the Phonology of the Bantu Languages*, being the English version of Grundriss einer Lautlehre der Bantusprachen by Carl Meinhof, translated, revised, and enlarged in collaboration with the Author and Dr. Alice Werner by N. J. van Warmelo. Dietrich Reimer/Ernst Vohsen, published under the auspices of The International Institute of African Languages and Cultures, The Carnegie Corporation of New York, and The Witwatersrand Council of Education, Johannesburg.

Miyagawa, Shigeru. 2009. *Why agree? Why move?: Unifying agreement-*

based and discourse-configurational languages. MIT Press.

Möhlig, Wilhelm J. G. and Jekura U. Kavari. 2008. Reference Grammar of Herero (Otjiherero). *Southern African Languages and Cultures* 3, Köln: Rüdiger Köppe Verlag.

Morimoto, Yukiko. 2000. Discourse conigurationality in Bantu Morphosyntax. PhD dissertation, Stanford: Stanford University.

Morimoto, Yukiko. 2016. Verb doubling vs. the conjoint/disjoint alternation. (Paper presented at Bantu 6, Helsinki).

Mouandza, Jean-Daniel. 2001. Elements de description du Iyaa (Parler bantu du Congo-Brazzaville). Université de Nice Sophia Antipolis.

Muriungi, Peter. 2008. Phrasal movement inside Bantu verbs: Deriving affix scope and order in Kîtharaka. PhD dissertation, University of Tromsø.

Mwansa, Joseph M. 2011. The acquisition of object markers by Bemba-speaking children. PhD dissertation, University of Leeds.

Myers, Scott. 1990. *Tone and the Structure of Words in Shona*. New York: Garland.

Ndayiragije, Juvénal. 1999. Checking economy. *Linguistic Inquiry* 30 (3): 399-444.

Ndonga Mfuwa, Manuel. 1995. Systématique grammaticale du kisikongo (Angola). Paris, Université René Descartes, Paris V.

Nikolaeva, Irina. 2001. Secondary Topic as a Relation in Information Structure. *Linguistics* 39 (1): 1–49.

Ngonyani, Deo. 2000. The structure of Kindendeule applicatives. In V. Carstens and F. Parkinson (eds.), *Advances in African Linguistics*, 61–76. Trenton, NJ: Africa World Press.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Nurse, Derek. 2006. Focus in Bantu: verbal morphology and function. *ZAS Papers in Linguistics* 43: 189-207.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford University Press, USA.

Pacchiarotti, Sara, Natalia Chousou-Polydouri and Koen Bostoen. 2019. Untangling the West-Coastal Bantu mess: identification, geography and phylogeny of the Bantu B50-80 languages. *Africana Linguistica* 25: 155–229

Pacchiarotti, Sara, Guy Kouarata and Koen Bostoen. forthcoming. Sound change versus lexical change for subgrouping: Word-final lenition of Proto-Bantu *ŋg in West-Coastal Bantu. submitted to *Language Dynamics and Change*.

Paulian, Christiane. 1975. *Le kukuya: langue teke du Congo*. Paris: SELAF.

Paulian, Christiane. 1994. Nasales et nasalisation en engungwel [Nasals and nasalisation in Engungwel]. *Linguistique Africaine* 13. CNRS-LACITO Paris. 83-129.

Paulian Christiane. 1997. L'expression de la qualité en küküa (langue bantu du Congo). *Linguistique Africaine*, Hors-Série : 197-234.

Paulian Christiane. 1998. La dérivation verbale dans une langue bantu atypique : le cas du küküa. In: *Faits de langues*, n°11-12. Les langues d'Afrique subsaharienne, sous la direction de Suzy Platiel et Raphaël Kabore, 377-390.

Paulian Christiane. 2001. Coordination et Subordination, Presentation handout at *Journée sur les langues bantu*, Lyon.

Polinsky, Maria and Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19 (3): 583-646.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the

structure of IP. *Linguistic Inquiry* 20 (3): 365-424.

Preminger, Omer. 2009. Failure to agree is not a failure: agreement with post-verbal subjects in Hebrew. *Linguistic variation yearbook* 9 (1): 241-278.

Preminger, Omer. 2011. Asymmetries between person and number in syntax: A commentary on Baker's SCOPA. *Natural Language and Linguistic Theory* 29: 917–937.

