Word order, information structure and agreement in Teke-Kukuya

Li, Z.

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]

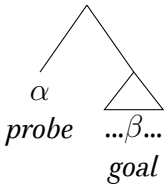
I proposed that the diachronic derivation starts from a basic cleft construction like in (1a) in which the relative clause and the predicative part of the sentence can be clearly distinguished. In (1b) the relative marker on the verb is deleted, which is accompanied by the fronting of the postverbal 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 1 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 φP with a 3rd person value is merged as the external argument in the vP, and the topical element that is co-referential with this 3rd person φP 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.

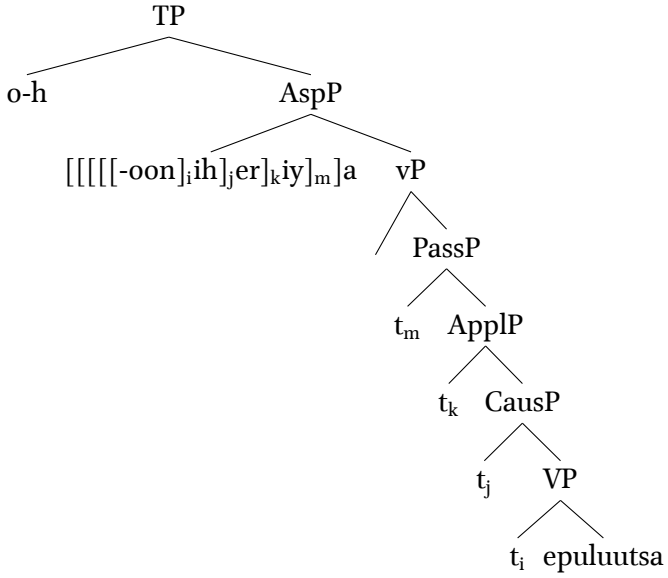
5.1.1 Position of the verb

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)

b.



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-fiemé mu ki-wá
1.child 1SM.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.

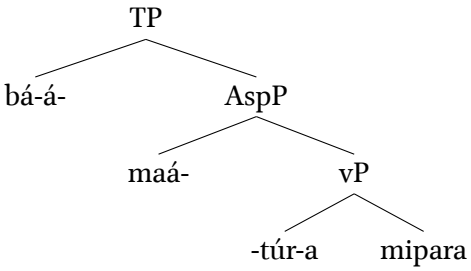
In addition to these inconsistent analyses on whether the function of FV in Teke is more towards tense or aspect marking, it should also be noted that the FV alternation is also attested in conditional and subjunctive clauses (see also chapter 2 section 2.6). One example is given in (5), in which the 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
1SG.PRO COND-1SG.SM-know-FV 1SG.COMP 1SG.PRO 14-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
bvheí́ 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.

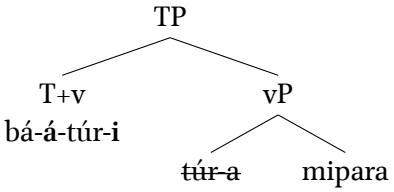
In (6) I propose that in the derivation of a basic SVO sentence, when there is an AspP that hosts a segmental aspect prefix, the verb just moves from V to v where it takes the default FV **-a**, and the perfective aspect prefix **maá**- is attached to the verb by phonological merger.

(6) a. Bó báá-maá-túr-á mí-pará.
2.PRO 2SM.PST-PFV-steal-FV 4-money
'They have stolen the money.'

b.


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.'

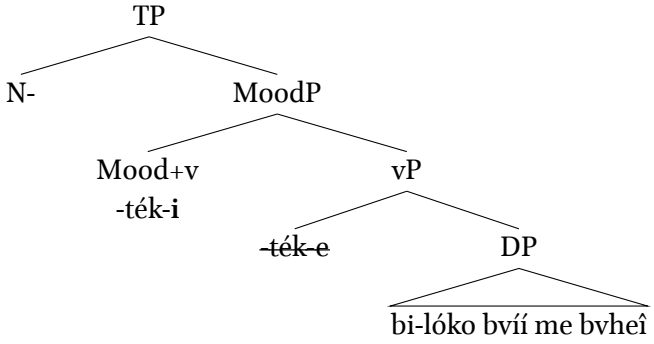
b.


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 bvheî́.
1SG.PRO 1SG.SM-sell-**SBJV** 8-thing 8.CONN 1SG.PRO 8.all
'...I would sell all my things.'

b.



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 account[1] (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

[1]For 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 Kukuya. 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(rgument)-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 Kĩîtharaka (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 Kukuya 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 maker 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
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?'

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 a thetic expression, and I have shown that a thetic 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 bhoî** báá-bák-ǐ 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-**líi**-li **líi**-fúúm-í **bhií**
4-pig 4REL-1PL.SM.RPST-COP 1PL.SM.RPST-buy-PST 1PL.PRO
'the pigs that we had bought'

b. mwáana wu-***báá/kíí**-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-wî 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-1SM.PST-kill.PST 1.leopard NEG
'Father did not kill the leopard.'

b. Ngo ka **taará** á-dwí ni.
1.leopard NEG 1.father 1SM.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).

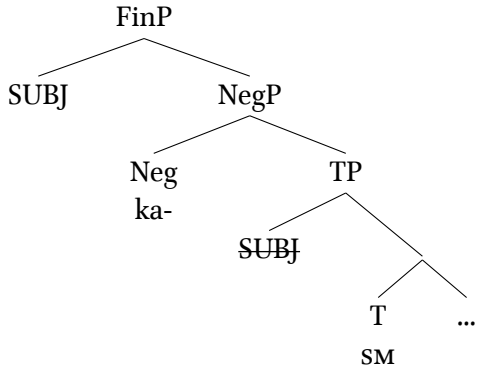
(19) a. Taará á-mún-i me.
1.father 1SM.PST-see-PST 1SG.PRO
'Father saw me.'

b. Taará ka-ká-mún-i me ni.
1.father NEG-1SM.PST-see-PST 1SG.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 and thetic interpretations.

