# SYNTACTIC PATTERNS OF ANAPHORIC RELATIONS IN LUBUKUSU: Representation and Interpretation in a Minimalist Perspective

**Justine Mukhwana Sikuku**

University of NAIROBI Library
0451782 7

UNIVERSITY OF NAIROBI
LIBRARY

A Thesis submitted in fulfillment of the requirements for the degree of PhD in Linguistics at the University of Nairobi

2011

# DEDICATION

I dedicate this work to my wife, Hellen S. Wanyonyi, and my children, Letyciah Khalayi, Lavendah Namaemba and Leonell Wamalwa for being the exceptional source of my strength

## Abstract

This thesis examines how LuBukusu, a Bantu language spoken in western Kenya, achieves anaphoric relations in syntactic constructions. The main idea is that syntactic constituents, mainly NPs, are interpreted as being co referential with other elements either within or outside the sentence in which both the anaphoric element and its antecedent occur. The hypothesis that LuBukusu anaphoric relations motivate specific structural and discoursal patterning that can be accounted for by the available models of syntactic analysis like Minimalism forms the basis of analysis. In the generative framework, such anaphoric relations were characterized in form of binding principles. Whereas numerous studies have been done on this phenomenon, based on European and Asian languages, little has been documented on African languages. This study is a contribution towards filling in this gap. In addition, an analysis that reformulates the binding principles in minimalist terms on the basis of feature interpretation and valuation is developed. Towards this end, the study makes a theoretical contribution.

Contrary to many Bantuists, this thesis considers both the reflexive and reciprocal markers as incorporated pronouns with the status of arguments that undergo an analysis similar to other arguments. They are therefore not valence reducing strategies, but are only used to maintain the status quo. This sort of analysis underlies the complex relationship between morphology and syntax in agglutinating languages leading to a redefinition of the whole concept of argument marking. Consequently the study, describes both the morphological and syntactic properties of the anaphoric NPs, examines the binding relations involving these NPs, and develops an MP based analysis to account for the facts hitherto discussed.

The data used for the study were purposively generated (based on characteristics of anaphoric elements and antecedents) by the researcher as a native speaker of LuBukusu, and verified for acceptability by other native speakers. The results were analysed using tables, tree diagrams and informed generalizations for results and findings. It was for example, noted that the reflexive and reciprocal markers, Agr-*eene*, and the phrasal

reciprocal all posses unique morphological and syntactic properties such as agreement, and co-occurrence possibilities. They enter into binding relations that involve both the anaphoric elements and their antecedents. The affixal markers are largely local while the non-affixal forms are either local or non local based on semantic factors such as focus.

This thesis also develops a minimalist analysis of the anaphoric NPs taken in isolation. The major idea is that since the affixes are incorporated pronouns, they are analyzed as DPs that move to spec *v*P position for feature checking. The free NPs are analysed as adjuncts, or arguments depending on the role they play in a construction. A similar analysis is extended, to contexts involving co-occurrences of the anaphoric elements motivated by focus or one of the valence increasing strategies such as causative or applicative markers. In either case the motivating feature triggers movement for feature checking.

## DECLARATION

**Declaration by the candidate**

This Thesis is my original work and has not been presented for a degree in any other university.

JUSTINE MUKHWANA SIKUKU

21st April 2011
DATE

**Declaration by Supervisors**

This Thesis has been submitted with the approval of:

PROF. MOHAMED. H. ABDULAZIZ
DEPARTMENT OF LINGUISTICS &
LANGUAGES.
UNIVERSITY OF NAIROBI

21st April 2011
DATE

Helga Schroeder

DR. HELGA SCHROEDER
DEPARTMENT OF LINGUISTICS &
LANGUAGES.
UNIVERSITY OF NAIROBI

21/4/2011
DATE

Copyright 2011 by
Sikuku. Justine Mukhwana

**All rights reserved.**

CONTENTS

# ACKNOWLEDGEMENTS

Support in the preparation of this thesis came from various sources including family, my supervisors, institutions and friends. It would be impossible, though, to name every one of them, so the following list is far from being exhaustive.

My first thanks go to the University of Nairobi, Department of Linguistics and Languages for giving me the opportunity and a conducive working environment to write this thesis. The numerous comments by both students and staff in the department, especially during seminar presentations of parts of this work, proved to be quite invaluable. Thanks also to several colleagues at Moi University, Department of Linguistics and Foreign Languages for their continous encouragement.

I would also like to thank my supervisors, Prof. Mohamed, H Abdulaziz for his patience and expertise, especially in dealing with data related issues, and Dr. Helga Schroeder for bearing a part of my cross, pushing, directing, correcting and generally going beyond the call of duty to ensure the realization of this thesis.

I doubt if this work would have been the way it is today without the crucial American iink. It all started with an e-mail to Michael Marlo, then in Michigan State University, for comments on my study. Michael did not only offer the comments but he also gave me another crucial link to Prof. Ken Safir of Rutgers University, Linguistics Department who has extensively worked on anaphora. Since then, the numerous exchanges vial e-mail have been quite helpful. As if this was not enough, Michael introduced me to Michael Diercks of Georgetown University whose interest in LuBukusu syntax has natured a fruitful partnership whose benefits are enormous.

I return every bit of thanks to members of my extended family. To my mum, Susan Nanjilula Sikuku, I say, what you worked for was not in vain! To my dad Stephene Sikuku Wamalwa, although you passed on before the dream was realized, thanks for instilling in me the enthusiasm and hunger for education. To my brothers and sisters, thanks for propping me up. You all provided the opportunity for me to see, through you, what sheer hard work can achieve. Grandpa, 'Perenato' Wamalwa, thanks for reminding

me to 'read all the books until there are no more left'. *Muryo muno mweesi.*

The best is no doubt saved for last. I thank members of my immediate family in a special way. To my wife, Hellen Sitawa Wanyonyi, my three children, Letyciah Khalayi, Lavendah Namaemba, and Leonell Wamalwa, I say thanks for being the pillar on which I always lean on whenever the going gets tough. Your patience and understanding when I thought I was working too much at your expense gave me the motivation to do it all for your sake. This sacrifice was not in vain!

All Glory be to God!

## Abbreviations and Symbols

### General Abbreviations

**GB**- Government and Binding

**MP**- Minimalist Program

**LAR**- (Lubukusu) Anaphoric Relations

**UG**- Universal Grammar

**BT**- Binding Theory

**SDA**- Short Distance Anaphora

**LDA**- Long Distance Anaphora

**UD-Forms**- Unbounded Dependent forms

**EPP**- Extended Projection Principle

**PF**- Phonetic Form

**LF**- Logical Form

**DbP**- Derivation by Phase

**uF**- Uninterpretable referential feature

**uΦ**- Uninterpretable features of number and person

**uC**- Uninterpretable case features

**PIC**- Phase Impenetrability Condition.

**ECM**- Exceptional Case Marking

**MLC**- Minimal Link Condition

**DP**- Determiner Phrase

**VP**- Verb Phrase

***vP***- Light verb Phrase

**T**- Tense

### Morphological Abbreviations

**Prfx**- Prefix

**Agr**- Agreement affix

**RFM**- Reflexive Marker

**SM**- Subject marker

**fv**- Final vowel

**OM**- Object Marker

**Asp**- Aspect marker

**Nom**- Nominative case

**Acc**- Accusative case

**Tns**- Tense marker

**PRES**- Present

**FUT**- Future

**Comp**- Complementizer

**RCM**- Reciprocal marker

**NEG**- Negative

**Caus**- Causative

**Appl**- Applicative

**HAB**- Habitual

**CL**- Noun class prefix

**TAM**- Tense, Aspect, Mode

Ø- Null element

**List of Tables**

# Chapter One

## Introduction

### 1.0 Background

In the last five decades, the study of anaphora has mainly focused on the development of locality conditions to regulate the interaction between NPs across languages. As a result, many linguists have concluded that the distribution of anaphoric elements (NPs with the ability to co-refer with other NPs either sentence internally or externally) in most languages is universal: certain forms occur only with antecedents that c-command them and which are in the same local domain (clause) as the anaphoric element, while other forms must be disjoint from their antecedents within the same domain. This is the conventional generative approach as instantiated in the Binding Theory (Chomsky 1981). Safir (2004:7) gives a clear picture of the research agenda:

> At times, theories and analyses of these distributions have been the driving force in the formulation of theoretical constructs while at other times, when new theories held out less promise in this domain, research into the principles regulating anaphora has continued almost as vigorously, just not at centre stage ... it is not always clear where the centre of the stage is but research into the **patterns** of anaphora appears to have part of it. (Own emphasis)

Indeed any research in anaphora that pays particular attention to the patterns thereof holds part of the centre in the anaphora field. It then follows that most studies in this area are guided by two fundamental objectives;

i) Describe in details the patterns of binding (coreference) relations in any given natural language.

ii) Determine how far the relations described are:-

a) Universal i.e. conform to universal set of binding properties.

a) Universal i.e. conform to universal set of binding properties.

b) Parametric i.e. do not conform to such properties.

In this study, the description in objective i) is both in terms of form and interpretation while objective ii) is determined by the application of theoretical assumptions instantiated in the Minimalist Program (henceforth MP).

**1.1 The Language**

Guthrie (1971) classifies LuBukusu (E31c) as one of the seventeen or so dialects grouped under the Luhya sub-group of the wider Bantu. The Luhya people occupy most parts of Western Province of Kenya. However it has been argued that there is no such language as Luluhya. Indeed, what identifies the Luhya is a group of dialects, some of which display remarkable linguistic variations. Most of them are however mutually intelligible. For this last reason, this study will retain the use of the term dialect to refer to any member of the Luhya group based on the assumption that they all belong to some Proto- Luluhya (recently defined in the ethnologue as a macro-language with the individual dialects being refered to as languages). Fedders and Salvadori (1980) give a precise description of the Baluhya as follows;

> They are not a single entity, that they do not constitute a homogeneous group of people. . There are variations in dialects, the ritual, the emphasis on economic pursuits and other such cultural trials ... but- and this but needs to be stressed- all Luhya share a common language and some of the other essential elements of a total culture'. (1980:99).

Some of the other Luhya dialects include; LuLogooli, LuNyole, LuKabrasi, LuWanga, LuNyaala, LwIsukha, LwIdakho, LuTachooni, LuMarama, LuMarachi, and LuSaamia. It is also worth noting that LuBukusu is closely related to LuGisu, a language spoken in

Eastern Uganda. In fact, it is a general belief that LuBukusu is more mutually intelligible with LuGisu than it is with other Luhya dialects. Further, note that the prefix *Lu-* in Lubukusu denotes 'language'. Other prefixes include; *Bu-* which refers to the place where the language is spoken; *Ba-* denotes the speakers of the language; and *Omu-* is a singular form for the speaker of the language.

Within LuBukusu there are subtle dialectical variations mainly related to geographical location. As a result, three groups are easily identifiable; i) Western (Sudi); ii) Central (Sirisia); and iii) Eastern (Webuye). The variations inherent in these groups are primarily phonological and lexical but do not radically affect intelligibility. Since the differences are not grammatical, this study treats LuBukusu as a homogeneous group whose description is adequately captured in the data available for the study.

### 1.1.1 Origin and History of BaBukusu

From the 1999 population census, it is estimated that LuBukusu is spoken by close to a million people as a first language and a couple of thousands of others as a second language (these are mainly neighbours like Batachoni, Ateso, and Sabaot). In western Kenya, BaBukusu are spread in at least four districts[1]: Bungoma (referred to as the home district) and Lugari districts (both in Western Province) and Trans-Nzoia and parts of Uasin-Gishu districts of Rift-Valley Province. A few others have settled in major towns mainly for employment reasons. It is also worth observing that most of the areas occupied by these people are conducive for a variety of agricultural activities which the people engage in. At one time BaBukusu were referred to as 'Kitosh' mainly because of

[1] District boundaries have since been adjusted with more districts being created out of the larger Bungoma and Trans-Nzoia districts. In the new constitution passed in 2010, administrative bondaries are recategorized into counties, with BaBukusu occupying enlarged counties of Bungoma, Trans-Nzoia, Lugari, and Uasin-Gishu.

their ruthless fights with Kalenjin neighbours. This term was vehemently opposed and was subsequently done away with.

In their traditional folklore, it is revealed that BaBukusu originated from 'Emisiri' (present day Egypt). Their migration from this area was necessitated by the need to find fertile land for cultivation and grazing of their numerous animals. In addition, it was possible that they were fleeing from hostile conditions like war and diseases. Indeed, evidence of their economic undertakings is enshrined in most of their oral stories that are full of agricultural and pastoral episodes. The language as it is today is abundant with vocabulary associated with the related socio-economic activities. The very nature of the area that they occupied is also quite revealing. Presently, the areas of Bungoma and Trans-Nzoia are some of the most agriculturally productive areas in Kenya.

Socially, the BaBukusu lived in fortified villages (*chingoba* (plural), *lukoba* (singular)) based on their numerous clans. Such villages were appropriate as they would keep away enemies and also promote harmonious living. These people had rules that governed their co-existence as they participated in such important social activities like marriage, and circumcision. The rules were often enforced by elders.

### 1.1.2 Language Features

Like most Bantu languages, LuBukusu has five vowels, [a, e, i, o, u] with a corresponding number of long vowels [aa, ee, ii, oo, uu] formed by doubling the vowels. There is also a contrastive distinction between short and long vowels. See (1) below for illustrations;

| | | |
|---|---|---|
| **(1)** | Khula- *grow* | Khuula – *uproot* |
| | Sima – get *put out* | Siima- *like* |
| | Sila- *be quiet* | Siila- *swell* |
| | Ruma- *send* | Ruuma- *jump* |
| | Mela- *grow* | Meela- *get drunk* |

Long vowels may also result from the fusion of two vowels across morpheme boundaries. For example *omweene* (Agr-*eene*) is derived from omu-+ -ene (class 1+ stem), which also involves glide insertion of (*w*). Its plural *babeene* is a result of the fusion of baba + ene (class 2 + stem).

In rapid speech some of the vowels undergo modification based on ease of articulation. The most notable process here is elision which involves the omission of a sound segment that would be present in the deliberate pronunciation of a word in isolation. See the structure in (2) below for illustration;

| | | | | |
|---|---|---|---|---|
| **(2)** | a- a- i - siim- a | → | e- e- siim- a | |
| | Agr-Tns-RFM- like-fv | | Agr/Tns-RFM-like –fv | |
| | S/he likes himself/herself | | S/he likes himself/herself | |

Here, two phonological processes take place: Elision and vowel harmony/coalescence. One of the first two prefixes is elided leaving one morph to serve the two roles of agreement and tense marking. In addition, the impossibility of moving from a lower vowel [a] to a high vowel [i] in the process of articulation, triggers a process of vowel harmony/assimilation to a more central [e].