Preminger, Omer. 2014. *Agreement and its failures*. Vol. 68. MIT press, 2014.

Preminger, Omer. 2019. What the PCC tells us about “abstract” agreement, head movement, and locality. *Glossa* 4 (1): 13.

Raharimanantsoa, Ruth. 2012. Aspects of phonology in Eboo-Nzikou (Bantu B74). Gothenburg: University of Gothenburg, preparatory research course paper.

Raharimanantsoa, Ruth. 2017. The prosody of tense marking in Teke-Eboo: A Bantu B70 language of Congo-Brazzaville. MA thesis, Gothenburg University, Gothenburg.

Reinhart, Tanya. 2006. *Interface strategies*. Cambridge, Massachusetts: MIT Press.

Richards, Marc. 2008. Defective Agree, case alternations, and the prominence of person. In M. Richards and A. L. Malchukov (eds.), *Scales*, 137–61. Leipzig: University of Leipzig.

Richards, Marc. 2015. Defective agree, case alternations, and the prominence of Person. In I. Bornkessel-Schlesewsky, A. L. Malchukov, and M. D. Richards (eds.), *Scales and Hierarchies*, 173–96. Berlin: De Gruyter.

Rizzi, Luigi. 1986. On the status of subject clitics in Romance. In O. Jaeggli and C. Silva-Corvalan (eds.), *Studies in Romance linguistics*, 391-420. Dordrecht: Foris.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar: A handbook of generative syntax*, 281-337. Dordrecht: Kluwer.

Roberts, Ian G. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Vol. 59. MIT Press.

Ross, John Robert. 1967. Constraints on Variables in Syntax. PhD. thesis, Massachusetts Institute of Technology.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1 (1): 75–116

Sabel, Joachim and Jochen Zeller. 2006. Wh-question formation in Nguni. In J. Mugane, J. Hutchison and D. Worman (eds.), *African Languages and Linguistics in Broad Perspective (Selected Proceedings of the 35th Annual Conference of African Linguistics, Harvard, Cambridge)*. Somerville: Cascadilla Proceedings Press, 271-283.

Sabimana, Firmard. 1986. The relational structure of the Kirundi verb (Morphosyntax, Bantu languages, GPSG, Burundi). Indiana University.

Safir, Ken. 2019. The A/ \bar{A} distinction as an epiphenomenon. *Linguistic Inquiry* 50 (2): 285-336.

Sasse, Hans-Jürgen. 1987. Thethetic/categorical distinction revisited. In: *Linguistics* 25: 511–80.

Sasse, Hans-Jürgen. 1996. *Theticity*. Arbeitspapiere Cologne: Institut für Sprachwissenschaft der Universität zu Köln.

Schadeberg, Thilo C. 1986. Tone cases in Umbundu. *Africana linguistica* 10 (1): 423-447.

Schadeberg, Thilo C. 1995. Object diagnostics in Bantu. In E. N. Emenjo and O. Ndimele (eds.), *Issues in African languages and linguistics*,

173-180. Aba: National Institute for Nigerian Languages.

Schurr, Hagay, Jason Kandybowicz, Abdoulaye Laziz Nchare, Tysean Bucknor, Xiaomeng Ma, Magdalena Markowska and Armando Tapia. 2023. Absence of Clausal Islands in Shupamem. *Languages* 9(1), 7.

Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7 (10): 534–59.

Selkirk, Elizabeth. 1995. The Prosodic Structure of Function Words. In J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.), *Papers in Optimality Theory*, 439-469. Amherst, Massachusetts: GLSA.

Şener, Sekan. 2010. (Non-)peripheral matters in Turkish syntax. Doctoral dissertation, University of Connecticut, Storrs

Shi, Dingxu. 2000. Topic and topic-comment constructions in Mandarin Chinese. *Language*, 76 (2): 383-408.

Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25: 701–721

Takizala, Alexis. 1972. Focus and relativization: the case of Kihung'an. *Studies in African Linguistics* 3 (2): 259-287.

Takizala, Alexis. 1974. *Studies in the grammar of Kihungan*. University of California, San Diego.

Tsao, Fengfu. 1987. A topic-comment approach to the *ba* construction. *Journal of Chinese Linguistics* 15: 1-54.