(20)



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 computated 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 Kukuya. 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-fúum-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** kâ-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'-position[2].

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

[2]Interestingly, 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 mhoî.
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 bvíí
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 complimentisers, 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** [wǔ-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 17-which 2SG.PST-hear-PST 2.COMP
[ku-ká-kwí ndé]?
17REL-1SM.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 dzándu ká-fúum-i.
1.PRO 6-wine l7.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]$_{FOC}$ kíí-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 1SM-RPST 1SM.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.

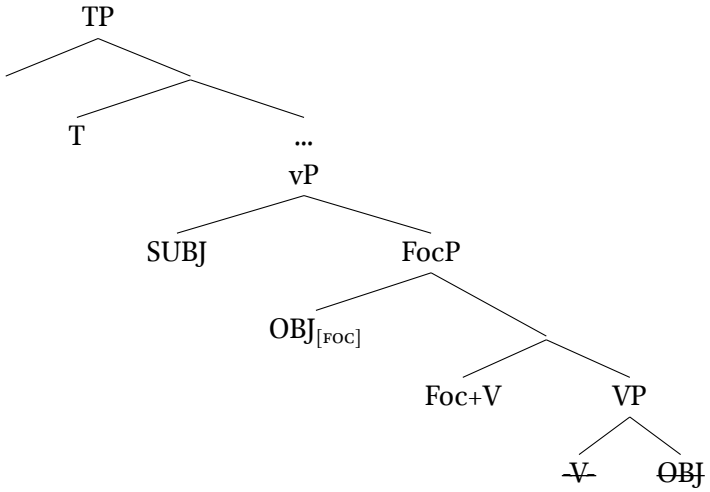
5.2.2.2 Low or high FocP?

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)

b.

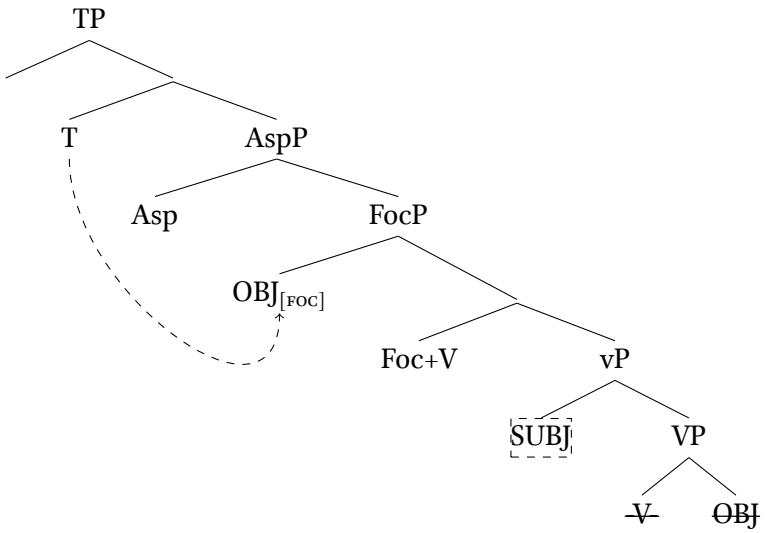


An obvious problem that this approach brings is that since the focused subject is considered to be located in a FocP, it should be attracted by some [Foc] feature from its base position in specvP, which stipulates that the Foc head must c-command the vP, otherwise the subject can never be visible to 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 Kukuya.

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)

b.



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 FV, 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 17-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-1SM.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.*)

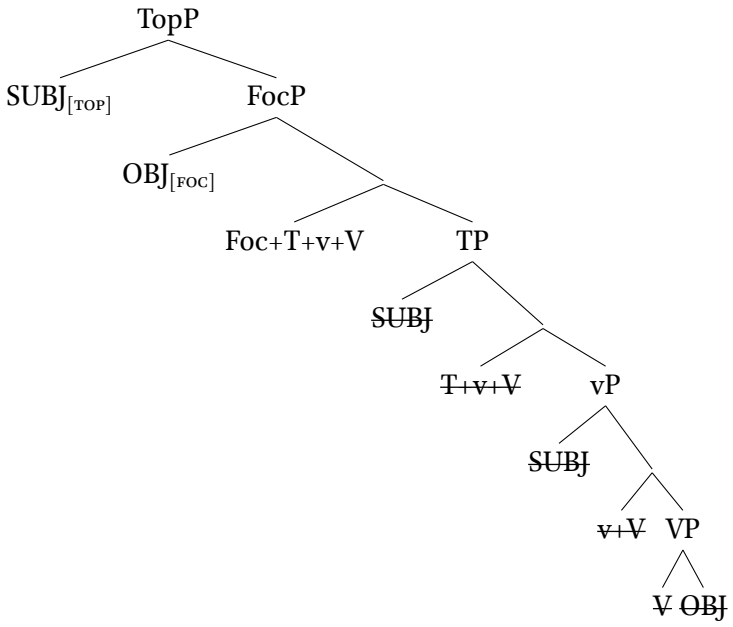
#Ki-lóko ngúku ka-ká-dzí ni.
7-thing 1.mother NEG-1SM.PST-eat.PST NEG

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.

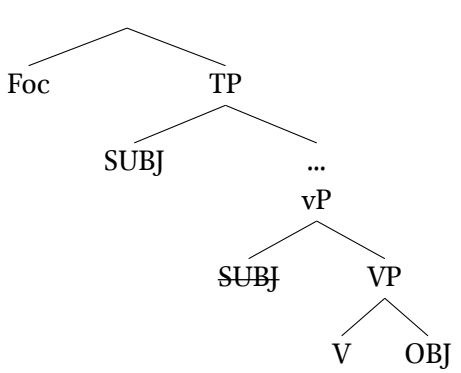
(38) a. TopP<FocP<TP<AspP<vP

b.