LuBukusu has twenty consonants characterized according to their place and manner of articulation and also according to whether they are voiced or voiceless. Further, they only occur at initial and medial word positions but never at the end. The table below is a

representation of Lubukusu consonants.[2]

| Manner of Articulation | Voiceless | Voiced | Place of Articulation |
|---|---|---|---|
| Plosives | P<br>t<br>k | | Bilabial<br>alveolar<br>velar |
| Fricatives | f<br>s<br>ʃ<br><br>x | <br><br><br>B | labio- dental<br>alveolar<br>palato-alveolar<br>bilabial<br>uvular |
| Affricates | tʃ | | alveo-palatal |
| Lateral | | l | Alveolar |
| Retroflex | | R | Alveolar |
| Nasals | | m<br>n<br>ŋ<br>ɲ | Bilabial<br>alveolar<br>velar<br>palatal |
| Pre-nasals | | mb<br>nd<br>nj<br>ŋg | Bilabial<br>alveolar<br>palatal<br>velar |
| Semi-vowels | | j<br>w | Palatal<br>Bilabial |

**Table 1: LuBukusu Consonants**

A noticeable feature of LuBukusu consonants is that the sounds /b/, /d/, /dʒ/ and /g/ do

---

[2] There is no standardized orthography for LuBukusu, but I use what is common practice in most writings in LuBukusu. These include: The use of Roman characters, and doubled vowel graphemes to symbolize long vowels as in *siima-* like/love, and the representation of certain unique consonants as follows;

i. The alveopalatal voiceless affricate [ʧ] is written as **ch**
ii. The voiced bilabial fricative [β] is represented as **b**.
iii. The velar nasal [ŋ] is represented as **ng**.
iv. The palatal voiceless fricative [ʃ] is written as **sh**.
v. The palatal nasal [ɲ] is written as **ny**.
vi. The uvular fricative[X] is represented as **x** or **kh**.

not occur as independent sounds instead, they must always be preceeded by the nasals /m/, /n/, or /ŋ/ as indicated in the table resulting in independent sounds also referred to as compound nasals or prenasalized stops.

It is also notable that most Bantu languages, LuBukusu included, have a very elaborate morpheme system characterized on virtually every lexical item, but it is perhaps more complex on the verbal form. Such morphemes normally occur as affixes attached to most words in different word classes and they mostly mark grammatical functions like agreement, noun class identification, tense, aspect, mood, negation, causation, passive and intensification. Of all the word classes, it is perhaps the verb (phrase) that is distinctly elaborate with both prefixes and suffixes. The rest of the classes mostly contain prefixes primarily marking agreement. For a detailed analysis of LuBukusu morphological patterns see Mutonyi (2000). Below, I briefly examine the noun and the verb, since they are directly relevant to this study.

#### 1.1.2.1 The Noun and Pronoun

LuBukusu nouns are elaborately classified into numerous classes normally referred to as noun class systems in Bantu. This categorization is based on pluralization patterns, agreement marking and patterns of pronominal reference. Demuth (2000:1) notes that;

> Noun classes are part of the larger concordial agreement systems where nominal modifiers, pronominals and verbs are all morphologically marked with the same noun class feature.[3]

[3] For a complete analysis of Proto- Bantu noun class systems, see Meeussen (1967), Guthrie (1967) and Welmers (1973).

In relation to the Proto- Bantu noun classes, the following classes can be identified in LuBukusu.

| CLASS | REPRESENTATION | PRE-PREFIX | PREFIX | EXAMPLES |
|---|---|---|---|---|
| 1 | Human/agentive- Sing. | O- | Mu- | Omuundu- person |
| 2 | Human/ agentive- Plur. | Ba- | Ba- | Babaandu- people |
| 3 | Non-human animates, bodyparts, things, plants. | Ku- | Mu | Kumukhono- hand |
| 4 | Non-human animates, body parts, things, plants. | Ki- | Mi- | Kimikhono- hands |
| 5 | Mass nouns, some animals, some body parts. some plants- Sing. | Li- | Li- | Litiisi- dam<br>Libuumbi- heap of rubbis |
| 6 | Mass nouns, some animals, some body parts, some plants – Plur. | Ka- | Ma- | Kamatiisi- dams |
| 7 | Things,instruments, some body parts – Sing | Si- | Si- | Sisiindu- thing |
| 8 | Things, Instruments, some body parts- Plur. | Bi- | Bi- | Bibiindu- things |
| 9 | Names of animals and Some plants | E- | N- | Enkaani- taboo |
| 10 | Things having length, some body parts, and utensils/ instruments. | Chi- | N- | Chinguulo- poles |
| 11/10 | " | Lu- | Lu- | Luluuchi- river |
| 12 | Diminutive, derogatory. | kha- | Kha- | Khakhaandu-small thing |
| 14 | Abstract nouns / State | Bu- | Bu- | Busiime- love |
| 15 | Gerundial/ infinitival verb forms. | Khu- | Khu- | Khusiima- to love |
| 16 | Locative- 'on' | Khu- | - | Khunju- on house |
| 17a) | Locative- 'at/by' | A- | - | Anju- at/by house |
| b) | Lacative-'towards' | Sya- | - | Syaanju- towards house |
| 18 | Locative- in/ from /to. | Mu- | - | Muunju- in house |
| 19 | Augmentative, derogatory, some plants, things and body parts. | Ku- | Ku- | Kuliango- big door |
| 23 | Locative-'at' | E- | - | Ebung'oma-at Bungoma |

**Table 2: LuBukusu Noun Classes**

In LuBukusu, there are different types of pronouns categorized as personal, reciprocal, reflexive, interrogative, demonstrative, and possessive. Some of these are illustrated in the table below:

| Pronoun Type | 1st Person | 2nd Person | 3rd Person |
|---|---|---|---|
| **Personal** | Ese Efwe | Ewe Enywe | Niye Nibo<br>Ni+Agr |
| **Reflexive/Agr-*eene*** | Samweene fwabeene | Wamweene Mwabeene | Omweene Babeene<br>Agr-eene |
| **Possessive** | Agr-ase Agr-efwe | Agr-owo Agr-enywe | Agr-ewe Agr-abwe<br>Agr+Agr |
| **Demonstrative (near)** | - | - | Oyu Aba<br>O/A/E +Agr |
| **Demonstrative (far)** | - | - | Oyo Abo<br>O/A/E +Agr+o |

**Table 3 LuBukusu Pronouns**

Note that *Agr* corresponds to the noun class affix of the noun that the pronoun stands in for. For the third person, the second line for each pronoun represents the non-human noun classes. Also notice that the possessive pronoun has two parts where the first represents the possessum while the second is the possessor.[4]

LuBukusu also allows a complex agreement form, Agr-*eene* on/with Agr-*eene,* as a reciprocal pronoun that often enhances the reciprocal reading marked on the verb by a reciprocal suffix, and is coreferential with an antecedent. This form will be the focus of much additional discussion as this dissertation proceeds.

### 1.1.2.2 The Verb

Like most Bantu languages, the LuBukusu verb is quite elaborate with a variety of affixes each representing a given grammatical function and more or less appearing in a

[4] As a strategy to break the triple vowel cluster [uee] in the reflexive pronouns, the glide [w] is inserted to form *omweene, wamweene,* and others.

predictable position in relation to the root, mainly as prefixes or suffixes. The prefixes include subject agreement, tense, aspect, and object agreement. The suffixes on the other hand are mainly the causative, applicative, intensive, reciprocal, passive, reversive markers, well known among Bantuists as verb extensions. In addition, there are tense and aspect markers and the final vowel which may also double as the mood marker. Several linguists working on the nature of affixes in Bantu (see Baybee 1985 and Baker 1985) argue that such affixes optionally or obligatorily occur in certain fixed slots in the verb (phrase). A typical Bantu verb has no less than eight slots for such affixes. In LuBukusu, the most notable are summarized in (3) below along the lines of Kraal (2005):

(3) a) **Pre- Initial:** Negative/ Relative marker conjunction
b) **Initial:** Subject Agreement
c) **Post Initial:** Tense, Aspect, Mode
d) **Pre Root:** Object Marker/ Reflexive Marker
e) **Root:** Verb root
f) **Post Root:** Derivational extensions (Causative, Applicative, Reciprocal, Passive, Intensive, Reversive)
g) **Pre- Final:** Tense, Aspect, Mode
h) **Final:** Final Vowel/ Mode
i) **Post Final:** Locative/ Negative Clitics

These facts can be illustrated in the examples below:

**(4a)** Ne- ba- a- mu- swen-el- ang- a- kho
Rel-SA-Tns-OM-step-Appl-Asp-fv-Loc
a b c d e f g h i
*When they stepped for him*

**(4b)** Se- na- a- ba –p- an- isy- an-a ta
Neg-SA-Tns-OM-beat-RCM-Caus-RCM-fv Neg
a b c d e f f f h i
*I did not cause them to fight against each other*

In the examples, the order of the elements is fixed except when elements of the same kind occur together. This is especially true for verb extensions in (f).[5]

Further. LuBukusu is primarily an SVO language with a very rich agreement system consisting of the marking of the nominal/ pronominal elements on the (pro)nominal modifiers, verbs and complementizers. Because of the rich agreement system sometimes the subject/object positions are syntactically null, but the content can always be understood from the linguistic context mainly by means of agreement. Consider the following examples:

**(5a)** Babaana ba- layi ba- lom- il- e ba- li ku- mu-liango
CL2-child Agr-good SM-said- Tns-fv Agr- that prefx-CL3-door
Ku- fun- ikh- e
Agr-break-Stat-fv
*Good children said that the door broke*

**(5b)** Wafula a- lom- ang- a a- li a- eny- a a- p- e Nafula
Wafula SM-say-HAB-fv Agr-that SM- want-fv Agr-beat- fv Nafula
*Wafula says that he wants to beat Nafula*

**(5c)** Wekesa a- e- siim- a omweene
Wekesa SM- RFM-like- fv Agr-own
*Wekesa likes himself*

(5d) Bibieene bi- a- kw- a mwi- loo mwe- embwa
Agr-own SM-Tns-fell-fv in- CL17hole Agr-of dog
**Themselves/they fell in the dog's hole*

[5] See Mutonyi (2000) for a detailed discussion of the morphology and occurrence constraints of verbal affixes.

In (5a), the *ba-* class 2 prefix is marked for agreement in the adjective *balayi*, the verb *balomile* and the complementizer *ba-li* while the *ku-* class 3 prefix is marked on the verb *kufunikhe*. In (5b), agreement with the subject '*Wafula*' is marked on the verb and complementizer of the matrix clause while the subject of the lower clause is marked on the verbs *enya* and *ape*. We also realize that the lower subject is not overtly represented, but its contents are recoverable from the agreement affixes. In (5c), the subject *Wekesa* is marked on the verb by the SM *a-* while the reflexive *omweene* is realized by the reflexive marker *-e-* attached to the verb. Finally, in (5d) the subject reflexive pronoun is marked on the verb while the locative prefix in *mwiloo* is also realized in the preposition *mwa*.

## 1.2 Statement of the Problem

The question of how NPs in a syntactic construction are interpreted as being coreferential with other NPs in the same construction or in the general discourse structure has always been a puzzle for linguistic theory, including Government and Binding (Chomsky 1981) and the Minimalist Program (Chomsky 1993, 1995, 2001). LuBukusu indicates relations between elements (mainly nominal) either morphologically, lexically or syntactically. The last two have received greater attention, especially in the Generative framework. Morphological marking, on the other hand, has received minimal attention, perhaps because it is mainly a feature of understudied languages such as Bantu. In addition, there are a number of complexities involved. Some of these complexities arise from the ability by most affixes to occur with an independent form playing a similar role. A reflexive marker (RFM) on the verb in LuBukusu, for example, optionally occurs with an *Agr-eene* form which raises interesting questions about how such cooccurrences should be syntactically interpretated and theoretically analyzed, and whether morphology plays any

general role in the mapping of syntactic processes. Consider the following examples:

**(6a)** *Wekesa* a-*e*-siim-a *omweene*
Wekesa SM-RFM-like-fv Agr-own
Wekesa likes himself'

**(6b)** *Wekesa* a-*e*-siim-a
Wekesa SM-RFM-like-fv
Wekesa likes himself'

**(6c)** Wekesa a-siim-a omweene
Wekesa SM-RFM-like-fv Agr-own
Wekesa likes him'

In (6a), the RFM occurs with the pronominal Agr-eene form which in turn is licensed by the RFM for it to have a reflexive reading. Conversely in (6b), the RFM does not depend on Agr-eene in a similar way. It can occur on its own and still succeed in marking reflexivity. The Agr-eene form alone in (6c) is free, and cannot be coreferential with *Wekesa.* Based on this, different questions arise: which of the two elements, Agr-eene and RFM, is the internal argument that is assigned case? What is the status of the remaining element? If the RFM can be an argument, how does this affect the language's word order? How does MP account for such facts? These and many other issues are part of the problem of this study.

Another problem concerns the usual distribution of anaphoric elements. It is common knowledge that all languages exhibit a given distribution of anaphoric elements and that this distribution may or may not be subject to universal conditions. Such universal conditions are numerous but a pair of opposing conditions can be generated to act as a springboard towards identifying even more intricate properties of LuBukusu anaphora. The conditions are stated in the generalization below:

**Generalization 1**

There are certain lexical items/ anaphoric elements that occur only with antecedents which c- command them and which are located in the same local domain (clause) as the anaphoric elements. Conversely, some other anaphoric elements are unable to take c- commanding antecedents within the local domain.[6]

These opposing conditions provide room for a reanalysis based on the view that LuBukusu displays both local and contextual anaphoric relations. The nature of such contradicting relations forms part of the puzzle in this study. As a whole, the following sets of issues form the background upon which the study problem is stated: The morphological and syntactic structure of anaphoric elements; the distribution of the anaphoric elements in terms of independent occurrences and co-occurrences resulting in varied binding relations; and the theoretical analysis of such relations on the basis of MP.

As a result, the following research questions are raised:

i) What is the structural composition of the varied anaphoric elements?
ii) What is the role of morphology in syntactic interpretation?
iii) Do the Anaphoric elements find interpretation strictly within the sentences that they occur in?
iv) What is the nature of the binding relations existing in short and long distance environments?
v) Which is the best way of conceptualizing a linguistic theory that can capture all the anaphora facts in LuBukusu, and by extension, in Bantu?

## 1.3 The Goal and Objectives of the Study

The study's main aim is to describe the general anaphoric phenomenon in LuBukusu as

[6] The relation of c- command is part of the Binding theory of GB (Chomsky 1981). It was largely a condition on the relation between a given anaphoric element and a potential antecedent

instantiated in overt anaphoric NPs with significant emphasis on delimiting the inherent binding relations available for expressing this phenomenon, all done within the context of MP.

The following specific objectives are thus formulated:

i) To describe the types of anaphoric NPs in Lubukusu
ii) To determine different binding relations involving the different NPs in LuBukusu.
iii) To identify the syntactic and discoursal constraints governing LuBukusu anaphoric relations.
iv) To work out the most appropriate analysis of LuBukusu anaphoric relations, and by extension, other Bantu languages, within MP.

## 1.4 Hypotheses

Based on the aforementioned objectives, this study is guided by the following general hypothesis: LuBukusu anaphoric relations motivate specific structural and discoursal patterning that can be accounted for by the available models of syntactic analysis like Minimalism. As a result, the following are the specific hypotheses.

i) All the anaphoric elements are realized in a structurally similar way.
ii) The anaphoric elements are construed within the minimum clause in which they occur.
iii) Anaphoric relations are constrained by both syntactic and discoursal factors.
iv) The arguments of the Minimalist Program can capture all the anaphora facts.

## 1.5 Justification for the study

The need for such a study arises from both empirical and theoretical reasons, which have incidentally provided more than enough reason for linguistic studies, for many years. The fact that this thesis investigates data from an understudied language such as LuBukusu, the immediate result is to extend the data base of mainstream theoretical linguistics, and make it necessary for a theory concerned with language universals to test its strength in the face of data from typologically interesting languages.

An important empirical reason for this study is that sometimes it describes data from a theory neutral stand point. This is important because it ensures that the study remains relevant even when certain theoretical constructs have reached the end of their usefulness. Description of data for its own sake is an important process towards language documentation.