Tsoue, Pamela Carmelle. 2017. Etude des marqueurs verbaux du Lètèyè [langue bantu parlée au Gabon (B71a)]. PhD Dissertation. Tours.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26: 79–123.

Vallduví, Enric. 1992. *The Informational Component*. New York: Garland

Van de Velde, Mark. 2018. Why relative verbs agree with their head noun in many Bantu languages. Presentation at *Syntax of the World's Languages 8*, Paris.

Van de Velde, Mark. 2019. Nominal morphology and syntax. In M. O. Van de Velde, K. Bostoen, D. Nurse and G. Philippson (eds.), *The Bantu Languages* (Second Edition), 237-269. London: Routledge.

Van de Velde, Mark. 2021. The Bantu relative agreement cycle. *Linguistics* 59 (4): 981-1015.

Van der Wal, Jenneke. 2009. *Word order and Information Structure in Makhuwa-Enahara*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT).

Van der Wal, Jenneke. 2011. Focus excluding alternatives: Conjoint/disjoint marking in Makhuwa. *Lingua*, 121 (11): 1734-1750.

Van der Wal, Jenneke. 2015. *Bantu Syntax*. Oxford Handbooks Online

Van der Wal, Jenneke. 2015. Bantu object clitics as defective goals. *Revue Roumaine de Linguistique* LX (2–3): 277–96.

Van der Wal, Jenneke and Jacky Maniacky. 2015. How 'person' got into focus: Grammaticalization of clefts in Lingala and Kikongo areas. *Linguistics*, 53 (1): 1-52.

Van der Wal, Jenneke. 2016. Diagnosing focus, *Studies in Language* 40 (2): 259–301

Van der Wal, Jenneke and Saudah Namyalo. 2016. The interaction of two focus marking strategies in Luganda. *Diversity in African languages*: 355-377.

Van der Wal, Jenneke. 2021. *The BaSIS basics of information structure*.

(<https://bantusyntaxinformationstructure.files.wordpress.com/2021/08/the-basis-basics-of-information-structure-complete-6.pdf>)

Van der Wal, Jenneke. 2021. On the use of idioms for testing focus. *Semantic Fieldwork Methods* 3 (2).

Van der Wal, Jenneke. 2022. *A Featural Typology of Bantu Agreement*. Oxford: Oxford University Press.

Vasina, Jan. 1966. *Kingdoms of the Savanna*. Madison: University of Wisconsin Press.

von Fintel, Kai. 2008. What is presupposition accommodation, again? *Philosophical perspectives* 22 (1): 137–170.

Wasike, Aggrey. 2007. The left periphery, wh-in-situ and A-bar movement in Lubukusu and other Bantu languages. PhD dissertation, Cornell University, Ithaca.

Watters, John R. 1979. Focus in Aghem: A study of its formal correlates and typology. Doctoral dissertation, University of Los Angeles.

Welmers, William E. 1973. *African language structures*. Berkeley: University of California Press.

Williams, Edwin. 2003. *Representation theory*. Cambridge: MIT Press.

Yoneda, Nobuko. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua*, 121 (5): 754–771.

Zeller, Jochen. 2008. On agreement, focus and subject positions in Bantu. Paper presented at the conference *Movement and word order in Bantu*, Leiden.

Zeller, Jochen. 2013. Locative inversion in Bantu and predication. *Linguistics* 51 (6): 1107–46.

Zeller, Jochen. 2013. In defence of head movement: Evidence from Bantu. *Diagnosing syntax*: 87-111.

Zerbian, Sabine. 2006. Expression of information structure in Northern Sotho. PhD dissertation, Humboldt University.

English summary

The first part of the thesis provides a grammar sketch of this language as a reference source for general readers and for Bantuists as well as typologists. The second part presents and discusses the interactions between word order variation and information structure, with particular interest in the IBV focus position, its historical origin and structural representation and the class 1 SM alternation that is associated with the use of the focus position.