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.

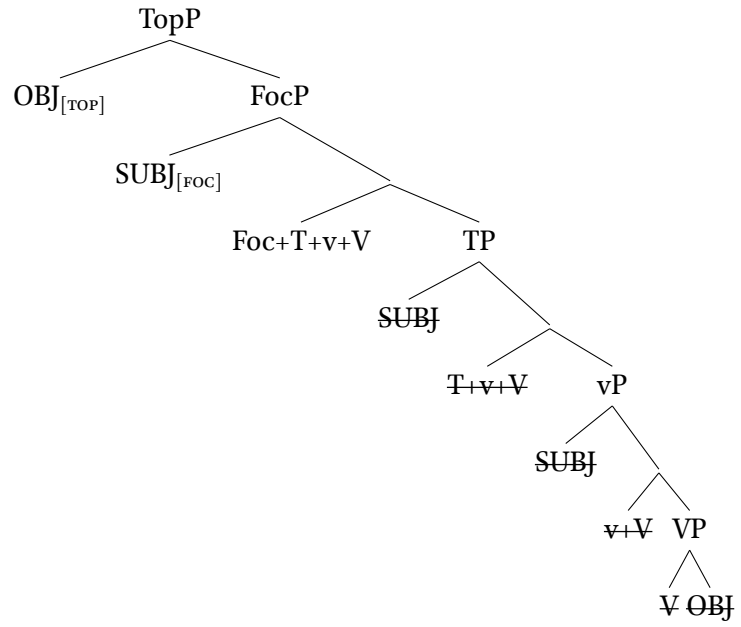
(39)



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.

(40)



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 Kukuya, 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-líí**-súruk-i] líí-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
tíí 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) $\text{HEAD}_i\ \text{REL}_i\text{-AGR}_i\text{-V}\ (...) \Longrightarrow \text{HEAD}_i\ \text{REL}_i\text{-V}\ (...)$

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).

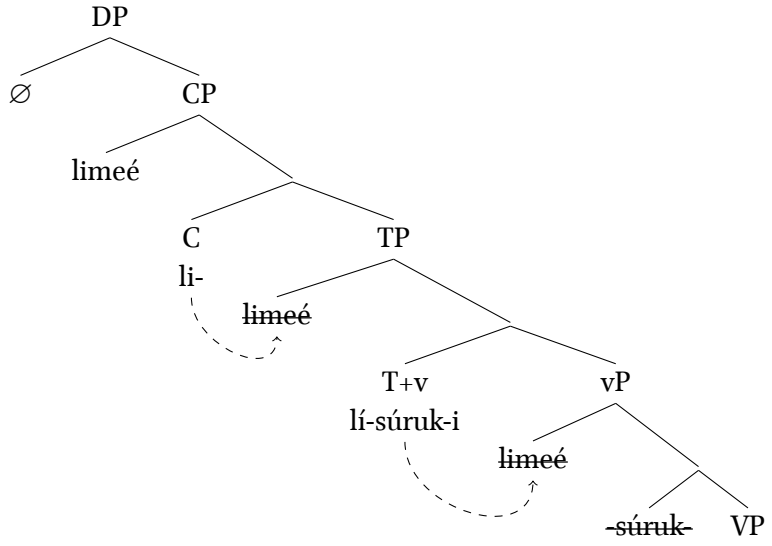
(45) Bhií **ba-lí**-kâ-sál-á máa-ntséke
1PL.PRO 2REL-1PL.SM-IMPF-work-FV 6-field
li-kâ-sílik-a bú-su.
1PL.SM-IMPF-wake-FV 14-front
'We who work in the fields wake up early.'

(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 [uφ] features and the two heads show distinct feature specification, I assume that they bear separate sets of [uφ] features (Henderson

2011; Carstens and Diercks 2013), thus I am not following the Feature Inheritance account (Chomksy 2005, 2008; Miyagawa 2010) in which the [uφ] 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 [uφ] and a [REL] feature, it probes and finds the subject in specTP, the subject values the [uφ] features on C and finally moves to specCP.

(47)



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

[3]Here 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) Yă nkúmá 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-**líi**-li **líi**-tí **bhií** bví
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** bví
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.

[4]Here 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-**kíí**/***báá**-fúúm-í **báana** báa-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-**kíí**/***míí**-dz-í ŋ**jé**
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**/**kíí**-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 lií-tsuk-í mu ku-**káá**/**kíí**-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-**báá**/**kíí**-fúúm-í **bó** míí-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 occurs 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
1SG.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. A 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-***kí**/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.

	[+**PERSON**]			[-**PERSON**]	
SM	1st	2nd	**3rd**	**class 1/2**	other classes
singular	-N-	-∅-	-**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.