Indeed the ultimate objective of most linguistic studies is to try and explain language acquisition. Safir (2004) asserts that the existence of certain interpretive aspects conditioned by syntax (read 'anaphora') form one of the more formal arguments for the existence of an innate linguistic faculty independent of other cognitive capabilities. A study of patterns of anaphoric relations can immensely contribute to an understanding of different aspects of Universal Grammar and by extension, issues of language acquisition

## 1.6 Scope and Limitations

This study describes representations and interpretations of various types of anaphoric binding in LuBukusu. It pays particular attention to overt anaphoric NPs which include reflexives and reciprocals in their affixal and non- affixal forms. Both the syntactic and

discoursal environments are examined to identify features that have never been reported before. In addition, such relations are discussed in the context of MP provisions contrary to the earlier practice which mainly made use of GB. To do this effectively, the morphological and syntactic properties of each anaphoric element are highlighted mainly to capture the characteristics that are both unique and crucial in determining the representation of LAR. The focus on reflexives and reciprocals is crucial as it not only narrows the scope but it also helps to capture intricate details that might have been otherwise overlooked. In any case, these elements form the basis upon which anaphora theories are formulated. Lastly, the delimitation of syntactic and discourse constraints helps in characterizing the very nature of LAR.

In morphologically agglutinative languages such as LuBukusu, all agreement markers are ideally 'anaphoric' in character i.e. they carry properties of some other lexical form appearing elsewhere either in the same sentence or in the entire communicative context. A SM, for example, represents a subject, an OM, an object and so on. Such forms of anaphora are excluded because of various reasons: a) The SM reperesents an R-expression whose occurrence does not require an antecedent for interpretation and is therefore not subject to binding relations; b) The OM is always free, and if it is referential it requires an extra sentential antecedent. Such elements will only be discussed whenever they relate with the anaphoric elements that are the subject of this thesis. The same is true with other verbal affixes such as tense, aspect, and the varied verbal extensions.

The idea then remains that anaphoric elements that have no potential of selecting an antecedent from the minimal clause are excluded as anaphors.

## 1.7 Definition of Key Concepts

In this section, I define some of the key concepts that form the basis of the study. These include:

*Agglutinating language:* A language that joins its morphemes together with no easily observable boundaries.

*Anaphoric Relations:* Co-referential relationships, both syntactic and discourse-based, that exist between noun phrases (and other syntactic categories) in a linguistic environment.

*Barriers:* These are mainly maximal projections which prevent government across them.

*C-command:* Deals with dominance relations between nodes in a tree diagram such that x c-commands y if neither dominates the other and the first branching node above x also dominates y.

*Co- indexing:* Assigning similar indices to two or more co-referential elements in structure.

*Coreferential:* When two or more expressions refer to the same real- world entity.

*Feature Checking:* The process of ensuring that the syntactic derivation has no uninterpretable features.

*Functional Category:* Syntactic unit expressing grammatical relations rather than lexical relations.

*Governing category:* The governing category of x is (roughly) the minimal clause containing x and its governor.

*Head:* basis of phrasal projection: lexical elements such as V, N, A, P are lexical heads; non-lexical elements such as *v*, I and C are also called functional heads.

## 1.7 Definition of Key Concepts

In this section, I define some of the key concepts that form the basis of the study. These include:

*Agglutinating language:* A language that joins its morphemes together with no easily observable boundaries.

*Anaphoric Relations:* Co-referential relationships, both syntactic and discourse-based, that exist between noun phrases (and other syntactic categories) in a linguistic environment.

*Barriers:* These are mainly maximal projections which prevent government across them.

*C-command:* Deals with dominance relations between nodes in a tree diagram such that x c-commands y if neither dominates the other and the first branching node above x also dominates y.

*Co- indexing:* Assigning similar indices to two or more co-referential elements in structure.

*Coreferential:* When two or more expressions refer to the same real- world entity.

*Feature Checking:* The process of ensuring that the syntactic derivation has no uninterpretable features.

*Functional Category:* Syntactic unit expressing grammatical relations rather than lexical relations.

*Governing category:* The governing category of x is (roughly) the minimal clause containing x and its governor.

*Head:* basis of phrasal projection: lexical elements such as V, N, A, P are lexical heads; non-lexical elements such as *v*, I and C are also called functional heads.

*Inherent case:* case assigned at LF and is dependent on the thematic properties of the head.

*Interpretability:* (Of features) reffering to essential properties of morphosyntactic objects used to license them in the course of the derivation.

*Light Verb:* In a VP shell structure, it is the head of the higher vP present to license the external argument.

*Local Anaphora:* Where the anaphoric element requires a clause mate antecedent.

*Logophoricity:* Semantic/ pragmatic conditions such as point of view, or source that determine the identity of an antecedent for a given anaphoric element.[7]

*Long Distance Anaphora:* refers to the interpretation of any given anaphoric NP outside the minimal clause domain in which it occurs, but restricted to the same sentence.[8]

*Native speaker intuition:* The innate ability of a native speaker of language to recognize or form acceptable structures in the language.

*Parameters:* language specific characteristics that make such a language unique.

*Theta-roles:* Lexico-semantic characteristics assigned to sentential constituents.

*Unbounded Dependent Forms (UD-forms):* Anaphoric elements that may allow an extra sentential antecedent in addition to a syntactic one and which are not syntactically restricted[9]

---

[7] The term 'logophoric pronoun' was coined by Hagege (1974) to describe a special sub-group of personal pronouns used in reported speech in some African languages to refer to individuals whose speech is reported or to the addressee of that speech.

[8] This categorization is motivated by the fact that in some languages, condition A of BT does not hold as stated. (See Cole, P et al 1998). Using the Chinese reflexive 'Ziji', they define long distance anaphors as those that can take antecedents outside their local domain. Most linguists (See especially Koster and Reuland 1991) looking at LDA realized that the anaphors share a number of characteristics across languages. Some of the characteristics include monomorphemicity, and subject orientation

[9] The term was coined by Safir (2004) to refer to anaphoric elements that can be anteceded beyond the domain of the tensed or indicative sentence, can appear as the subject of a tensed or indicative sentence, can have their antecedent outside of a relative clause in which they are embedded, among other characteristics.

*VP Shell:* A complex internal constituent structure of VP where the VP splits into layers of VP projections representing the subject, object and other arguments.

## 1.8 Research Methodology

In this section I explain the methodological design used in the study.

### 1.8.1 Methods of data collection

The study used two methods of data collection: i) The researcher's native speaker competence to generate appropriate data displaying Lubukusu LAR; ii) Native speakers of Lubukusu as informants to verify the generated data.

The choice of these methods was based on the notion that the researcher, as a native speaker, knows what he is looking for and hence saves a lot of time and prevents the possibility of unwanted data. Moreso, Performance data alone may not easily display the language features which are under investigation. Commenting on the possibility of having the native speaker linguist as the informant, Horrocks (1987:11) notes:

> It is simply absurd to wait for native speakers to produce utterances which would allow linguists to infer whether some language has a particular grammatical characteristic when it is perfectly possible for the linguistic as a native speaker to ask all important questions and answer them himself.

The choice of these methods is also in line with a lot of work on generative syntax which incidentally relies on native speaker judgements on acceptability. However, one has to be cautious with such native speaker judgements because occasionally, they may turn out to be quite subjective. Therefore, to prevent generalizations that are based on incorrect data, I use the methodology advice presented in Featherston (2007). To him, linguists should, as much as possible, use data that is uniform and representative. To guarantee this, some

standards are set, generally classified as Essentials and Desirables. The former includes the use of multiple informants, and multiple lexical variants of the structures. The latter relates to task and scale which spell out response to input and multiple degrees of well formedness. In this study, the use of native speaker verification and variation in sentence structures is aimed at conforming to the essentials, while the inclusion of structures with high degrees of well formedness meets the standard set in the desirables.[10]

The study targeted data from LuBukusu displaying anaphoric relations. The phenomenon is mainly represented by reflexives and reciprocals in certain positions in a sentence. Only data that, as much as possible, gave an exhaustive representation of the different positions of the anaphoric NPs was included. In addition, cases of anaphora marked inherently on the verb without an NP were also considered. In this regard, the generated data was constrained to include the following:

(i) Reflexive Marker

(ii) Reciprocal Marker

(iii) Lexical Reflexives

(iv) Inherent Reciprocals

(v) Phrasal Reciprocal

(vi) Agr-*eene*

Data collection was therefore done in two phases based on the sample for the study. In the first phase, native speaker intuition from introspection was used. For each of the anaphoric relation, twenty-five structures (a number deemed to be as representative as

[10] In this thesis, degrees of wellformedness are clearly indicated as follows: wellformed is unmarked; doubtful but looks ok uses (?); illformed is preceded by (*).

possible) were provided based on the following characteristics:

(a) Different positions of the anaphoric element in a sentence

(b) Relative positions of potential antecedent

(c) Varied person and number features for the anaphoric element and antecedent.

(d) The distance between the antecedent and the anaphoric element in terms of the complexity of sentences.

(e) Agreement between the antecedent and the anaphor.

These characteristics were to ensure that the data generated represented the LAR phenomenon as exhaustively as possible.

The second phase of data collection concerned the verification of the data using adult native speakers selected from Tongaren Division of Bungoma North District. This was mainly because it was believed that adult native speakers have high levels of competence. In addition, the selected region is my home area, which means that I know the people well hence making the selection of informants easier. Ten literate native speaker informants were selected based on the resercher's knowledge. They were given a list of the structures generated by the researcher with instructions to indicate against each structure the degree of acceptability. Those that were generally acceptable were unmarked, the doubtful ones were marked with a question mark, while the unacceptable ones were shown with an aesterick. They were also free to offer alternatives especially in cases where they did not agree with a given structure. The results were then analyzed for structures that were acceptable to a majority of the informants. Such structures then

became the core data for the study. On the other hand, the doubtful and unacceptable cases were used to support negative generalizations on what LAR is not.

### 1.8.2 Procedures for data analysis

From the isolated data, an analysis was done based on the objectives of the study. To achieve the objectives, two procedures were used to analyze the structures selected for each anaphoric element:

(i) The use of the conventional tree diagram (phrase structure) representations for each of the sentences. This was to show the syntactic domains (representation) of the anaphoric elements vis-à-vis their antecedents (interpretation).

(ii) Generalizations were made about the nature of LAR and how they are handled by the different language models. On such generalizations Radford, (1981: 22) notes:

> Unfortunately, there is no known set of inductive procedures which the linguist (or anyone else) can apply to a given set of data to find generalizations. The simple answer is that you have to make an intelligent, informed ' guess' about what principle or rule might be needed to account for a particular phenomenon.

The informed 'guess' in this case is the knowledge of MP as it applies to other languages. Some comparison is therefore made with other languages, especially Kiswahili, in order to determine LuBukusu parameters.

## 1.9 Theoretical Framework

Although, the study intended to primarily use the Minimalist framework to describe anaphoric relations in LuBukusu, such a description is done against the background of GB assumptions mainly because MP arose as a reaction to the shortcomings in GB. Further, the choice of theory has been motivated by the fact that MP assumptions are

largely precise and can therefore deal with a wide range of phenomena with adequate clarity. In addition, MP is a more recent development in syntactic theories and hence is able to solve many problems associated with more traditional theories of grammar. Also the fact that MP is the most current theory in mainstream syntax provides an additional motivation for using it.

In this study, GB theory refers to the version presented in Chomsky (1981, 82) and subsequently revised and modified by Chomsky and other linguists in later years. Both the original and revised formats are used depending on their suitability to handle the data at hand. The Minimalist Program on the other hand refers to the model presented in Chomsky (1995, 2000, 2001a, and b) and illustrated by different linguists like Reinhart (1997), Safir (2004), Reuland (2001) and others.

It is clearly stated in generative literature that the main aim of the Minimalist Program (MP) is to offer a very different view of UG from that offered by GB based theories. It is also true that MP consists of different 'versions' because it is a program and not a theory[11]. In its analysis of syntactic phenomena, it makes use of certain methodological ideals that are reflected in concrete principles used in minimalist models.

As a starting point, we note that MP is derived from the success of GB. How do we quantify such success? Remember that the main goal of generative theories was to provide answers to a critical problem: how is it that children are able to acquire language up to high levels of grammatical competence yet the nature of the data that acts as input

[11] According to Chomsky MP is still evolving, and is not yet a full fledged theory. This study contributes to the evolution.

is impoverished.[12] In addressing this critical problem, GB's success receives its quantification. This is where the explanatory adequacy requirement for linguistic theories is realized. The idea is simple: Children are biologically equipped with a set of principles of grammar construction (UG). Metaphorically speaking, the principles of UG have open parameters easily equated to sets of societal values e.g. respect, and marriage values. Specific grammars (e.g. LuBukusu and English) arise once the parameter values are specified (i.e. contextualized in a specific language just as societal values are contextualized within a specific ethnic group). The Primary Linguistic Data is the basis upon which the parameters are determined. Using our sociological metaphor – the societal values are determined on the basis of their expression in real life situations. A language specific grammar is therefore simply a channel or vector specifying the parameters that the principles of UG leave open. Indeed it is true that the environment and a child's cognitive capabilities all play a significant role in the acquisition process. It is also true that syntactic research since the 1970s has aimed at elaborating such a picture and demonstrating its viability. In order to set the stage for a further elaboration of MP, it is perhaps crucial to give a brief overview of GB.

### 1.9.1 Distinctive Features of Government and Binding

The following are the general features of GB: a) GB is Modular: It is divided into various sub-components sensitive to a variety of well-formedness requirements; b) GB has a general Transformational component: It contains movement rules (i.e *move alpha*) and construal rules which index nominal expressions to another; c) GB has four distinct Levels of Representation: D- Structure, S- Structure, Phonetic Form (PF) and Logical

[12] The data that constitutes the language input for the child has been referred to as Primary Linguistic Data (PLD) and the problem of explaining language acquisition is what Chomsky calls 'Plato's Problem'. Chomsky (1986b).

Form (LF) which are assigned to sentences in any natural languge; and d) The Central Grammatical relation in GB is Government: This is a 'common factor' relation appearing in every module of the grammar. Government provides conceptual unity to rather diverse components. It mainly deals with the sphere of influence that lexical heads have over their complements. It is this influence hat enables such a head to assign grammatical properties such as case or theta roles to its complements.

In any discussion of anaphora, the starting point is always the Binding theory (henceforth, BT) developed by Chomsky in a series of works from 1973 on. BT's main concern was the characterization of typologies of NPs with the main conclusion that there are three types of NPs each governed by a single Principle, labelled as A, B, and C, and popularly known as the Principles of BT. Below, we give the formulation of the Principles as stated in Chomsky (1981).

A. An anaphor is bound in its governing category.
B. A pronominal is free in its governing category.
C. An R-expression is free.

The definition of 'governing category' is given in section 1.7. Chomsky (1982) and Reinhart (1983) both argue that condition C dealing with referential expressions is not part of the grammar. On the other hand, It is noted that forms that require an antecedent such as 'themselves' are treated as syntactic anaphors and subject to the locality principle A. This instantiates what is commonly referred to as local anaphora. Further, the binding relation between an anaphor and its antecedent is licensed in terms of c-command defined in section 1.7.

Turning now to principle B of BT that deals with pronouns, it is noted that they are not allowed to be coreferential with any antecedent that binds them in the same local domain

since they are in complementary distribution with anaphors. In other words the occurrence of pronouns is excluded in environments where anaphors occur.

In order to justify the choice of MP in dealing with anaphoric relations in LuBukusu, I, next, briefly outline the shortcomings of GB, and more specifically those of BT.