The grammar sketch in chapter 2 first introduced the phonology of the language, including segmental phonology and the prosodic system. Compared with many other Teke varieties, Kukuya shows less phonological reduction, such as the preservation of CV- shape nominal prefixes and optional vowel nasalisation. The language employs five fixed tone patterns which are mapped onto certain prosodic domains. Interestingly, the postlexical prosodic domain in Kukuya consists of a stem with the following prefix. The stem-initial prominence is also discussed, which can be reflected in the facts that only the stem-initial consonant has the full inventory and the tone on the stem-initial syllable is always stable and not subject to tone change rules. The description on the nominal morphology shows that Kukuya is typical of a Bantu language in that it makes use of the noun class system including locative classes, and I also introduced the nominal derivation rules such as the formation of diminutives. I also presented the functions of the adnominal modifiers in the noun phrase

and their concord patterns with the head noun. The verb shows less agglutinative morphology than in a prototypical Bantu language, since object marking and verbal extensions are completely lost, which is a tendency attested more widely in northwestern Bantu. Both segmental and tonal strategies are attested in expressing TAM distinctions. In the introduction of the clause structure, I presented the agreeing *say*-complementisers, the structure of relative clauses, as well as different kinds of adverbial clauses.

Chapter 3 contributes to a description of the expression of information structure in Kukuya. The chapter is mainly concerned with different kinds of word order variation in the language that are influenced by information structure. I first discussed the canonical word order of the language and what discourse functions it can have. Then I provided an overview on the functions and interpretation of the dedicated IBV focus position, showing that the IBV position is available for arguments including subject and object, as well as for adjuncts to get focused. The IBV focused element often has an identificational/contrastive focus reading. I also introduced different kinds of topical elements in the preverbal domain, proposing the distinction between primary and secondary topics. I presented two specific constructions that can function as the equivalent of a passive, namely the OSV and the impersonal *ba*-constructions. The language also makes use of cleft constructions to express focus. I presented the structure and functions of the basic cleft, (reverse) pseudo-cleft and the so-called reduced cleft. In general, the Kukuya language behaves as a discourse-configurational language.

Chapter 4 continued the discussion on the IBV focus strategy, and concentrated on the investigation of its historical origin. I hypothesised that the IBV focus construction originates from a cleft, and I corroborated this hypothesis by presenting segmental and tonal evidence that connects these two constructions. I showed that the two constructions share many grammatical properties such as the H tone prefix on the focused element, the class 1 SM alternation and the H tone insertion on the SM, as well as the verb-final H tone in the context of non-subject focus. Based on these shared features, I proposed a grammaticalisation pathway from a basic cleft to the IBV focus construction, arguing that the latter is developing from a biclausal cleft towards a monoclausal focus construction. At the

end of the chapter I also made reference to some other West-Coastal Bantu languages in which the IBV focus position is also attested, displaying some micro-variation with regard to this focus strategy.

The syntactic analysis of the IBV focus construction and the associated class 1 SM allomorphy is given in chapter 5. I first discussed the structural representation of the canonical SVO order, deciding the structural position of the verb and the grammatical subject. Then I discussed the structural position of the IBV focused element. I showed that the IBV element undergoes A'-movement, and there is one unique IBV position that is available for both focal subject and focal non-subject. I compared the low FocP and the high FocP approaches to accommodate the IBV focus element, and proposed that the IBV element is located in the specifier of a high FocP which is in the extended TP domain above TP. The rest of the chapter investigated the syntax of the class 1 SM alternation. Since this allomorphy is a shared grammatical feature with the cleft/relative constructions and the subject agreement patterns in relatives are also interesting in Kikuyu syntax, I first presented the subject agreement morphology in subject and non-subject relatives. Based on the Defective Goal approach and the distinction between Full and Minimal Agree, I proposed the presence of an extra [Person]-layer on certain pronouns and provided an analysis of the agreement asymmetries between preverbal and postverbal subjects. I also generalised the lexical insertion rules for the spell-out of SMs. At the end of the chapter, I probed the structure of the reduced cleft and the IBV focus construction, investigating the class 1 SM alternation in both constructions. I proposed that the initial topical element in the IBV focus construction is always base-generated and under non-subject focus it can control a clause-internal ϕ P that function as the true argument of the verb. The SM *ka-* is spelled out as a result of the agreement between T and the ϕ P that equals to the 3rd [Person] feature.

Samenvatting in het Nederlands

Dit proefschrift beschrijft de grammatica van het Teke-Kukuya en analyseert woordvolgorde en werkwoordscongruentie met het oog op de uitdrukking van informatiestructuur.