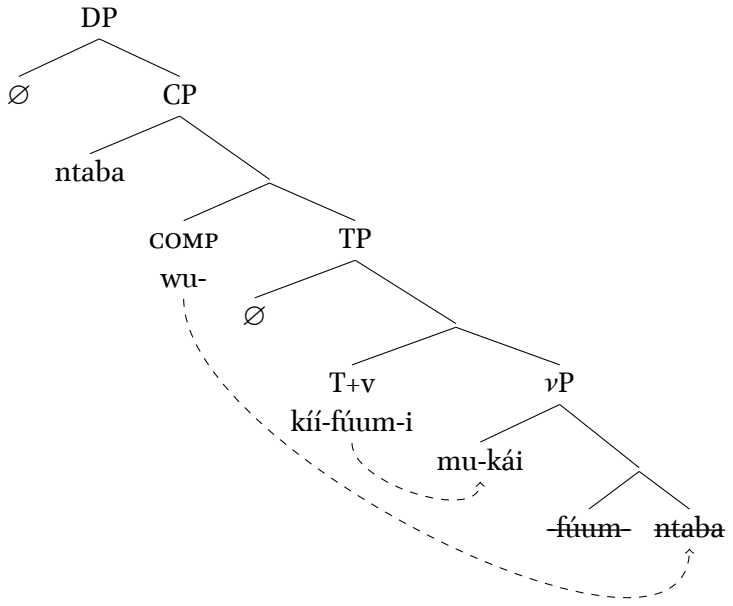
One may wonder why T cannot agree with the subject and raise it to specTP first, and then the verb just moves from T to C to satisfy the linear order requirement, since the specTP position is active in this language, and T has separate φ-features relevant for subject-verb agreement. I suppose that T 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'

b.



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. ʔal-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 ńko: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-enga.
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.

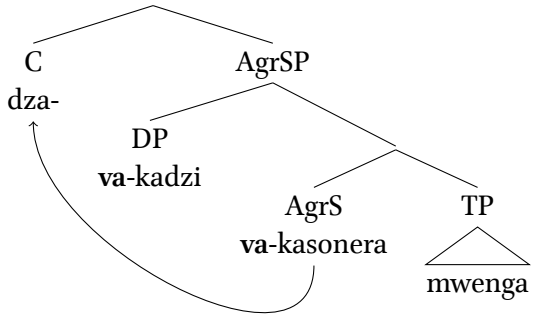
(67) a. kit ki-**a**/(*ka)-swiim-in **Kipes**
7.chair 7REL-1SM-buy-PST 1.Kipes
'the chair that Kipes bought'

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.

(68) (Zeller 2013: 96)



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 [uφ] 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.

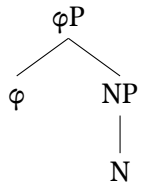
5.3.2.2 How does T agree with a postverbal subject?

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.

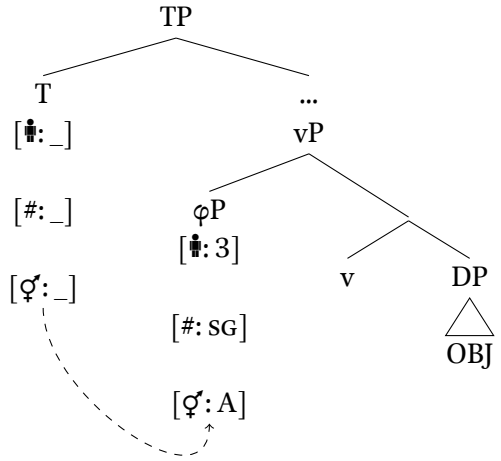
To illustrate, a pronoun can be a φP which is structurally smaller than a pronominal DP (Déchaine and Wiltschko 2002), as shown in (70), which only contains ϕ-features and no D-feature. When it agrees with a probe that contains [uφ] features and values them, the probe will contain the same valued features as the φP and possibly more than those, while the φP itself does not contain any features that are not present on the probe, thereby counting as a defective goal of the probe.

(70)



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 φP pronoun that contains the [Person], [Number] and [Gender] features which form an (improper) subset of the features on T, this subject φP counts as a defective goal of T.

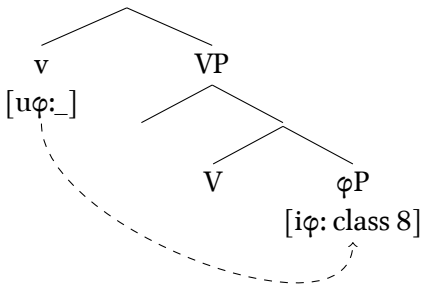
(71)



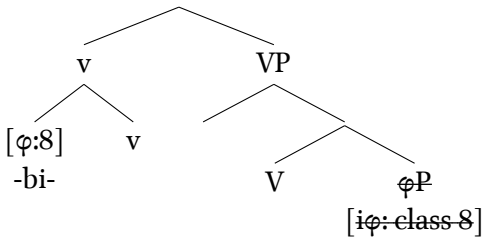
Upon agreement, there are two indistinguishable copies of the same set of features shared on the probe and the goal which form a chain, this leads them to undergo the PF-operation of chain reduction (Nunes 2004). The chain reduction takes place and the φ-features are only spelled out on the *higher* 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

[uφ] features, probes down to find the internal argument with [iφ]. If the object is a φP 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)

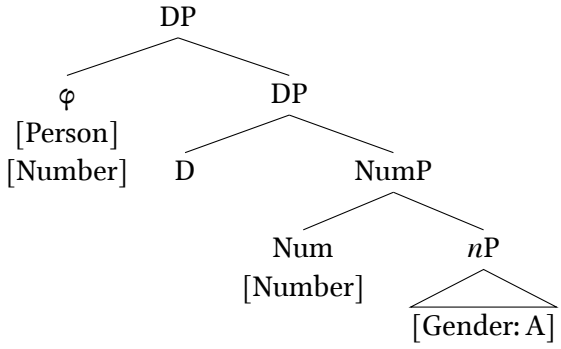
As mentioned, Iorio (2014) has applied the defective goal approach to account for subject marking in Bembe. In (73), we see that in Bembe object relative clauses the SM on the verb and the postverbal lexical subject DP are in complementary distribution. Iorio proposes that in (73a) what T agrees with is a defective subject ϕP 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 Kukuya, 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 Kukuya