### 1.9.2 Shortcomings of Government and Binding Theory in dealing with Anaphora

It has always been argued that BT principles apply either at S-structure, the level that feeds both PF and LF, or at LF, the place for semantic interpretation. With the elimination of S-structure and D-structure within minimalism (see Safir 2004 and Hornstein et al 2005 for description of how the two levels were rendered redundant in MP) then LF remains the only level at which BT must hold. If this is the case then there is need for a radical restatement of the working of binding relations. I suggest possible ways of reformulating BT principles based on MP assumptions.

Principle A of BT does not take into consideration variations across languages and even within languages. In LuBukusu we have noted cases where the anaphoric elements are bound outside the conventional governing category. Within English, the so-called picture NPs also present similar problems to the definition of governing category.

For empirical reasons, the procedure in BT principles is naturally out of step with MP thinking. In MP there has been a tendency to avoid rules of construal and instead try to constrain the output of the syntax by algorithms defining derivations (Epstein and Seely (2002); Frampton and Gutmann, (2002); Chomsky (2001a,b)). To the contrary, however, binding principles have always been stated as filters on representations. Binding in a filter-

based GB analysis is for example only allowed if there is a c-command relation within a defined domain. This contrasts sharply with MP assumptions that words/morphemes selected for use in the derivation (the (N)umeration) are inserted into the tree by a derivational structure building process of adjoining known as Merge or Move. MP is thus regulated by the twin procedure of move and merge.

The use of indices to mark construal relations also runs into problems in MP especially with the Inclusiveness Condition. The condition is stated in the generalization below.

> **Generalization 2: The Inclusiveness Condition**
> Any structure formed by the computation is constituted of elements already present in the lexical items selected for N [Numeration]: No new objects are added in the course of the computation apart from the rearrangement of lexical properties. (In particular, no indices). (Chomsky 1995:228).

This raises a serious question, therefore, of how to account for anaphoric relations in syntactic theory, a question which this thesis aims in part to address.

### 1.9.3 The Minimalist Program

Against the background of the shortcomings, MP comes in to try to explain anaphoric relations in different terms. Specifically it shows that BT conditions that hold for anaphoric elements can be established in narrow syntax, indices can be eliminated and the interpretation can be restricted to phases. Following the arguments of Heinat (2006), we shall assume that anaphor-antecedent relation is a form of agreement relation (Reuland 2001), in the same sense as the case-Φ-feature relation (Chomsky 2000, 2001 a, b), and that all unvalued features act as probes. In this regard, the requirements of BT principle A are part of the

syntactic computation, which, within narrow syntax, does not distinguish between pronouns and anaphors.

Generally, GB has been successful in meeting the demands of explanatory adequacy. However, much emphasis on such methodological standards has consigned other standards of theory evaluation to the periphery. Standards such as simplicity, naturalness, elegance, and parsimony had almost taken a back seat. In MP, these become the center of focus and towards this end, the question of how to quantify such evaluative notions is the preoccupation of MP proponents. To simplify this preoccupation, MP is based on three considerations: a) Minimal conditions of adequacy which takes the assumption that any theory of grammar must accommodate certain minimal requirements; i) UG has a principles and parameters format; ii) Sentences are the basic linguistic units; iii) Sentences are pairings of sounds and meaning; iv) There is no upper limit to the number of sentences in any given language; v) Sentences show displacement properties in the sense that expressions pronounced in one position are interpreted in another; vi) Sentences are composed of words organized into larger units with hierarchical structure; b) Methodological economy which has to do with issues such as simplicity and elegance where the general rule is 'less is always better than more'. One module is better than many. Perhaps here is where the challenge against GB is based; and c) Substantive economy which is based on the assumption that when an operation involves least effort then it is far much better than that that does not. For example movement only applies when it must and no expression occurs without purpose. Wastage is not allowed.

All these considerations create a specific strategy of trying to get the simplest theory whose operations have a least effort touch and which accommodates the *minimal* conditions about grammar. In the past decade linguists working in Generative Grammar have been preoccupied with the actualization of the strategy. In this study I focus on how the strategy works using anaphoric relations in LuBukusu.

Further note that any adequate account of the morphosyntax of a given language must take into consideration sets of features that are said to represent phonetic and semantic properties. Chomsky et al, (2001a) give a clear picture of the notion of features. To them, languages have a dislocation property i.e. they are interpreted at two levels: phonetically at the edge (some kind of surface structure) and semantically (thematically) at the local position or the position of merge. It is this dislocation that is the result of the computational system. The concern of minimalism is about how languages implement the dislocation property or what the computational system looks like. When I think about the procedure of implementation, the analogy of a recipe comes to mind. One always requires ingredients and the procedure. The ingredients here will consist of three types of features; a) Feature identifying a target and determining the kind of expression that moves to it. (Mainly redundant features of agreement); b) Feature identifying thing to be dislocated. (Usually identified by structural case) c) Feature deciding whether the target has an extra- position or not. (Extended Projection Principle (EPP) feature). All the three are inflectional features that are said to be uninterpretable (abstract). They are only present to implement dislocation because all expressions must be interpretable at the interface for sensory motor level to interpret it. And once their job is finished, the uninterpretable features are eliminated by the computational system.

We have already noted in the introduction to this section that languages have displacement properties in the sense that sentences are expressed in one position and interpreted in another. How is this accommodated? In GB, movement and construal processes are used to explain this. MP requires that all these explanations be unified as one explanation is better than two separate explanations. Therefore construal is either movement or movement is construal. Many linguists have favored the latter version. See Chomsky (1986b). This study views anaphora on the basis of such explanations to determine the extent at which varied issues could be unified. The version of MP that I use is the Derivation by Phase (Chomsky 2001). I review how it works next.

### 1.9.4 Derivation by Phase (DbP)

The use of the term *derivation* is a direct consequence of the fact that the syntactic component is considered to be strictly built step by step within units identified as phases that contain elements known as probes and goals charged with the responsibility of implementing the derivation. An understanding of the key concepts, probes, goals and phases is therefore important if the whole process is to be made explicit.

#### 1.9.4.1 Probes and Goals

In order to define these concepts, it is pertinent to review MP thinking on lexical items and their features. Reuland (2001:16) gives a summary of MP basics in relation to features. These are outlined below:

I. Lexical items are associated with a set of grammatical features (formal features);

a) In the lexicon, lexical items are listed with inherent features (e.g. person and gender for N);

b) Upon insertion into the numeration, optional features are added (e.g. number and case for N, person and number for V);

II. Features come in two kinds; Interpretable (e.g. number for N) and uninterpretable (e.g. case for N, all F – features for V);

III. At the interface the uninterpretable features must be deleted;

IV. Movement is triggered by an attracting feature;

V. Covert movement moves formal features only;[13]

VI. The features of a complement move at most as far as (the features of) its head.

In addition, note that Chomsky (1995:297) also assumes that feature movement is governed by attraction stated below:

**Generalization 3 Attraction**

K attracts F if F is the closest feature that can enter into a checking relation with a sub-label of K. There is no separate checking operation. After checking the feature is automatically deleted when possible. A deleted feature is erased when possible.

From this summary, one can conclude that syntactic dependencies between constituents (Such as anaphoric relations) are derivative of dependencies between features that they contain.

Pursuing the notion of interpretable and uninterpretable features, Chomsky (2001 a, b) stipulates that the heads v and T are probes because they have uninterpretable F- features

[13] This movement of formal features was abandoned by Chomsky (2000, 2001), in favor of claiming that all movment moves the full DP, but in the case of covert movement, because all movement involved copying a DP into another position, it is simply pronouncing the lower copy of a DP, in the newer framework.

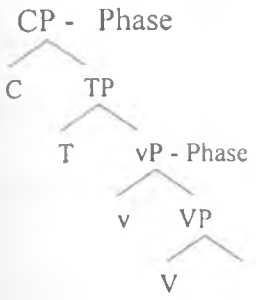
(henceforth uF). These uF must be checked/ deleted for the derivation not to crash at spell-out. The heads with uF therefore look (i.e probe) for matching features in their domain i.e. the structure that they c-command. Remember that Nouns (DPs) have interpretable F-features but uninterpretable case features (henceforth uC). uC make them active goals for searching probes. When a probe finds a goal the probe gets a value for its uF. They thus become valued and are marked for deletion. The goal in turn gets a value for uC, which also becomes interpretable and is also marked for deletion. We shall explore the implementation of the movement relations in chapter four. The next section explains the concept of phases.

**1.9.4.2 Phases**

The syntactic derivation normally proceeds in successive cycles, which Chomsky (2000, 2001) explains in terms of 'phases'. To Chomsky, a phase is either a CP or vP and the operation is in such a way that only the next lower phase head and its specifier(s) are available for operations. This conditional constraint is called the Phase Impenetrability Condition (PIC) that is a near parallel to the BT principles of 'barriers' which basically delimits the sphere of relations between syntactic elements to the CP (Chomsky 1986). Chomsky provides a detailed account of the conditions that constrain government and movement. More specifically, the notion of 'barriers' is introduced as a key feature in specifying such boundaries. The main difference, however, is that PIC is within minimalist thinking but both are conditions on the distribution of NPs. According to PIC, the complement of a phase head is sent to spell- out for phonological and semantic interpretation at the next higher phase, and a probe can find a matching goal in the next lower phase but not further down. At spell-out, all unvalued features must have a value

otherwise the derivation will crash. The structure of phases is represented below in a model representation of a sentence.

(7)




As a consequence of PIC, only the spec(s) of v and the head of v are available goals for probes in the next higher phase. v and its spec(s) are sent to spell-out when the next higher phase is merged.

Having given the general framework, I briefly illustrate how the properties outlined apply to LuBukusu data.

Note that structure building operates with two basic operations, Merge and Move. Although there is current debate on exactly how Move is to be formulated, the differences between the current proposals are minute. The following discussion follows Chomsky's original proposal. Merge is a function that takes two objects (say α and β) and merges them into an unordered set with a label (either α or β, in this case α). The label identifies the properties of the phrase. This operation is shown in (8).

(8) Merge $(\alpha, \beta) \rightarrow \{\alpha, \{\alpha, \beta\}\}$

Here the labels α and β represent sentential constituents such as VP. Consider the following sentence for illustration:

(9) Wafula a-a-lim-a buulime
Wafula SM-Tns-plough-fv CL14land
*Wafula ploughed the land*

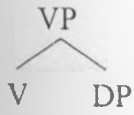
For example, Merge can operate on the lexical items '*aalima*'(α) and '*buulime*' (β) to give '*aalima buulime*' (α, β). Note that the phrase '*aalima buulime*' behaves more like the verb '*aalima*' than like the noun '*buulime*'. That is, wherever we can put the verb '*buulime*' we can also put the phrase '*aalima buulime*'.

So, we identify the phrase with a label. In this case, the label is '*aalima*' since the phrase acts as a verb. Conventionally, this phrase is a verb phrase, or VP. (8) can therefore take the form in (10) below:

**(10)** Merge (aalima, buulime) → {aalima, {aalima, buulime}}

Again, since we already know that 'aalima' is a verb (V), and 'buulime' is a noun phrase (NP) or (DP). A more general representation will then be as shown in (11).

**(11)**

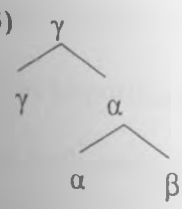



The next step will be to merge the external argument, the DP *Wafula* (γ) with the structure already built in line with similar thinking represented abstractly in (12)

**(12)** Merge (γ, {α, {α, β}}) → {γ, {γ, {α, {α, β}}}}

This corresponds to the structure in (13) below:

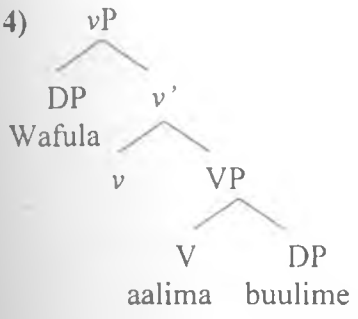
**(13)**




UNIVERSITY OF NAIROBI
LIBRARY

Here, γ corresponds to a general label specifying the identity of the whole structure. In minimalist thinking, this is a kind of verb phrase (known as light *v*P), a functional category present only to licence the external argument and implement movement. Its specifier is usually the external argument. This line of argument is especially appropriate for morphological languages such as LuBukusu where the subject DP is optional implying that a structure x consisting of the subject and a verb phrase is the same as a structure y consisting of the verb phrase only. In other words, x = y. The structure of the tree after merging the external argument is represented in (14) below.

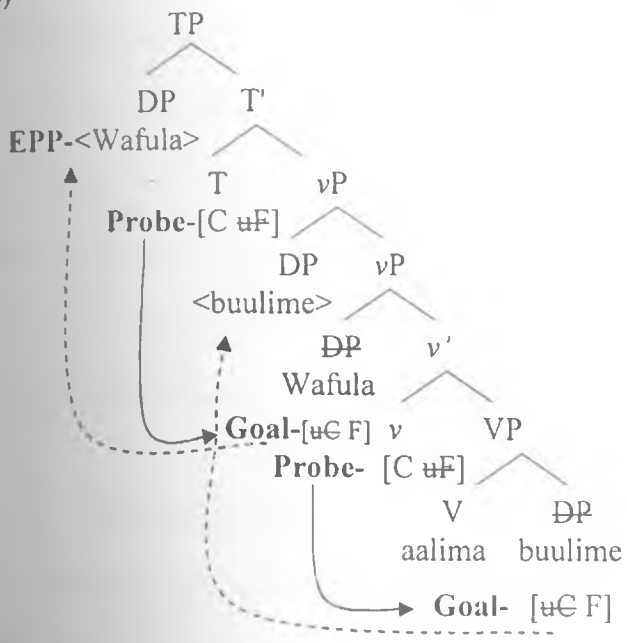
(14)




Notice, that the structure has two forms of VP, vP and VP, in what Chomsky calls a VP shell whose significance will become known shortly as we implement the operation 'move'. However, this kind of representation is overly simplified: Nothing is said about the types of features that each of the categories in the computation has that leads them to behave in the ways that they do. In other words, what are the licensing conditions for the operations hitherto stated? And more crucially, how is the whole computation procedure implemented to reflect both language specific and universal characteristics? In part to address these important questions, we apply the operation 'Move' which requires a probe and a goal both with feature specifications that must be checked in the course of the derivation. The v has interpretable case (C) features but uninterpretable phi features (uF)

(formal features such as number and person), it therefore acts as a probe that searches down the tree for a matching goal. This goal is found in the object DP which attracts the v by virtue having uC and F features. The two match and the uninterpretable features are deleted. The object DP moves to spec vP. The merging of the external argument which also has uC and F triggers another probe-goal derivation which is implemented by T that has C and uF. T is the probe that searches for a goal. The goal is found in the external argument. Matching then takes place, and the DP moves to Spec TP because of an EPP feature. The computation is finally captured in the following structure:

(15)




Note that the arrow indicates the direction of movement where the probe first searches the goal and in turn the goal moves out of the position of merge to a position where its uninterpretable features become interpretable. The analysis of anaphoric relations in LuBukusu follows this procedure and is shown in chapters four and five.

## 1.10 Literature Review

Most of the literature available on anaphora has developed in the last three decades. In this section I review the literature on anaphora both within and outside the generative framework. In addition, I focus on aspects of Bantu morphology and syntax with specific attention to anaphoric elements in order to draw comparisons with LuBukusu. Finally, I examine varied linguistic works in LuBukusu, to put this study in perspective.