Het eerste deel van het proefschrift geeft een grammaticale schets van het Teke-Kukuya en kan als referentiebron gebruikt worden door Bantuïsten, typologen, en verdere geïnteresseerden. Het tweede deel presenteert en bespreekt de interacties tussen woordvolgordevariatie en informatiestructuur, met bijzondere aandacht voor de focuspositie direct voor het werkwoord ('immediate before verb', IBV), haar historische oorsprong en structurele representatie en de alternantie in de onderwerksmarkering van klasse 1 die geassocieerd wordt met het gebruik van de focuspositie.

De grammaticaschets in hoofdstuk 2 introduceert eerst de fonologie van de taal, inclusief de segmentele fonologie en het prosodisch systeem. Vergeleken met veel andere Teke-variëteiten vertoont Kukuya minder fonologische reductie, wat blijkt uit het behoud van CV-vormige nominale voorvoegsels en optionele klinkernasalisatie. De taal heeft vijf vaste toonpatronen die zijn gekoppeld aan bepaalde prosodische domeinen. Interessant is dat het postlexicale prosodische domein in Kukuya bestaat uit een stam met het daaropvolgende voorvoegsel van het volgende woord. De staminitiële prominentie wordt ook besproken; dit is zichtbaar in het feit

dat alleen de staminitiële medeklinker de volledige inventaris van klanken heeft en het feit dat de toon op de staminitiële syllabe altijd stabiel is en niet onderhevig aan toonveranderingsregels. De beschrijving van de nominale morfologie laat zien dat Kukuya een typische Bantoetaal is in de zin dat het gebruik maakt van het naamwoordklassensysteem inclusief locatiev klassen. Ook de nominale derivatie wordt besproken, zoals de vorming van verkleinwoorden. Daarnaast worden de functies van de adnominale modificeerders in de naamwoordgroep en hun concordantiepatronen gepresenteerd. Het werkwoord vertoont minder agglutinatieve morfologie dan in een prototypische Bantoetaal, aangezien objectmarkering en werkwoordderivaties volledig verloren zijn gegaan, wat een tendens is die meer voorkomt in het noordwestelijke Bantoegebied. Zowel segmentele als tonale strategieën komen voor bij het uitdrukken van onderscheid in tijd, aspect, en grammaticale wijs ('tense-aspect-mood', TAM). In het gedeelte over de zinsbouw zijn de congruerende voegwoorden van 'zeggen' interessant, alsook de structuur van betrekkelijke bijzinnen en verschillende soorten bijwoordelijke zinnen.

Hoofdstuk 3 beschrijft de uitdrukking van informatiestructuur in Kukuya. Het hoofdstuk gaat voornamelijk over verschillende soorten woordvolgordevariatie in de taal die beïnvloed worden door informatiestructuur, dat wil zeggen hoe de boodschap verpakt wordt zodat de luisteraar zo goed mogelijk kan horen wat oude/nieuwe en contrasterende informatie is. Eerst bespreek ik de canonieke woordvolgorde van de taal en welke informatiestructuur daarmee uitgedrukt kan worden. Daarna volgt een overzicht van de functies en interpretatie van de specifieke IBV-focuspositie, waaruit blijkt dat de IBV-positie beschikbaar is voor argumenten inclusief onderwerp en lijdend voorwerp, en ook voor adjuncten om gefocuseerd te worden. Ik laat daarbij zien dat het IBV-focus-element vaak een identificatieve/contrastieve focuslezing heeft. Ik heb ook laten zien dat verschillende soorten topics in het preverbaal domein kunnen voorkomen en dat een onderscheid gemaakt kan worden tussen primaire en secundaire topics. Verder zijn er twee constructies die kunnen functioneren als het equivalent van een passieve zin, namelijk OSV woordvolgorde en de onpersoonlijke ba-constructie. De taal maakt ook gebruik van cleftconstructies om focus uit te drukken. We kunnen drie soorten clefts onderscheiden: de basiscleft ('het is X die ...'), de (omgekeerde) pseudocleft ('degene die ... is X') en de zogen-

aamde gereduceerde cleft, waarin bepaalde kenmerken ontbreken. In het algemeen kunnen we zeggen dat de Kukuya-taal zich als een zogenaamde discourse-configurational taal gedraagt.