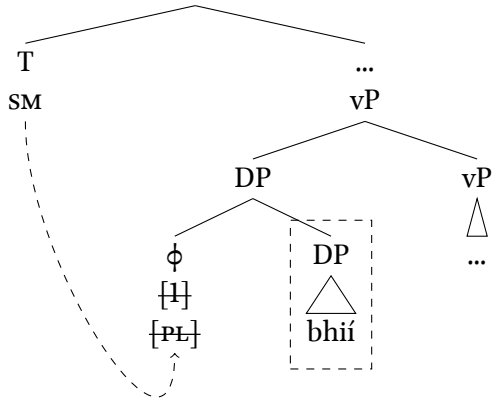
The DP with an extra φ-layer in (74) has been analysed as a "big DP" in different languages (Uriagereka 1995; Belletti 1999; Bleam 1999; Cechetto 2000) and Bax and Diercks (2012: 196) adopted this big DP hypothesis in accounting for Manyika object marking, in which the extra layer is itself

the object marker. As for the [Person] pronouns in Kukuya in (74), I suppose that the extra φ-layer is intrinsically a [Person]-layer that contains different person values, following van der Wal (2015, 2022). However, the [Number] feature can always percolate from the embedded NumP to this φ-layer, otherwise we would not expect the number distinction in the agreement morphology of person pronouns (see section 5.3.3.3 on the feature specification of SMS). Notably, as illustrated in (74), I propose that unlike many other Bantu languages for which the extra layer was proposed to contain/copy all the φ-features from the DP (van der Wal 2022: 54), in Kukuya only [Person] and [Number] features are present on this extra ϕ-layer, but [Gender] is systematically embedded in the lower part of a DP, namely in the *n*P that contains the lexical root, but does not head its own projection (Carstens 2008, Fuchs and Van der Wal 2022) or percolate to the ϕ-layer to be visible to the probe, which accounts for the systematical lack of [Gender] specification on the SM in a non-subject relative.

In example (75), when the postverbal subject DP is a [+Person] pronoun, it contains the extra layer that contains [Person] and [Number] features. T probes down and finds this extra layer as the goal, as shown in (76). Since the φ-layer is a defective goal of T that contains a full set of [uϕ], the agreed and valued [Person] and [Number] features are spelled out as the SM on T, while the ϕ-layer itself gets deleted at PF, and the rest of the DP (in the dashed square) is spelled out in its base position, resulting in the SM doubling. Under this analysis, the agreement with an *in situ* [Person] pronoun, which I previously assumed to show full agreement, is actually not specified for [Gender], but [Person] and [Number] features are specified enough for the SM to be spelled out, see section 5.3.2.3 on the lexical insertion rules.

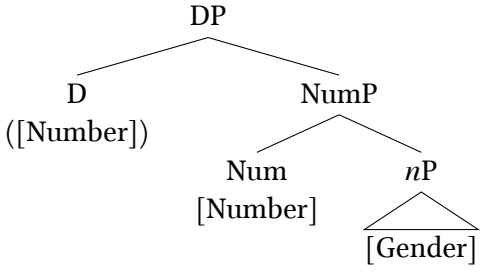
(75) ma-meé ma-**líí**-tí **bhií** bví
6-stone 6REL-1PL.PST-launch.PST 1PL.PRO 9.falling
'the stones that we threw away'

(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).

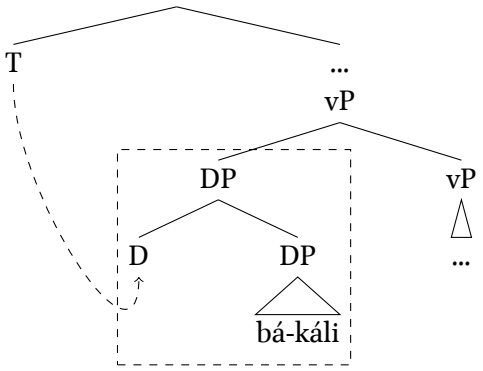
(77) Structure of lexical DPs and class pronouns in Kukuya



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-1SM-build.PST 2-father 2.CONN 1PL.PRO
'the houses that our fathers built'

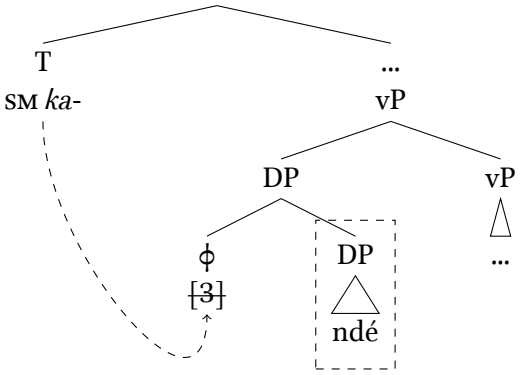
b. **U**-faan-i we sal-i.
1SM-should-PST 2SG.PRO work-SBJV
'You have to work.' [Teke-Eboo B74, Raharimanantsoa 2012]

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***-.

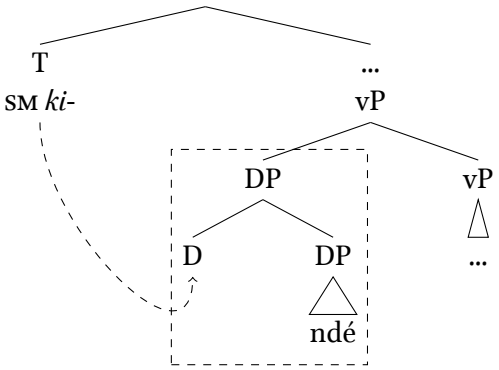
As for the proposed [+/-Person] differentiation between the two sets of pronouns which have the forms *ndé* and *bó*, the choice of SMs can be similarly explained as follows. In example (82), when the postverbal subject *ndé* is equipped with the extra ϕ-layer that contains a 3rd [Person] feature and 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 lií-tsuk-í mu ku-**káá**/**kíí**-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 maker 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 ϕP, 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).