### 1.10.1 Anaphora

As noted in section 1.7, anaphora generally refers to referentially dependent expressions in natural language which contribute their meaning by identifying another expression to give them their semantic value. In the literature, studies abound on the nature of anaphora cross-linguistically. From these, two broad categories emerge: Discourse anaphora and syntactic anaphora. The latter has been dealt with extensively within the generative framework, while the former is less unified. I attempt a formalized review of literature in the two categories next.

#### 1.10.1.1 Anaphora in the Generative Framework

In this section, I outline several theories that have been formulated to account for binding phenomena in the generative framework. These include; a) Binding Theory of GB, Chomsky (1981; 1986a); b) Long Distance Anaphora Koster & Reuland 1991; c) Predicate and Chains, Reinhart & Reuland (1993), and d) Koster (1993). All these attempt to explain the binding domain and differences in this domain across languages.

The Binding theory was formally introduced in Chomsky (1981), and its major achievement

was the formulation of three well-known Binding principles dealing with the distribution of anaphors, pronominals and R-expressions. Consequently, the principles were labeled, A, B and C respectively. Principle A stated that an anaphor must be bound in its governing category; principle B set a restriction on pronouns requiring them to be free in their governing categories. Lastly R-expressions were expected to be free everywhere.[14]

However, the binding principles were shown in a variety of languages to be questionable. As Gelderen (2000) notes, such variations exist in the famous `snake'-sentences where the choice between a pronoun and an anaphor in a PP complement construction varies greatly from language to language, and even between pronoun type in any given language. According to de Jong (1996) for Romance, there are more languages which are problematic. Ideally, the governing category for an element should not be so different for different languages.

From this review, it is already clear that the definition of governing category for anaphoric elements forms a significant point of departure for studies in anaphora, including this thesis. Whereas the binding theory deals with the straight forward cases occurring in a minimal clause, nothing is said about forms such as Agr-*eene* in LuBukusu that are conditioned by both syntactic and semantic/pragmatic factors. Further still, data that supports the principles is mainly derived from Romance languages. This study expands this data base from an agglutinating language to test the validity of BT principles.

Faced with the problem of defining an all inclusive governing category as the relevant domain for anaphoric binding, a number of alternative approaches were developed mainly in

[14] Before the inclusion of SUBJECT in the domain, the Specified Subject Condition and Tensed Sentence Condition (Chomsky 1973) were used to prevent the generation of sentences with unacceptable binding relations

the early 1990s. Most notable are perhaps Rizzi (1990a), and Manzini, (1992). Rizzi's proposal is one of the most influential and is popularly known as Relativized Minimality. This proposal breaks away from the earlier rigid formulation of minimality especially in relation to the type of governor. Apart from a c-command requirement, binding relations are also controlled by a similar referential index.

Generally, the elements used to illustrate the minimal conditions are mainly empty categories in pro or PRO positions, together with Wh- movement, both in complement or adjunct positions. Still this does not account for coindexation across finite clauses in morphological languages such as LuBukusu where the SM intervenes regardless of the distance between the antecedent and a coreferential argument. This study bridges this gap.

To account for the `snake'-sentences, mentioned above, different types of solutions have been proposed. Reinhart & Reuland (1991) argue that Binding Theory should be formulated as a condition on predicates (the verb and its arguments) rather than as a condition on anaphors and pronouns. Binding relations are therefore such that if two arguments of a predicate are coindexed then the predicate is reflexive with either a lexically reflexive verb or a SELF-anaphor argument.

Reinhart & Reuland's definition differs from the three principles of BT in that the governing category, i.e. the binding domain, is reformulated as the arguments of a predicate and the predicate itself. In addition, Reinhart & Reuland claim that there is a Chain Condition that allows pronouns to be used anaphorically if they are not fully marked for Case and phi-features. This is a quite significant contribution especially in enhancing the understanding of the binding domain in terms of an argument structure. Nevertheless, the theory does not

account for anaphoric affixes such as the RFM/ RCM in languages like LuBukusu which are part of the argument structure of a predicate. Further still, it is assumed that SELF- forms are always arguments. In LuBukusu, these forms, characterized as Agr-*eene* are anaphoric but may occur in non-argument positions, which renders some of the generalizations redundant. This study opens up new ways of analyzing anaphoric relations in previously understudied languages.

As was stated in the preceeding sections, perhaps the most significant shift within generative grammar occurred in the mid 1990s with the advent of the Minimalist Program (Chomsky 1993, 1995) which led to a significant change in GB based theories including the Binding theory. I mainly review the most influential minimalist literature in anaphora, next.

Koster (1993) reformulates the notion of governing category in Minimalist terms and crucially uses Case checking. He argues that morphologically marked anaphors are strong and must be checked with AGR (eement). To him, languages differ as to where the feature is located. If it occurs with AGRs, the position responsible for subject agreement, non-argument pronouns cannot function as anaphors; if it occurs with AGRo, the position responsible for object agreement, as in English, non-argument pronouns function anaphorically. This argument is crucial as it explains the distribution and subsequent interpretation of argument and non-argument anaphoric elements, and it partially contributes to the characterization of Agr-*eene* in LuBukusu as an adjunct.

Thus, both Reinhart & Reuland and Koster argue that domains can vary in terms of whether or not adjuncts are included and that inherent Case marking enables a pronoun to serve as an anaphor. Since the presence of inherent Case varies from language to language, pronouns

function anaphorically in some but not in other languages.

Another minimalist approach to Binding is provided by Burzio (1996). He argues that the antecedent is important and that anaphora is a kind of agreement between the anaphor and the Subject/Inflection complex. This claim is also made in Reuland (2001, 2005). If verbal agreement is strong (as in many Indo-European languages), pronominal reflexives are less likely than if it is weak (as in East Asian languages). This explains why pronominal free NPs are rare in LuBukusu because of the strong agreement.

Safir (2004) on his part introduces a slightly new approach in dealing with anaphora. To him anaphora is a direct consequence of a set of competitive theories. He differentiates between dependent identity and coreference. His contention is that the former is what is represented in the syntax and makes predictions to the absence of coreference (not blocked identity). Because of the distinction, indices as a basic indicator of coconstrual are eliminated from the syntax in conformity with minimalist thinking as captured in the Inclusiveness condition. A separate principle known as FTIP (Form to interpretation Principle) is then formulated to account for the failure of dependant identity or non-coreference. The basic idea behind this principle is that in order to achieve dependent identity, a form selected as the anaphor must be the most dependent form when put in competition with other forms.

Unlike other competitive theories, the FTIP is first situated in MP and relates syntax and semantics. It also predicts the complementarity between pronouns and anaphors without reference to the BT principles.

In addition, Safir distinguishes between dependent interpretations such as codependent

covaluation, dependent identity and indistinctness readings and subsequently concludes that the scale of dependence is relative to such types. Another crucial distinction is made between anaphors (subject to principle A) and UD forms, subject to specified logophoric conditions. Although this study does not adopt wholly Safir's approach, certain ideas are borrowed especially those relating to the characterization of Agr-*eene* and the phrasal reciprocal as UD forms.

On the representation of sentences, Chomsky (1992: 173), based on Pollock (1989) and Chomsky (1989), i.e. 'early' Minimalism, argues that all Case is checked in a Spec-Head relationship. For this purpose, several Functional categories, such as AGRs and AGRo, are introduced. NPs move to the Specifier and verbs to the Head positions. Nominative Case is checked against AGRs and objective against AGRo. Verbal agreement is checked in a Head-Head relationship between V and AGR after the verb incorporates into the AGR Head. These arguments form the basis for the representation of LuBukusu anaphoric elements developed in this study. However because of the congestion imposed by AGR categories, I assume a more unified approach where multi vP categories, in a VP shell, are assumed with the v head checking all the VP internal features including those of the subject and object. Chomsky (2000, 2001 a & b) also assumes such an analysis in his popular Derivation by Phase model (DbP). To him, binding conditions minimally occur within a phase, identified as a vP or CP containing both a probe and a goal capable of checking the relevant features for the reflexive reading to be achieved.

Heinat (2006) on his part uses minimalist views to draw a distinction between pronouns and anaphors. To him, both lack a referential feature R, but unlike the pronouns, anaphors

have the R feature valued within the syntactic context. This means that instead of positing separate principles for pronouns and anaphors, the two are unified under feature checking. The reformulation of the BT principles that I propose in this study is founded on grounds established in narrow syntax, hence eliminating indices and restricting the interpretation of anaphoric elements to phases. In this regard, the requirements of BT principle A are part of the syntactic computation, which, within narrow syntax, does not distinguish between pronouns and anaphors.

### 1.10.1.2 Long Distance Anaphora

For a long time studies on anaphora focused on locality conditions on anaphors across languages, with the main focus being on the extent at which the cross- linguistic anaphoric systems conformed to Principle A of GB. It was however soon noted that in some languages, anaphors can take antecedents outside their local domain in what is popularly known as Longdistance anaphora. In the literature, two texts stand out in their exhaustive handling of LDA. These include; Koster and Reuland (1991), and Cole et al (2001). I highlight some of their major findings next, and how they bear on the present study.

As a result of the workshop on LDA hosted by the Department of Linguistics of Groningen University in June 1987, a book edited by Koster and Reuland was published in 1991. Among other issues, the major themes discussed were aimed at answering the following fundamental questions: i) what is the nature of the domain restriction on long distance reflexives? ii) what sorts of logophoric or other discourse conditions are necessary for longdistance anaphora?

As a follow up to the Groningen workshop, another workshop was held in Cornell in 1997,

and the results were published in Cole et al (2001). The major differences between the two texts were the languges of focus. Whereas the former mainly dealt with European languages especially Italian and Scandnavian languages, the latter focused on Asian languages such as Mandarin, Hindu- Urdu, Kannada, Riau Indonesian, and others.

Most important however is that both volumes share predominant themes, which among other things include: i) The difference between pronouns and bound anaphors seen in the light of the domain of interpretation. Unlike pronouns, Bound anaphors were generally assumed to require a binding relationship with a c-commanding subject, allowed extra-sentential antecedent, and favoured both a sloppy and strict reading under VP ellipsis; ii) The role of discourse in relation to syntax in licensing LDA, was also of great concern. Such discourse conditions are popularly known as logophoric conditions and they affect the choice of antecedent mainly for bound anaphors, but never pronominals; iii) The concern with LDA characteristics which include monomorphemicity, c-commanding subjects and blocking effect, all affecting only bound anaphors.

For most of the contributors in the two texts, the starting point for theoretical explanations is Chomsky's BT. To them, it is suitable especially in the characterization of bound anaphors in local environments. In addition they favour a semantic analysis along the lines of Reuland (1993). They also posit a movement procedure for anaphors to a position near their potential antecedents.

Whereas LuBukusu does not diplay classical features of LDA, both Agr-eene and the affixal markers can refer to subjects in a higher clause in a form of longdistance anaphora necessitated by a rich agreement system. I extend the same arguments to LuBukusu, mainly

chapter three, in order to determine the true nature of long distance anaphoric relations in this language. This concern is also unique as it is concerned with an African language hitherto not a basis for anaphora theories. As it were, this study breaks new ground in the field of anaphora that previously concentrated primarily on Asian and European languages.

### 1.10.1.3 Discourse Anaphora

Discourse anaphora mainly deals with coreference beyond the boundaries of a sentence, and is therefore not subject to the traditional BT. Several studies have been done in this area with varied results. Partee (1984), for example identifies what he calls temporal anaphora which mainly applies to pronouns referring back to time NPs in separate sentences by use of expressions such as then, when, before, while, and many others. In addition, null anaphora is also discussed and is seen mainly in the use of demonstratives, and epithetics.

In dealing with the computation of discourse anaphora, Cooper and Parsons (1976) posit an assignment function which designates a referential value for any free variables within its domain. In case similar reference is assigned to an element as an earlier one, then coreference is achieved. This sort of analysis is similar to that of Evans (1980) who argues for a category of pronouns called E-Type pronouns which pick out a specific entity in the discourse. Grosz, Joshi and Weinstein (1995) deal with pragmatic anaphora where ones own knowledge of the world provides the missing link necessary for the interpretation of expressions as coreferential.

Notice that in all these studies, the conditions for coreference are freer and are not purely controlled by syntactic factors, and hence fall outside the scope of the study. Some of the semantic factors may however be important in the description of extra-sentential anatcedents

for anaphoric elements in LuBukusu..

### 1.10.2 Studies in Bantu Morphosynatax

Studies on the morphosyntactic constraints in Bantu have included ideas inspired by the "Mirror Principle" advanced by Mark Baker (1985, 1988), suggestions about morphological template, and proposals that perhaps thematic information is implicated in the order of the verb extensions (for relevant discussion, see Hyman 1991, 2003 Hyman and Mchombo 1992; Mchombo 2004; Ngunga 1997; Sibanda, 2004).

With regard to the Mirror Principle the suggestion is that "morphological derivations must directly reflect syntactic derivations (and vice versa)" (Baker 1985:375). This approach ties morphological structure to syntactic derivation, probably determining semantic scope too. These ideas are greatly utilized in the representation of anaphoric elements that are licensed by verb extensions.

The idea of a morphological template that fixes the preferred order of the morphemes on the basis of principles independent of syntactic derivation or semantic composition can be seen from the work of Hyman (1991, 2003). Based on studies of various African languages Hyman has noted a recurrence of the order causative, applicative, reciprocal, passive (CARP) for the verbal extensions or suffixes. The suggestion is that in the absence of over-riding factors, this is generally the preferred order of those morphemes.

On the nature of prefixes in Bantu, a lot of research has concentrated on the status of the object marker, and other affixes. The question is whether they are just grammatical agreement markers or they are pronominal arguments of the verb participating in the

verb's subcategorization structure. Recent studies on the former characterization include Diercks (2006) and Riedel (2009). They content that because the OM largely co-occurs with free NPs, in several languages such as Kiswahili, in the same way as SMs, they should be characterized as agreement markers rather than arguments.

The most influential work for the argument in support of OM as a pronominal argument comes from Bresnan and Mchombo (1987) (henceforth B&M) on Chichewa (a Bantu language of East and Central Africa). They work within the LFG approach. Similar arguments are presented in a Principles and Parameters approach notably by Jelinek (1984), and Baker (1996). The basic ideas are generally similar. B&M, for example, argue that because the SM is obligatory in most Bantu languages including Chichewa, it is used as an agreement marker. On the other hand, since the OM is largely in complementary distribution with a lexical NP, it functions as an argument.

Corbett (2006) further states that the canonical diagnosis for syntactic agreement as opposed to pronominal argument is the need to be obligatory and context free. These sentiments are also supported by Baker (1988b), albeit from a theoretical perspective. The analysis of SM as a grammatical agreement affix and the OM as an argument is a direct consequence of this diagnosis.

In this study, I use similar arguments for the characterization of the SM in LuBukusu. I conclude that it is not an incorporated pronoun and hence not an argument the way the OM is.[15] These views are supported by Diercks (2010) and Diercks and Sikuku (in

[15] The diagnosis described does not apply to the OM in a straight forward manner as it does with the SM. Its categorization depends on a host of other factors such as the presence of the free lexical NP, semantic properties of the

preparation)

### 1.10.3 Studies in LuBukusu

I have noted a growing literature on the general linguistic properties of LuBukusu, but none of the studies deals with anaphora, more so within a minimalist perspective. In phonology, most notable studies include Austen (1974a), De Blois (1975), Mutonyi (1992, 1996, 2000), and partly, Wasike (2004). Morphology and syntax studies include Wasike (1992, 2007) on the LuBukusu simple sentence and aspects of A-bar movement and left periphery, Sikuku (1998, forthcoming) on anaphora and issues of lexical reduction, respectively, Bell (2004) on negation, Khaemba (2008) on the LuBukusu *pro*-drop parameter, Diercks (2010) on issues related to agreement with subjects, and Diercks and Sikuku (in preparation) on Object marking in LuBukusu.