Hoofdstuk 4 vervolgt de discussie over de IBV-focusstrategie en concentreert zich op het onderzoek naar de historische oorsprong ervan. Ik onderzoek de hypothese dat de IBV-focusconstructie voortkomt uit een cleft, en ik bevestig deze hypothese door segmenteel en tonaal bewijs te presenteren dat deze twee constructies met elkaar verbindt. De twee constructies, IBV focus en cleft, hebben veel grammaticale eigenschappen gemeen, zoals de hoge toon op het prefix van het gefocuste element, de allomorfie van de klasse 1 onderwerpsmarkeerder en het toevoegen van een hoge toon op de subjectsmarkeerder, evenals de werkwoordsfinale hoge toon in de context van niet-subject focus. Gebaseerd op deze gedeelde kenmerken stel ik een grammaticalisatiepad voor van een basiscleft naar de IBV-focusconstructie, met als argument dat deze laatste zich ontwikkelt van een tweeledige cleft naar een monoclausale focusconstructie. Aan het eind van het hoofdstuk verwijs ik naar enkele andere Westkust Bantoetalen waarin de IBV-focuspositie ook voorkomt en die enige microvariatie vertonen met betrekking tot deze focusstrategie.

De syntactische analyse van de IBV-focusconstructie en de bijbehorende allomorfie van de klasse 1 onderwerpsmarkeerder wordt gegeven in hoofdstuk 5. Eerst wordt de structurele representatie van de canonieke SVO-volgorde besproken, waarbij de structurele positie van het werkwoord en het grammaticale onderwerp worden bepaald. Daarna bespreek ik de structurele positie van het IBV-element. Ik toon aan dat het IBV-element A'-verhuizing ondergaat, en er is één unieke IBV-positie die beschikbaar is voor focale constituenten, of ze als onderwerp fungeren of een andere rol hebben. Ik vergelijk de lage FocP en de hoge FocP benaderingen als landingsplaats van het IBV focus element, en concludeer dat het IBV element zich bevindt in de specifier van een hoge FocP die zich in het uitgebreide CP domein boven TP bevindt. De rest van het hoofdstuk onderzoekt de syntaxis van de allomorfie van de klasse 1 onderwerpsmarkeerder. Omdat deze allomorfie een gedeeld grammaticaal kenmerk is met de cleft/relatieve constructies en de onderwerpscongruentiepatronen in betrekkelijke bijzinnen ook interessant zijn in de Kukuya syntaxis, wordt eerst de onderwerpscongruentie in verschil-

lende betrekkelijke bijzinnen gepresenteerd. Gebaseerd op de Defective Goal benadering en het onderscheid tussen Full en Minimal Agree, stel ik de aanwezigheid van een extra [Persoon]-laag op bepaalde voornaamwoorden voor en analyseer ik de congruentie-asymmetrieën tussen preverbale en postverbale onderwerpen. Aan het einde van het hoofdstuk onderzoek ik de structuur van de gereduceerde cleft en de IBV-focusconstructie, specifiek de allomorfie van de klasse 1 onderwerpsmarkeerder in beide constructies. Ik stel voor dat het initiële topicale element in de IBV-focusconstructie altijd op de hoge positie gegenereerd wordt onder niet-onderwerpsfocus een zinsinterne ϕ P kan aansturen die functioneert als het ware argument van het werkwoord. De onderwerpsmarkeerder *ka-* is de reflectie van de overeenkomst tussen T en de ϕ P van het [Persoon] kenmerk.

Curriculum vitae

Zhen Li (李臻) was born on 2nd August, 1994, in the city of Tianjin in China. He attended Peking University from 2013 to 2017, where he obtained a bachelor degree in French language and world history. During his undergraduate studies, he spent one exchange semester at the Paris Institute of Political Studies in France. In September 2017, he went to the School of Oriental and African studies (SOAS, University of London) in the UK and obtained a master degree in African Studies (major in African languages). In October 2018, he came to the Netherlands to pursue a PhD degree in African linguistics at Leiden University. He was a member in Dr. Jenneke van der Wal's Vidi project *Bantu Syntax and Information Structure* from 2018 to 2022. During his PhD, he conducted fieldwork in Brazzaville and Lékana in the Republic of Congo, studying the grammar and information structure of the Teke-Kukuya language. Some of his findings have been presented at various conferences in Europe, Africa and North America such as ACAL52-54, Bantu8-9 and CALL51-53. This dissertation is the outcome of his PhD research. At the time of his defence, he has taken up a Boya postdoctoral researcher position in the School of Foreign Languages at Peking University.