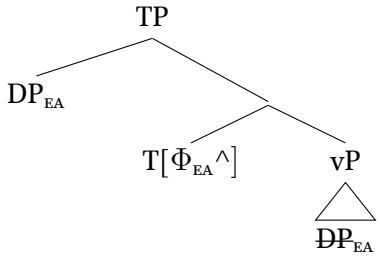
(84) **Ngúku** áá-télek-i [ntsúú wu-**káá**/**kíí**-ték-i]?
1.mother 1SM.PST-prepare-PST 1.chicken 1REL-1/7SM.PST-sell-PST
'Did mother prepare the chicken to sell?'

(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?

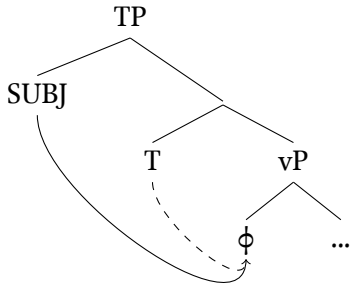
Here I adopt the defective goal approach again as an alternative to the EPP-associated agreement, I follow van der Wal (2022: 204) and Perry (*to appear*)'s suggestion in proposing that the *movement* of the goal over the probe triggers spell-out of the features on the probe as well. That is to say, 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 ϕP 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.

(86)



(van der Wal 2022: 205)

(87) *Agreement with a preverbal subject DP*



After illustrating how T agrees with different types of *in situ* subject DPs, the remaining question for understanding the differential subject marking strategies for preverbal and postverbal subjects is how we account for the different spell-outs of the different featural specifications, especially the class 1 *a-*/*ka-* SM alternations.

5.3.2.3 Lexical insertion rules

Now I discuss how the valuation of ϕ-features on T is is spelled out in the morphology, and I assume that there are some lexical insertion rules that can account for the appearance of T features as SMs in different derivations. Note that here I only discuss rules that work for the spell-outs of SMs but not for C-agreement.

Some lexical insertion rules that are relevant to the present discussion

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**-/$\Longleftrightarrow$[1]
/**li**-/$\Longleftrightarrow$[1], [pl]
/**Ø**-/$\Longleftrightarrow$[2]
/**ma**-/$\Longleftrightarrow$[C], [pl]
/**ki**-/$\Longleftrightarrow$[D]
/**bvi**-/$\Longleftrightarrow$[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**-/$\Longleftrightarrow$[3]
/**a**-/$\Longleftrightarrow$[A], ([3])
/**ba**-/$\Longleftrightarrow$[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 ϕP 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 ϕP 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**]$_{\text{FOC}}$ [mu-kái]$_{\text{TOP}}$ **káá**-fúum-i.
4-pig 1-woman 1SM.PST-buy-PST
'(It was) some PIGS that the woman bought.' [reduced cleft]

b. [Mu-kái]$_{\text{TOP}}$ [**mí-fémé**]$_{\text{FOC}}$ **káá**-fúum-i.
1-woman 4-pig 1SM.PST-buy-PST
'The woman bought the PIGS' [IBV object focus]

c. [Mi-féme]$_{\text{TOP}}$ [**mú-káí**]$_{\text{FOC}}$ **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

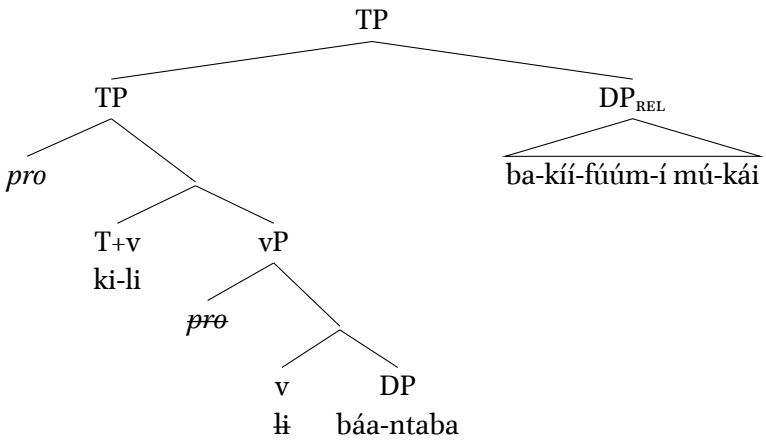
I start from investigating the syntactic derivation of clefts. A construction that I have labeled as a basic cleft is repeated in (91), which looks like a lot an *it*-cleft construction in English as shown in the translation. However, if we look at the construction carefully, we find that it differs 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.'

(93)



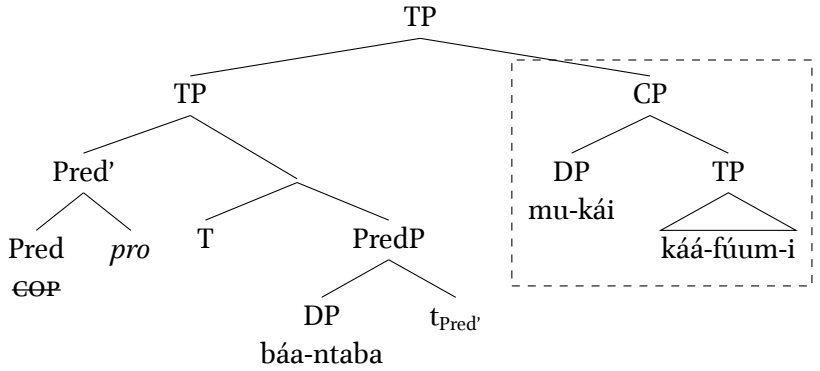
Now let us consider the derivation of what I have labeled as the reduced cleft in (94) and (95). In these two examples there is no overt copula preceding the initial predicative NP, and no relative pronoun on the verb.