The obvious implication of this review is the need for more studies on different aspects of LuBukusu that will lead to a better understanding of the language and also provide a basis for testing some of the theoretical frameworks against new data from understudied languages.

### 1.10.4 Conclusion

From this review, we have noted the following:

a) The definition of governing category for anaphoric elements as shown in the Generative literature, forms a significant point of departure for studies in anaphora, including this

---

noun such as animacy and the syntactic environment. In Kiswahili, for example, the OM is only obligatory with animate objects especially humans (Marten et al 2007). Riedel (2009) reports similar arguments for Sambaa (a Bantu language spoken in Tanzania).

thesis. Whereas the binding theory deals with the straight forward cases occurring in a minimal clause, nothing is said about forms such as Agr-*eene* in LuBukusu that are conditioned by both syntactic and semantic/pragmatic factors. Further still, data that supports the principles is mainly derived from Romance languages. This study expands this data base from an agglutinating language to test the validity of BT principles.

b) Reinhart & Reuland claim that there is a Chain Condition that allows pronouns to be used anaphorically if they are not fully marked for Case and phi-features. This is a quite significant contribution especially in enhancing the understanding of the binding domain in terms of an argument structure. Nevertheless, the theory does not sufficiently address anaphoric affixes such as the RFM/ RCM in languages like LuBukusu which are part of the argument structure of a predicate.

c) Further still, it is assumed that SELF- forms are always arguments. In LuBukusu, these forms, characterized as Agr-*eene* are anaphoric but may occur in non-argument positions, which renders some of the generalizations redundant. This study becomes theoretically relevant in so far as it opens up new ways of analyzing anaphoric relations in previously understudied languages.

d) Koster (1993) reformulates the notion of governing category in Minimalist terms and crucially uses Case checking. This argument is crucial as it explains the distribution and subsequent interpretation of argument and non-argument anaphoric elements, and it partially contributes to the characterization of Agr-*eene* in LuBukusu as an adjunct.

e) Another crucial distinction is made by Safir (2004) between anaphors (subject to principle A) and UD forms, subject to specified logophoric conditions. Although this study does not adopt wholly Safir's approach, certain ideas are borrowed especially

those relating to the characterization of Agr-*eene* and the phrasal reciprocal as UD forms.

f) The reformulation of the BT principles that I propose in this study is founded on grounds established by Heinat (2006) in narrow syntax, that succeed in eliminating indices and restricting the interpretation of anaphoric elements to phases. In this regard, the requirements of BT principle A are part of the syntactic computation, which, within narrow syntax, does not distinguish between pronouns and anaphors.

g) Some of the semantic factors established in the literature on discourse anaphora are important in the description of extra-sentential anatcedents for anaphoric elements in LuBukusu.

# Chapter Two

# Anaphoric Elements in LuBukusu

## 2.0 Introduction

What can generally be called Bantu anaphora usually distinguishes two broad categories of anaphoric elements. They can either be bound morphemes or independent lexical forms all corresponding to typical NP positions. Falling in the former category are reflexive (RFM) and reciprocal markers (RCM), while the latter contains a distinct class of Agr-forms that are either reflexive or reciprocal.

This chapter focuses on the identification and categorization of LuBukusu anaphoric elements in a purely descriptive sense. Four such elements are identified, with each discussed in a different section of this chapter according to morphological realization and syntactic properties. The sections are organized as follows: Section 2.1 examines the reflexive marker giving its underlying form and the inherent variations mainly motivated by adjacent phonemes of other morphemes. The syntactic relations that the RFM enters into with other elements in a syntactic configuration are also discussed with great emphasis on the licensing conditions, constituent ordering, grammatical marking, and grammatical functions and valence operations. In the same vein, section 2.2 deals with the reciprocal marker, section 2.3 with the lexical *Agr-eene* form, section 2.4 with the phrasal reciprocal usually represented as *Agr-ene* on *Agr-eene*. The conclusions are given in section 2.5.

## 2.1 The Reflexive Marker

The reflexive marker is perhaps one of the least understood forms occurring on the verb morphology. Most of the literature on aspects of LuBukusu morphology (notably Mutonyi 2000 and Wasike 2007) do not even mention it, perhaps with the believe that it should be analyzed as part of the object agreement morphology. This is true to some extend especially when one considers the fact that in monotransitive constructions the OM and RFM are in complementary distribution implying that they occupy the same position and that their roles are similar.[16] A closer look however reveals a different picture: The RFM does not participate in agreement the way object markers do; in multitransitive constructions, both the RFM and OM can cooccur. In this section I shed light on what is the true nature of the LuBukusu RFM. In order to do this effectively, I examine it according to both the morphological and the syntactic properties.

### 2.1.1 Morphological Properties

The realization of the RFM in LuBukusu is through prefixation reminiscent of many Bantu languages (see data on Digo, CiNsenga, and Kiswahili elsewhere in this section). The forms *–i-* and *–e-*, are used to represent this morpheme, sometimes with a preceeding glide in phonologically defined environments. Details of this kind of pattern together with the question concerning the specific underlying form will be answered after a consideration of varied data.[17] Consider the following examples:

---

[16] The complementarity only occurs in transitive verbs that take the RFM and RCM as objects and hence cannot allow another object.

[17] The variation in the data is based on aspects such as number, person, tense and aspect marking and the nature of the verb stem all aimed at capturing a typical context within which the RFM occurs and hence correctly accounting for its realization.

(1a) Yohana a-a-e-bon-a
John SM-Tns-RFM-saw-fv
*John saw himself*

(1b) Mu-a-e- bon-a
2plSM-Tns-RFM-saw-fv
*You saw yourselves*

(1c) Khu-a-e- bon-a
1plSM-Tns-RFM-saw-fv
*We saw ourselves*

Notice that the RFM which attaches immediately before the verb stem in all the examples is morphologically realized as *–e-*. Before I discuss the motivation for this kind of realization. Let us consider the alternative data below:

**(2a)** N-i- siing-ang-a
1sgSM-RFM-wash-Asp-fv
*I wash myself*

**(2b)** W- i- siing-ang-a
2sgSM-RFM-wash-Asp-fv
*You wash yourself*

**(2c)** Ba-e- siing-ang-a
3plSM-RFM-wash-Asp-fv
*They wash themselves*

The RFM is realized as *–i-* in (2a) and (2b), but as *–e-* in (2c). Further, the RFM does not vary according to person or even participate in noun class agreement the way object markers do. Nevertheless, sentences (1) and (2) are evidence that the affix varies considerably depending on the phonological environment of adjacent morphemes, which include the SM, or even the tense and aspect markers. This strategy results in phonotactic combinations that are acceptable in the language. In this case, the modification is caused by the unacceptable possibility of vowel clusters or diphthongs. In order to understand

such modifications, I will take as a starting point the position that the RFM is underlyingly represented as *–i-* with *–e-* as its allomorph. Such a position is informed by both historical and empirical evidence. The former relates to a comparison with other Bantu languages whereas the later deals purely with language internal factors. I examine the two factors next.

### 2.1.1.1 Phonological Evidence

The most straight forward language factors that support the *–i-* hypothesis comes form the context where a consonant precedes the RFM. Imperative formation is a crucial factor mainly because the imperative forms of verbs normally represent the basic structure without any modifications. Consider the data in (3) below:

**(3a)** i-bon-a
RFM-see-fv
*See yourself*

**(3b)** i-siim-a
RFM-like-fv
*Like yourself*

**(3c)** i-an-a
RFM-give-fv
*Give yourself (out)*

The realization of the RFM is *–i-* in such imperatives, and because the form is not preceded by any other phoneme, then it is the basic reflexive realization. Notice that (3c) has a verb beginning with *–a-* yet the structure of the RFM remains unaffected.

Other supporting evidence for the *i-* hypothesis comes from contexts with consonants preceding the RFM. This is quite significant because in typical language situations, vowels coming before any morpheme are much more likely to affect the adjacent

phonemes. Typical examples are shown in (2a) and (2b) above, where the 1st and 2nd person singular subject markers preceed the RFM respectively. In both cases, the RFM is basically *–i-*. The situation is however quite different in (4) where there is an intervening vowel.

**(4a)** N-a-e- siing-ang-a
1sgSM-Tns-RFM-wash-Asp-fv
*I washed myself (habitually)*

**(4b)** W-a- e- siing-ang-a
2sgSM-Tns-RFM-wash-Asp-fv
*You wash yourself (habitually)*

The past tense morpheme represented by *–a-* before the RFM lowers the representation of the RFM to *–e-*. Thus *–e-* is marked in special contexts especially those with lower vowels.

### 2.1.1.2 Evidence from other Bantu Languages

Evidence from other Bantu languages indicates a representation that is much closer to *–i-* than *–e-*. Consider the example in (5) below[18]

**(5a)** Etta a-li-**ji**-pig-a — **Kiswahili**
Etta SM-Tns-RFM-beat-fv
*Etta beat herself*

**(5b)** Etta o-**zi**-nyany-a — **CiNsenga**
Etta 3sg-RFM-hate-fv
*Etta hates herself*

**(5c)** Anache a-**dzi**-pig-a — **Digo**
CL2child SM-RFM-hit-fv
*(The) children hit themselves*

[18] Except for Kiswahili and CiNsenga, data from Digo (a mijikenda language spoken at the East African coast) was directly elicited from a native speaker. Kiswahili data is derived from my own judgement as a speaker of the language, and verified by Prof. Abdulaziz (my supervisor and a native speaker of the language). CiNsenga (A language spoken in eastern Zambia and western central Malawi) is derived from 'The CiNsenga file of the Afranaph project' at www.africanaphora.rutgers.edu.

If *–i-* is the underlying representation then any other variant realization is conditioned by specific factors. I argue that *–e-* occurs as a result of phonological conditioning motivated by the nature of the adjacent sounds on morphemes such as tense, aspect and subject marking. Such an environment is summarized in the paradigm below that helps to determine a general phonological trend described at the end, and which accounts for the form of the RFM.

| Tense | TAM | RFM | V.Stem | Example |
|---|---|---|---|---|
| **Immediate future** | La | e | Siima | Khu-le-e-siim-a<br>We'll love ourselves |
| **Intermediate future** | Kha | e | Siima | O-khe-e-siim-e<br>you'll love yoursef |
| **Remote future** | Li | i | Siima | A-li-i-siim-a<br>he'll love himself |
| **Immediate past** | kha | e | Lola | Na-khe-e-lol-a<br>I have checked myself<br>Khwa-khe-e-lol-a<br>we've checked ourselves |
| **Remote past** | A | e | Lola | E-e-lol-a<br>he checked himself |
| **Persistive Aspect** | Si | i/ye | Lola | E-si-i-lol-a (esyelola)<br>I am still checking myself<br>Khu-si-i-lol-a (khusyelola)<br>we are still checking ourselves |
| **Stative/ Completive Aspect** | A | e | Lola | We-e-lol-a (wayelola)<br>you'(ve) still check(ed) yourself<br>Mwe-e-lol-a (mwayeelola)<br>you'(ve) still check(ed) yourselves |

**Table 4 Tense/Aspect Prefixes in LuBukusu**

The analysis highlighted reveals that the form of the RFM is largely determined by phonological conditions normally triggered by a sound adjacent to the RFM, and which

usually belongs to a tense, aspect or person morpheme. Notice that the RFM is realized as *–i-, -e-* or *ye* in specified environments. Such contexts can be represented by the following simple generalization:

**Generalization 4**

a)...[*a*]/*RFM* →[ *e*]

b) .../*RFM* → *i/ G/e*

These rules are a conjunction of the sounds that make up the context in which the RFM is realized. In a), such a context is made up of the morpheme immediately preceding the RFM and whose final vowel is *–a-*, consequently leading to the realization of the RFM as *–e-*. The idea is simple: Since [a] is an open vowel, it lowers the underlying *–i-* to *–e-*, and also succeeds in changing to *–e-* which is an intermediate height in between the low vowel [a] and the high vowel [i].

The remaining context is specified in b). The position preceding the RFM has an unspecified value i.e. it can be any of the other four vowels: *e, i, o,* or *u* in which case, *-i-* is selected or *–e-* so long as it is preceded by a glide or a consonant.

These are, however, not the only variant forms of the reflexive. The infinitive clause and the 'self' nominals motivate different realizations. I turn to these below.

### 2.1.1.2 The Reflexive Marker in Infinitives

LuBukusu marks the infinitive by means of the prefix *khu-* . The form is also quite related to the infinitive marking in most Bantu languages. Kiswahili, for example, uses *ku-*. In English, it loosely translates to the infinitive *to*. Again, just like in other languages, Bukusu infinitives share certain characteristics with nouns such as having

double prefixes, in a few cases, and an agreement system reminiscent of other Bantu noun classes (Noun classes are exemplified in details in chapter one). In fact, the infinitive is analyzed as a class 15 gerundial noun. Consider:

(6) Khu- khu-inyukh-a khu-eewe khu-li khu-a kalaa
Prprfx-prfx-stand-fv Agr-his Agr-is Agr-of slow
*His (way of) standing is slow*

Here, the infinitive form displays both functional and structural features of nouns: It occurs as the head of a NP which in turn is the subject of the matrix clause. The form also has double prefixes which are marked as agreement affixes on the possessive modifier within the subject NP, on the copular verb, and on the preposition. Note that double prefixation in infinitives is conditioned by phonological factors i.e. it occurs in verbs whose stems have an initial vowel e.g. *–iikhala* (sit), *ikalaho* (close) which then become the infinitives *khukhwiikhala* (to sit) and *khukhwiikalaho* (to close) respectively.[19]

On the other hand, the infinitive construction could still be analyzed as a verb heading an IP, hence a clause. In addition it consists of the verbal features conceptualised in the following rule:

**Generalization 5**

The infinitive is verbal because It takes certain aspectual and mode markers, but not tense, and it can always co-occur with other conventional verbal affixes such as the object marker, applicative, reflexive, reciprocal, and causative.

The example below helps to illustrate this.

[19] There is debate concerning the best way to represent the mixed properties of infinitives: one possibility is to rely on phrase structures characterized by category switch (Pullum 1991), possibly enriched with functional structure sharing (Bresnan 1997); another one is to set up a new category (head value), which is neither verb nor noun, but inherits from less specified nominal and verbal categories at the same time (Malouf 2000).

(7) Khu-mu-siim-is-il-ang-a kamalwa khu-a-mu-reer-er-e
Inftv-OM-like-Caus-Appl-Asp-fv CL6beer Agr-Tns-OM-bring-Asp-fv
butinyu
CL14problem
*Pleasing him with beer brought him problems*

Notice that the aspects described in the rule are actualized in this example. However, research is yet to establish whether these characteristics are common to Bantu languages or not. Creissels and Godard (2005) discuss Tswana infinitives and note similar features, only differing with respect to the specific affixes allowed within their morphological structures, the most common being the tense marker and object marker. Mugane (2003) gives a contrasting analysis of Kikuyu infinitives concluding that they have an impoverished morphology. Whatever the analysis, we conclude here that the form of the infinitive in LuBukusu is formed according to the generalization in 6 below:

**Generalization 6**

[*khu*+...*V*...] Inf.