(94) **Bá**a-ntaba mu-kái **káá**-fúum-i.
2-goat 1-woman 1SM.PST-buy-PST
'It was some goats that the woman bought.'

(95) Ngâŋwa taará **káá**-fúúm-í báa-ntaba?
9.truth 1.father 1SM.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).

(96)



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 mɔ́n **òkáár** bààn
day that PST see woman children
'the day that the woman saw the children'

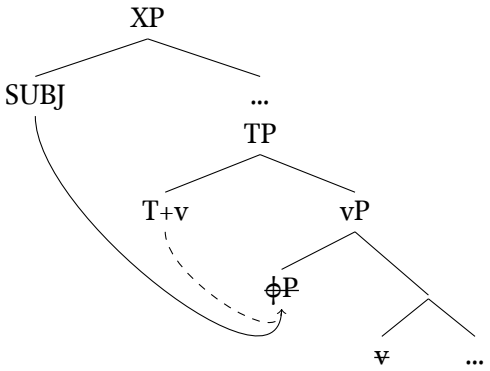
b. èsúú nà **òkáár** ò mɔ́n **ńdé** bààn
day that woman PST see she children
'the day that the woman saw the children'
[Nzadi B865] (Hyman 2012: 113)

(98) nɛ̌ **bààr** ó mɔ̂n?
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 φP which is structurally smaller (see (70) above) than a pronominal DP (see (74) above). This φP counts as a defective goal of T and agrees with the corresponding [uφ] on T, as illustrated in (99). The valued φ-features on T are spelled out as SM, while the φP as the lower end of the chain has no component left to be spelled out thus becoming phonologically null.[5]

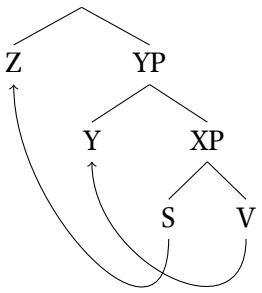
[5] As for Nzadi, the lack of subject agreement (Hyman 2012) in this language suggests that its T head has completely lost the [uφ] features, so the absence of the postverbal anaphor in (98) suggests that the pronominal element is simply deleted, rather than being reduced to a φP and agreeing with T. Alternatively, in Nzadi non-subject relatives there was once also a postverbal φP that is connected to the preverbal subject, and due to the loss of [uφ] on T, the unagreed φP can only be spelled out as an overt postverbal anaphor.

(99) *Agreement with a postverbal φP in Kukuya*



The Subject-Verb-φP 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 φP. 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.

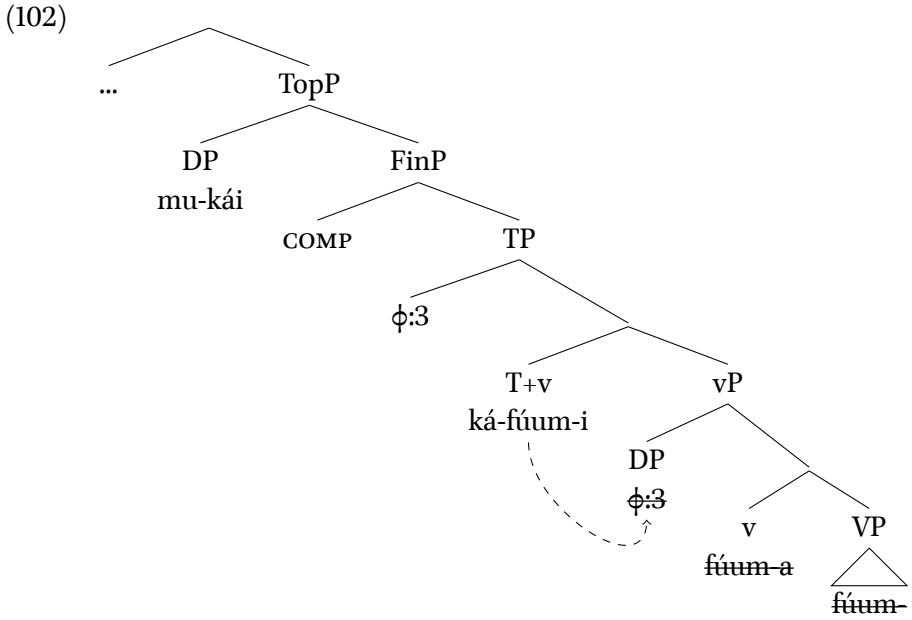
(100) *SVs Structure* (Hyman 2012: 104)



(101) ibaa na **mbvá**$_{i}$ ŋgo té **n**ɔ̌$_{i}$
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 φP 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 φP and the φP 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 φP 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 φP 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 φP, considering that the φP cannot be the trace/copy of a base-generated DP.

(102)



I propose that the φP 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 φP 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 ϕP 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 φP 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 ϕP 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 ϕP 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 ϕP. 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

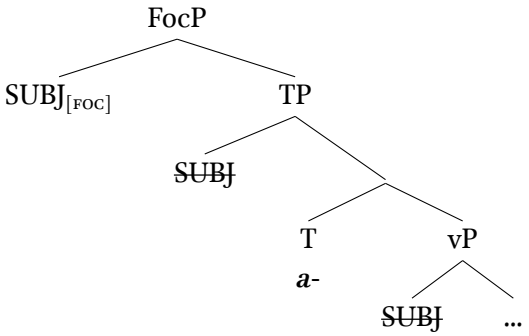
Now I discuss the class 1 subject marking alternation in the IBV focus construction, as repeated in (103).

(103) a. [Mu-kái]$_{\text{TOP}}$ [**mí-fémé**]$_{\text{FOC}}$ **káá**-fúum-i.
1-woman 4-pig 1SM.PST-buy-PST
'The woman bought the PIGS.' [IBV object focus]

b. [Mi-féme]$_{\text{TOP}}$ [**mú-kái**]$_{\text{FOC}}$ **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.