This simple rule interprets an infinitive as a structure consisting of *khu-* appearing in the initial position of the category (whose identity is, depending on several factors (see Muller 2005), either a NP or a clause). This is then followed by optional morphological material that give way to the verb stem that is in turn followed by more optional suffixes. The type and nature of the affixes is uniquely specified from language to language. Such suffixes could be both inflectional and derivational affixes common in Bantu morphology. Following from these, one can easily predict that since an OM is possible, so is a RFM, but the two are in complementary distribution occupying a similar pre-verbal position. Consider the following examples:

(8a) Wanjala a-reeb-a Nekesa khu-khw-i-fumy-a
Wanjala SM-asked-fv Nekesa to-RFM-praise-fv
*Wanjala asked Nekesa to praise herself*

(8b) Wanjala e-ny-a khu-mu-fumy-a
Wanjala SM-wants-fv to-OM-praise-fv
*Wanjala wants to praise him*

The most important question here is: What is the inherent realization of the RFM? On the surface this appears straight forward. One would easily point out that the affix *–khwi-* is the RFM, viewed as an allomorph of the conventional *-i-*, and that it is grammatically conditioned. Let me briefly consider another line of argument that will lead to a slightly different conclusion. For the sake of argument, I take the RFM in infinitives to be similar to the normal one i.e. *-i-*. This means that the *–khw-* form attached requires justification. One explanation would be to consider it an agreement affix having the same features as the infinitive marker. This is however not plausible because agreement features should be marked on another head (lexical item), but not on the same head as is the case here. The only explanation then is one that takes this element as a case of double prefixation. Remember however that the main motivation for such double prefixes is an initial vowel on the verb. The verb *–inyukha* has an initial vowel and therefore with the addition of the infinitive *–khu* there is reduplication that results in *khukhwinyusya*. Similarly, the insertion of the RFM on an otherwise consonant initial verb such as *fumya* in (8) triggers similar reduplication to become *khukhwefumya*. I now take the RFM as an initial vowel so that whenever an infinitive is added, it will be reduplicated on the strength of the presence of the RFM vowel. (See Mutonyi (2000) for a detailed analysis of reduplication patterns in LuBukusu).

More evidence for the conclusion that the RFM in infinitives is no different from the conventional *–i-* can be found in infinitive verbs whose stem has an initial vowel as is the case in (9a). In such verbs the reduplication is already present. The addition of the RFM will yield (9b) below.

**(9a)** Khu- khu-inyusy-a khu-eewe khu-li khu-a kalaa
Prprfx-prfx-RFM-raise-fv Agr-his Agr-is Agr-of slow
*His (way of) raising is slow*

**(9b)** Khu- khu-i-inyusy-a khu-eewe khu-li khu-a kalaa
Prprfx-prfx-RFM-raise-fv Agr-his Agr-is Agr-of slow
*His (way of) raising himself is slow*

The RFM remains the underlying *–i-* that is not attached to the preceding prefix. I conclude that, in infinitive constructions, the form of the RFM does not change significantly but it may trigger reduplication of infinitive marking especially in verbs with an initial consonant.

### 2.1.1.3 The Reflexive Maker in 'Self' Nominals

Related to infinitives is a general class of deverbal nouns that are derived via affixation of a selected category of noun class affixes. In fact, the majority of noun classes participate in deverbal nominalization, as shown in table 5 below. The form that this process takes is quite complex, involving not only the addition of prefixes but also final vowels or even consonant modifications. It is also common to have a single verb undergoing several instances of nominalization depending on the affix that it takes. The most common is the creation of 'agent' and 'patient' nouns by suffixation of [i] and [e] respectively, suffixation of [o] especially in non- human classes, and the use of [fu] suffix in verbs of

'feeling'. See the noun class paradigm below derived from the verbs *teekha* (cook), *lima* (dig), and *khupa* (beat), for illustrations.

| Class | Deverbal Noun | | |
|---|---|---|---|
| | **Agent** | **Patient** | **Reflexive** |
| 1 | Omuteekhi | Omuteekhe | Omw-i-teekhe |
| 2 | Bateekhi | Bateekhe | Baba-e-teekhe |
| 3 | | Kumulime | Kumw-i-lime |
| 4 | | Kimilime | Kimi-i-lime |
| 5 | | Likhupe | Lili-i-khupe |
| 6 | | Kamakhupe | Kama-e-khupe |
| 7 | | Siteekhe | Sisi-i-teekhe |
| 8 | | Biteekhe | Bibi-i-trrkhe |
| 9 | | Endime | ?E-i-ndime |
| 10 | | Chindime | ?Chi-i-ndime |
| 11 | | Lukhupe | Lulw-i-khupe |
| 12 | Khateekhi | Khateekhe | Khakha-e-teekhe |
| 14 | Bulimi | Bulime | Bubw-i-lime |

**Table 5: Selected Verb Nominalization Patterns in LuBukusu**

Notice that in this data, nominalization with a reflexive is quite productive. In fact it occurs in all noun classes that also take a patient nominal. This contrasts sharply with the unproductive agent nominals which only occur in the human classes together with the abstract class 14. Further, this selection of deverbal nouns reveals a number of issues on deverbal nominalization in LuBukusu:

a) Nominalization of verbs is a process of prefixation and final vowel modification;

b) In transitive verbs, the final vowel modification depends on whether the noun occurs as an agent or a patient. This distinction is however only possible in noun classes that allow the referents to be agents, otherwise in the majority of the classes only the patient variant is allowed.[20]

[20] Intransitive verbs are mostly limited to animate classes mainly because they can perform actions as agents. Stative verbs are therefore excluded.

c) Like the infinitive, some of these deverbal nouns can be reflexivized. Notice that reflexivization only occurs on the patient nouns. However, unlike the infinitive, an OM is not allowed in the same position, perhaps underlying their typical nominal features as opposed to noun/ verb- like features characteristic of infinitive.

The other forms of nominalization patterns which take the suffix [fu] and [o] are shown below.

| Noun Prefix | Verb | Deverbal Noun | 'Self' Nominal |
|---|---|---|---|
| Omu | Suubila- *believe*<br><br>Rekeresya- *listen* | Omusuubilifu- *trusted one*<br><br>Omurekerefu- *Good listener* | *Omu-i-suubilifu- *self trusted one*<br>*Omu-i-rekerefu- *self listener* |
| Si | Kaba- *give* | Sikabilo- *used for giving*<br><br>Sikabo- *given talent* | Si-sye-kabilo- *used for self giving*<br>?Si-sye-kabo- *self given talent* |
| En | Teekha- *cook* | Endeekhelo- *used for cooking* | *E-i-ndeekhelo- *used for self cooking* |
| Si | | Siteekho- *wedding* | Sisi-i-teekho – *Self wed* |

**Table 6: More 'self' nominals in deverbal nouns**

This kind of nominalization is more irregular than the earlier pattern. Consequently, the formation of reflexive nominals is less productive. Notice that the nominals ending in [fu] cannot be reflexivized while those ending in [o] have mixed results. Presently, it looks plausible to conclude that the nature of the deverbalization process determines whether or not a deverbal noun can take a RFM. Based on the available evidence the conclusion in form of Rule 4 below seems plausible.

c) Like the infinitive, some of these deverbal nouns can be reflexivized. Notice that reflexivization only occurs on the patient nouns. However, unlike the infinitive, an OM is not allowed in the same position, perhaps underlying their typical nominal features as opposed to noun/ verb- like features characteristic of infinitive.

The other forms of nominalization patterns which take the suffix [fu] and [o] are shown below.

| Noun Prefix | Verb | Deverbal Noun | 'Self' Nominal |
|---|---|---|---|
| Omu | Suubila- *believe*<br><br>Rekeresya- *listen* | Omusuubilifu- *trusted one*<br><br>Omurekerefu- *Good listener* | *Omu-i-suubilifu- *self trusted one*<br>*Omu-i-rekerefu- *self listener* |
| Si | Kaba- *give* | Sikabilo- *used for giving*<br><br>Sikabo- *given talent* | Si-sye-kabilo- *used for self giving*<br>?Si-sye-kabo- *self given talent* |
| En | Teekha-*cook* | Endeekhelo- *used for cooking* | *E-i-ndeekhelo- *used for self cooking* |
| Si | | Siteekho- *wedding* | Sisi-i-teekho – *Self wed* |

**Table 6: More 'self' nominals in deverbal nouns**

This kind of nominalization is more irregular than the earlier pattern. Consequently, the formation of reflexive nominals is less productive. Notice that the nominals ending in [fu] cannot be reflexivized while those ending in [o] have mixed results. Presently, it looks plausible to conclude that the nature of the deverbalization process determines whether or not a deverbal noun can take a RFM. Based on the available evidence the conclusion in form of Rule 4 below seems plausible.

**Generalization 7**

The RFM is possible in all deverbal nouns ending in a patient nominalizer [e], partially possible in those ending in [o], and totally excluded in those ending in agentive [i], and [fu].

In order to capture the nature of the context hosting RFM in nominals, consider the following examples:

**(10a)** Lili-i-suubila li-a Andrea li-a-belel-isy-a Maria
CL5-RFM-believe Agr-of Andrea SM-did-annoyed-Caus-fv Maria
*Andrew's self-confidence annoyed Mary*

**(10b)** Lili-i-manyisya li-a Andrea li-a-sim-isy-a Wekesa
CL5-RFM-introduction Agr-of Andrew SM-Tns-please-Caus-fv Wekesa
*Andrew's self introduction pleased Wekesa*

Notice that the first of the two *–li-* in the subjects of the two sentences stands for the noun class marker. However the second is a reduplication of the first, motivated by the presence of the –i- RFM. In (10a), the basic form of the subject – *Lisuubila,* requires a syntactic category to express the role of agent that is obligatorily assigned by the deverbal noun. This role is correctly assigned by the head of the PP complement, *Andrea.* With the addition of the reflexive, the role is assigned to the *-i-* affix which finds its antecedent in *Andrea.* The two are therefore construed. The same thing applies to (10b). The conditions under which such construal takes place are discussed in details in chapter 4. Meanwhile, I turn to the syntactic context of the RFM.

### 2.1.2 Syntactic Properties of the Reflexive Marker

This section examines the syntactic relations occurring between the RFM vis-a-vis other elements in a grammatical structure. The relations are discussed in terms of; a) The

licensing conditions/ factors; b) constituent order; c) grammatical marking, functions such as subject/ object and valence operations.

### 2.1.2.1 Licensing Conditions

In this section, I describe the syntactic conditions that license the occurrence of the RFM in specified positions in a syntactic configuration. These include transitivity, locality and agreement.

#### 2.1.2.1.1 Transitivity, Locality, and Agreement

The RFM is licensed by verb transitivity i.e. a verb that requires two participants in the real world, whether distinct or merged. The most common idea is that if a verb can take an overt object or an OM, then that verb will also host the RFM, in which case there cannot also be an OM on the verb in question. This by extension means that both the OM and the RFM are in complementary distribution. Consider the following examples:

**(11a)** A-a-ir-a Wekesa
SM-Tns-kill-fv Wekesa
*He killed Wekesa*

**(11b)** A-a-mu-ir-a
SM-Tns-OM-kill-fv
*He killed him*

The verb here is transitive i.e. it can take either an overt object or an OM. Consequently, the addition of the RFM is possible, as seen in (12) below.

**(12)** A-a-e-ir-a
SM-Tns-RFM-kill-fv
*He killed himself*

The complementarity in distribution between the OM and RFM in transitive verbs is shown in (13) below.

(13) *A-a-mu-i-ir-a
SM-Tns-OM-RFM-kill-fv
*He killed him/ himself*

Note. however, that the complementarity in distribution is only present if a verb is mono-transitive i.e. requiring only one object. In which case, the ungrammaticality stems from an apparent competition for the same argument position between the RFM and the OM. However, if one of the valence increasing affixes such as causative is added, then the RFM and the OM can co-occur. This is seen in (14) below.

(14) A-a- ba- e- ir- isy-a
SM-Tns-OM-RFM-kill-Caus-fv
*He made them kill themselves*

Details of such co-occurrences are discussed in chapter three with the theoretical implications explained in chapter four. I briefly turn to an explanation as to why there is complementary distribution between the OM and RFM. In the literature, the explanations are quite varied. The Lexical Functional Grammar (LFG) account concerning incorporation. for example, is based on the assumption that since object markers are incorporated pronouns (Baker 1988), it follows that a verb with an OM çannot take another object. It has seemingly become intransitive or has had its valence lowered.[21] A

[21] Lexical Function Grammar (LFG) describes an utterance in terms of several distinct representations. The constituent structure (c-structure) represents the surface constituency relationships that hold between the words of an utterance, whereas the functional structure (f-structure) represents the function-argument structure of an utterance. The c-structure is a conventional tree structure with words as its terminals, and the f-structure is a recursive attribute-value structure, which resembles the hierarchical frame structures used in knowledge representation systems such as predicate, case, and agreement.

Incorporation has been used to account for such complementarity. An incorporated pronoun or pronominal inflection is a bound morpheme that specifies a complete pronominal f-structure. The functional specification of a pronoun is incorporated with the functional specifications of the stem to which the morpheme is bound (In this case the verb stem). The functional specification of a pronoun includes semantic features, binding features which constrain the range of possible syntactic antecedents and case and agreement features. The latter features are shared by lexical nominals, and include gender, person, number, animacy, and specificity features. Only pronouns carry the pronominal binding and semantic features. Because of the principle of Functional Uniqueness, an incorporated pronoun will preempt the appearance of any c-structure constituent of the same function having a conflicting semantic feature predicate value, within the functional domain of the head to which the incorporated pronoun is bound.

minimalist account on the other hand stems from a related motivation: That both the OM and RFM are pronominal elements fully specified for both interpretable and uninterpretable features. The interpretable features consist of agreement features such as person and number, in addition to a referential feature that distinguishes various types of pronominal expressions (Hyman 2000). The uninterpretable ones concern case as a semantic feature assigned by the verb. The idea behind such facts is then that since both the RFM and OM have case features that must find value by matching with the same verb, then the result will be competition for interpretation leading to ungrammaticality.

As widely acknowledged in the Binding Theory of GB, and discussed in details in chapter one, anaphoric relations in many languages are largely local i.e. they occur within the same clause (Chomsky (1981), Haegeman (1991), Black (1998)). The principles of the binding theory specify the elements whose construal is determined by some antecedent, and the requisite structural conditions for successful construal. Binding has, traditionally, involved reflexives and pronominals. The former have their construal with an antecedent resolved within an appropriately defined local domain. The latter, on the other hand, appear to be more involved in discourse structure. Earlier versions of binding theory included reciprocals in the typology of bound elements, grouped together with reflexives as requiring an antecedent within some local domain. The identification of the reflexive with the reciprocal as constituting the class of anaphors is one that continues to undergo revision in light of their grammaticalization and semantic properties. This section adds to this endevour.

As discussed in section 2.1.1 above, the LuBukusu RFM is *–i-* with

morphophonologically determined variations. It appears in the position of the OM i.e. to the left of the verb stem (VS). The OM and the VS comprise the Macrostem (Goldsmith & Sabimana 1985). In Bantu languages the OM has been analyzed as an incorporated pronominal argument anaphorically bound by an antecedent within the discourse structure (Bresnan & Mchombo 1986, 1987 (Chichewa); Dlayedwa 2002 (Xhosa); Letsholo 2003 (Ikalanga); Chimbutane 2003 (Changana); Deen 2004 (Kiswahili)). The main arguments for this kind of thinking can be summed up in the form of rules outlined and illustrated below.

**Generalization 8**

The free object NP can only occur in its base position following a verb in a clause when the OM is absent.