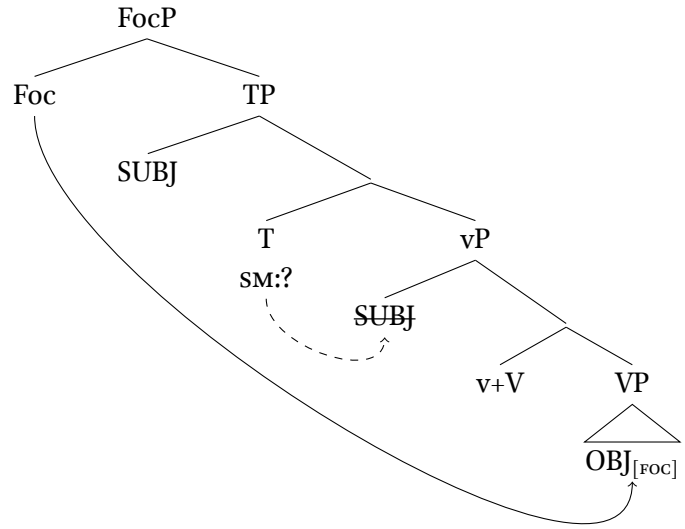
(104)



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,

(105)

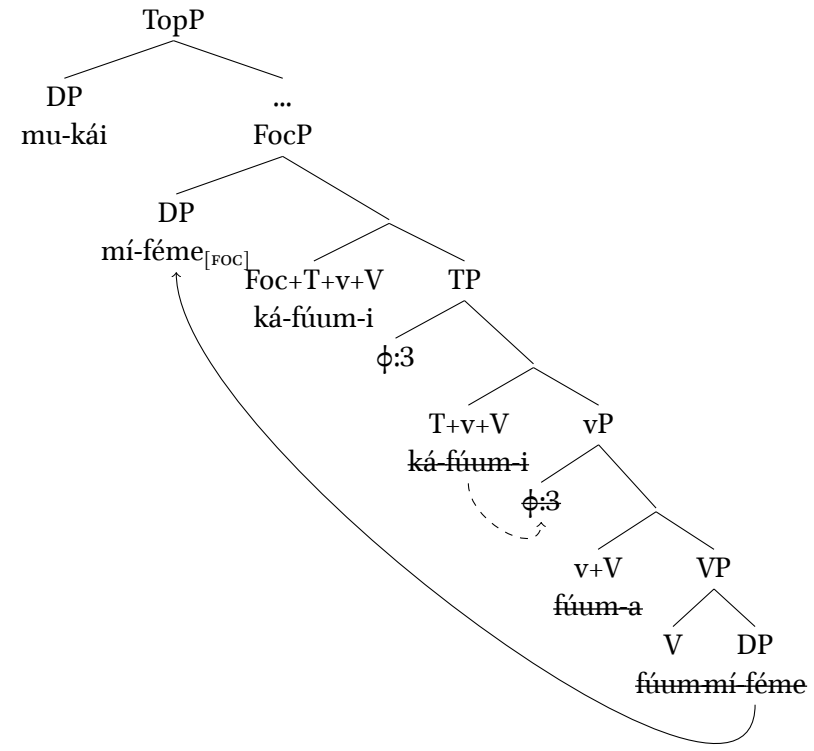


Recall that in the previous subsection I proposed that in a reduced cleft, the SM *ka-* represents the agreement between T and a ϕP that equals to the 3rd [Person] feature. Since diachronically the reduced cleft is supposed to be the precursor of the IBV focus strategy, I propose that the subject agreement 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 ϕP 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 ϕP that equals to 3rd person [Person] feature and encodes the animacy and givenness of the agent in the context. This ϕP 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] ϕP.

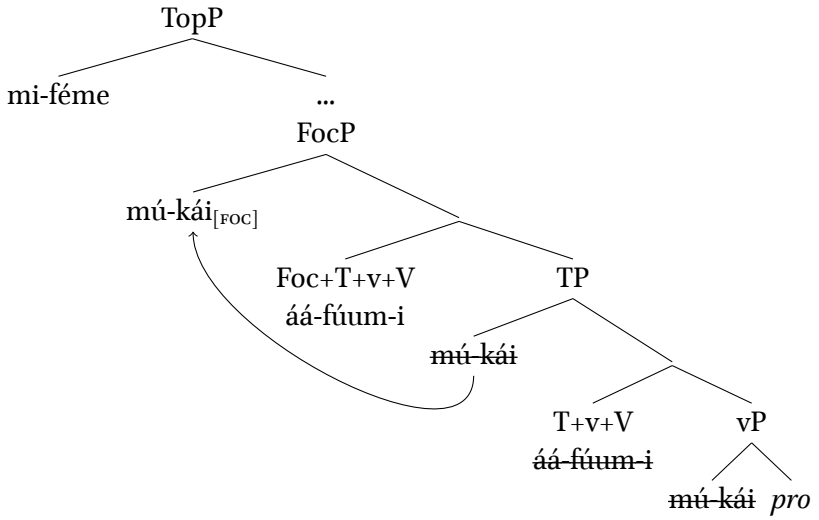
(106)



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 ϕP 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.

(107)



The separation of the initial topic from the rest of the sentence and the use of the ϕP 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 ϕP 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 ϕP 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) Kǐ-sî 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-wî?
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 ϕP that is controlled by the initial topic in the IBV focus construction.

(111) Mi-fémé mí-kâ-tsuomó **bóri** me ŋjé 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 ϕP. 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, mwáánà kí-má 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 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 lií-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 ϕP 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 ϕP 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 *míí-*, 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 **míí**-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 ϕP 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 coexist. 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.