Consider the following:

**(15a)** Wekesa a-a-p-a omwaana
Wekesa SM-Tns-beat-fv CL1child
Wekesa beat (the) child

**(16b)** *Wekesa a-a-mu-p-a omwaana
Wekesa SM-Tns-OM-beat-fv CL1child
Wekesa beat (the) child'

Notice that when the OM and the NP *omwaana* occur together in (5b) the structure is illicit. It is however quite interesting to note that if this NP was a pronoun, then the structure would become grammatical. I argue that this is as a result of the semantic properties of pronouns allowing them to be used for emphasis unlike complete referential expressions.

**Generalization 9**

There are constraints on word order whenever the free NP is shifted to different

positions in a clause.

The following data helps to illustrate this.

**(16a)** A-a-p-a omwaana Wekesa
SM-Tns-beat-fv CL1child Wekesa
*Wekesa beat (the) child*

**(16b)** *Wekesa omwaana a-a-p-a
Wekesa CL1child SM-Tns-beat-fv
*Wekesa beat (the) child*

**(16c)** *A-a-p-a Wekesa omwaana
SM-Tns-beat-fv Wekesa CL1child
*Wekesa beat (the) child*

**(16d)** *Omwaana a-a-p-a Wekesa
Wekesa SM-Tns-beat-fv CL1child
*Wekesa beat (the) child*

Only (16a) is grammatical because of the rule that object NPs are base generated in the VP internal position and must always follow the verb. They can only appear in a different position if there is a substitute form playing this role. This form is usually the OM. The examples in (17) further help to illustrate this argument.

**Generalization 10**

If the overt NP occurs with the OM, there is significant freedom in the variation of word order.

**(17a)** A-a-mu-p-a omwaana Wekesa
SM-Tns-OM-beat-fv CL1child Wekesa
*Wekesa beat (the) child*

**(17b)** Wekesa omwaana a-a-mu-p-a
Wekesa CL1child SM-Tns-OM-beat-fv
*Wekesa beat (the) child*

**(17c)** A-a-mu-p-a Wekesa omwaana
SM-Tns-OM-beat-fv Wekesa CL1child
*Wekesa beat (the) child*

**(17d)** Omwaana, a-a-mu-p-a Wekesa
Wekesa SM-Tns-OM-beat-fv CL1child
*Wekesa beat (the) child*

Using similar arguments we note that the grammaticalization of the RFM, appearing in the position of an incorporated pronominal argument, supports its treatment as a pronominal argument whose construal is determined by principles of syntactic binding. Consider the following:

**(18a)** Baabini ba-a- e-kis-a mu- musiru
CL2witches SM-Tns-RFM-hid-fv CL18-CL3forest
The witches hid themselves in the forest

**(18b)** Baabini ba-a- ba-kis-a mu- musiru
CL2witches SM-Tns-OM-hid-fv CL18-CL3forest
The witches hid them in the forest'

**(18c)** Baabini ba-a- bool-el-a Nanyama ba-li ba-a-e-kis-a
CL2witch SM-Tns-told-Ben-fv Nanyama Agr-that SM-Tns-RFM-hid-fv
Mu-musiru
CL18-CL3forest
The witches told Nanyama that they hid themselves in the forest'

The relevant antecedent for the RFM in (18a) and (18c) appears to be the SM attached to the verb hosting the RFM. The antecedent itself is a functionally ambiguous agreement marker and an incorporated pronominal argument (Bresnan & Mchombo 1986, 1987; Mchombo 2004). The pronominal argument status of the SM underlies the apparent 'long distance' relationship between the reflexive and the constituent that determines its referential value in (18c). In other words, the subject of the matrix clause (*baabini*) is able to antecede the RFM in the lower clause via *ba-* SM hosted on the same verb as the RFM. The binding conditions here can be looked at in two ways: First, they are local and therefore subject to principle A of BT. In this case the apparent long distance binding is reduced to a local one by means of the agreeing SM. Secondly, in LuBukusu agreement

marking over long distances reduces barriers for binding hence imposing only one condition: That the antecedent of the RFM is any NP whose agreement features are marked on a local SM (SM closest to the RFM). How such issues are implemented will be dealt with in chapter four, so will be the question of what exactly serves as the antecedent; the SM or the marked subject. The story is however slightly different in reflexive nominals. Here, the antecedent licensing the RFM is the NP contained in the PP complement of the reflexivized nominal as shown below:

**(19)** Lili-i-suubila li-a Andrea li-a-belel-isy-a Maria
CL5-RFM-believe Agr-of Andrea SM-Tns-annoyed-Caus-fv Maria
*Andrew's self-confidence annoyed Mary*

Here, the prepositional complement *Andrea* is the antecedent of *-i-* .

Still in GB terms, the main principle regulating anaphoric relations is principle A of GB. This principle specifies the conditions under which binding must take place. In this case the domain of binding is the subject NP. Two significant questions arise; i) Are these really universal conditions? ii) How are the conditions reformulated within minimalist thinking? These issues are dealt with in details in chapter four.

The notion of agreement in language has been the center of focus for many studies mainly because of the crucial role it plays in licensing relationship patterns in language. In order to understand its licensing role for the RFM, I present a brief discussion on what agreement is.

Agreement involves identity of features. Based on this assumption, agreement can broadly be classified into three main categories exemplified in (20) below:

I. The grammatical subject/ verb agreement;

II. Phrasal agreement between dependants and the head;

III. The semantic pronominal agreement that is usually anaphoric and may either occur in a specified syntactic domain or can stretch over long distances in discourse and depends on non- linguistic factors such as societal beliefs and conventions.

**(20)** Chikhaafu chin-daayi chi-a-mu-siim-a
CL10cow Agr-good SM-Tns-OM-like-fv
*Good cows like him*

Subject/ verb agreement is noted where the CL10 SM occurs on the verb. Replacing this with another class affix will render the sentence ungrammatical because of the mismatch in agreement[22]. Further, agreement within a phrase is shown in the subject NP where the prefix on the adjective modifier agrees in terms of noun class with the head of the NP. Lastly, semantic agreement is realized in the OM, where the entity refered to is realized in the discourse context beyond the sentence. In addition, in an English sentence such as *The country recovered from the hardships because its/ her people are hardworking*, the choice of the pronoun depends on people's views about the gender of the referent and convention. Such notions do not affect the nature of genuinely grammatical agreement

[22] Anderson (2004: 2) suggests that the best definition of agreement is a theoretical reconstruction. To him, the theory must deal with the following pertinent issues:

a) Domain: What is the configuration within which agreement takes place?
b) Directionality: Does agreement go from one element to another, or is it just a relation? If it is directional, what is the directionality?
c) Locality: What limitations are there on the relations between agreeing elements?
d) Features: In what properties do grammatical elements agree?
e) Strictness: Under what conditions does agreement involve less than strict identity in the features involved?
f) Conflict: What happens when two principal (typically semantic and syntactic) agreements differ?
g) Referentiality: To what extend does agreement material have a referential function? How is referential agreement related to other sorts of anaphora/ binding?
h) Relation to clitics: What is the relation between pronominal clitics and agreement?
i) Variation: What happens when more than one agreement pattern is grammatical, and does the difference correlate with anything else?
j) Function: What good is agreement?
k) Change: Where does agreement come from, and what happens to it over time?
l) Realization: How is agreement implemented formally?

like that between the subject and the verb. A good account of agreement must thus distinguish such types.

Whatever the type of agreement, it is noted that the interaction between grammatical elements in many agglutinating languages depends very much on the agreement features carried by the individual elements within the syntactic configuration. Typically, lexical items carry agreement features which may include person, number, case, noun class, gender, and even tense. In LuBukusu, some of these features play a crucial role in licensing the occurrence of the RFM, although some of them may not be overtly realized on the surface forms of the anaphoric elements. Consider the following:

**(21)** **A-a-e-siim**-a
SM-Tns-RFM-love-fv
*He loves himself*

According to the present analysis, incorporated affixes such as the RFM enter into the derivation with feature specifications, normally marking agreement. In MP, lexical elements enter the computational process with three sets of features; head, specifier, and complement features (Chomsky 2000). Checking then occurs for uninterpretable features. Once checked, they are erased. This means that the features must match/ agree for the structure to be grammatical. The licensing condition in this case, then, is agreement. In example (21), the RFM has interpretable person and number features, but uninterpretable case feature. On the other hand, the verb with which it is merged has uninterpretable person and number features but interpretable case feature. When merging takes place in the course of the derivation, agreement between the matching features takes place and the structure converges.

### 2.1.2.2 Constituent ordering

The syntactic environment is also defined by the order in which morphemes attach to the verb stem. The RFM, occurs as a prefix, but is always next to the verb stem regardless of the number of other prefixes occurring with it. This is represented in section 2.1.1.1 above. In other words, it is part of the verb's internal structure as opposed to the SM which can be separated from the root verb by other prefixes such as tense and aspect markers. See chapter one for details of such ordering.

### 2.1.2.3 Grammatical Marking, Functions and Valence Operations

In this section, I examine the status of the RFM in terms of grammatical categorization, functions and valence operations. The important question is whether the RFM is an agreement affix or an incorporated pronoun. Or further still, whether it occupies an argument position as an affix or it only represents an argument. For many languages it is difficult to distinguish pronouns from agreement affixes. (See, for example, Mchombo 2004 for details on such difficulties). Before I consider such issues, it is important to note that there is no direct correlation between the function of a particular device in one language and superficially similar devices in other languages- though there are generalities that can be made. For example, free pronouns in English function roughly like person marking does in LuBukusu. Whereas verb agreement in English cannot constitute the only reference to the participant, in languages such as LuBukusu, agreement affixes marked on the verb are sufficient as subject or even object indicators, making it possible for a verb to occur alone as a clause. Consider sentence (22) below.

**(22a)** **Wekesa a-itaal-ang-a kumupiira**
Wekesa SM-kick-Asp-fv CL3ball
*Wekesa kicks the ball*

**(22b)** a-itaal-ang-a kumupiira
SM-kick-Asp-fv CL3ball
*He kicks the ball*

Notice that in LuBukusu, person marking on the verb is an anaphoric device i.e. it may count as the only reference to the subject or object as shown in (25b). In English person marking merely agrees with an independent subject. In GB terms, LuBukusu unlike English is a pro-drop language. Because of such features, LuBukusu free pronouns are used very rarely in discourse, and may easily be described as Focus constructions. The RFM, for example, involves incorporation of the object NP into the structure of the verb. (See section 2.1.2.1.2 for evidence on this kind of analysis).

Still another question on categorization arises: Could the RFM be analized as an anaphoric clitic? An anaphoric clitic is a device that is not free morphologically- it must attach (cliticize) to another word. This then means that a clitic will normally function either at the phrasal or clausal level. Further, like pronouns, clitics are in complementary distribution with full noun phrases. In LuBukusu, there are various pronominal affixes that are either in complementary distribution or in free variation with noun phrases or with each other, and which can therefore be analyzed as clitics. The affixes include the SM, RFM, OM, and RCM.

These affixes constitute the most common markers of agreement morphology in Bantu languages. Perhaps one of the most detailed studies in this area is offered by Bresnan & Mchombo (1987) on subject and object marking in Chichewa. To them the SM is ambigously both a grammatical agreement affix and a topic anaphoric pronoun (Clitic). On the other hand, the OM is purely a pronominal argument. Evidence given to support

such assertions arises from complementarity possibilities with free NPs playing similar roles. The general idea is that if the presence of an affix excludes the free NP, then the former is purely argumental, with the latter only serving a disambiguating role. Conversely, if the two do not exclude each other, then the question of whether they are pronominal or agreement markers is determined by other factors. We assume presently that this also applies to LuBukusu as shown in (23) and (24) below.

**(23a).** Wekesa a-a-ch-a engo
Wekesa SM-Tns-go-fv home
*Wekesa went home*

**(23b).** a-a-ch-a engo
SM-Tns-go-fv home
*He went home*

In (23a), the overt NP *Wekesa* occurs with the SM which reflects the usual agreement features of person, case and number (note that LuBukusu lacks overt gender marking). In this instance, the SM is an agreement form. However in (23b), the SM is the only reference to the subject, and is therefore a pronominal argument that is anaphoric in status i.e. it is able to select a referent in discourse. In the literature, it has also been analyzed as an anaphoric clitic because of its obligatory attachment to the verb and apparent function at a phrasal level, characteristic of clitics cross- linguistically.[23]

Further, the status of the OM is exemplified below:

**(24a).** Wekesa a-a-p-a Omwaana
Wekesa SM-Tns-beat-fv CL1child
*Wekesa beat (a) child*

[23] The phrase level function is seen where the VP must always have the SM, implying that it is crucial in the formulation of the VP. See especially Mchombo (2004) for arguments in favour of SM as a clitic.

**(24b).** Wekesa a-a-mu-p-a
Wekesa SM-Tns-OM-beat-fv
*Wekesa beat him*

**(24c).** *Wekesa a-a-mu-p-a omwaana
Wekesa SM-Tns-OM-beat-fv CL1child
*Wekesa beat the child*

Unlike the SM, the OM is clearly in complementary distribution with an overt NP in the object position, implying the two have the same function as the internal argument of the verb. However, because the OM's function is merged with a verbal element then it is analyzed as an incorporated pronoun. This complementarity, however, only holds with R-expressions. As seen in (25) below, free pronouns are allowed in the verb's internal position even with the presence of an OM.

**(25)** Wekesa a-mu-u-p-a ni-ye
Wekesa SM-OM-Tns-beat-fv pro-Agr
*Wekesa beat him*

Several questions arise: i) Why is the complementarity selective between R-expressions and free pronouns? ii) What is the role and interpretation of the free pronoun when the argument is already expressed? Is this role similar to that of the subject NP?

To answer both questions, one needs first to realize that the SM and OM play distinct roles (as seen above). In addition, the SM maps the features of an NP that is independently selected, and is hence obligatory as a language specific parameter. Conversely, the OM is itself pronominal, and any free NP occurring with it can only help to reiterate/ emphasize the OM. In other words, the difference in the distribution of the SM and OM lies in agreement directionality. For the SM, it takes place from the NP to the SM. The reverse is true for the OM. Further, only pronouns can be emphatic, because

R-expressions are complete functionally, and are hence excluded whenever their semantic component has already been expressed. This is what happens for the OM. Following such arguments, I deductively assume that since the RFM is parallel to the OM in terms of shared characteristics (identity in occurrence position and, conditioned complementarity with free NPs) it is also characterized as pronominal. These issues of grammatical marking are very crucial especially in determining how the anaphoric elements are to be represented and hence their interpretation. In this study, the RFM assumes the full status of a lexical head heading a phrasal category that may or may not have the specifier position filled. This conclusion becomes central to the representational proposal presented in chapter four.

As is generally the case in Bantu systems, the RFM can only establish relations between the subject of the predicate and some other co- argument. Further complications arise with respect to how the RFM corresponds to arguments that are not otherwise expressed (i.e., are phonetically null such as the droped subject in pro-drop langauges). As will be seen (in chapter three), there is no neat one-to-one matching between argument positions and the number of elements representing such positions. This is because several elements can cooccur and impact the anaphoric interpretation of a single argument position. (The RFM can co-occur with Agr-eene). Interesting questions arise as to how each marker affects the argument structure of the verb, in addition to questions concerning what each affix contributes semantically (for example what semantic contribution does the RFM make when it co-occurs with the phrasal reciprocal to form a reciprocal reading?). These are issues related to grammatical functions and valence operations.

Valence has always been characterized as either syntactic or semantic or both (Payne 1997). Semantic valence refers to the number of participants involved in the action or state of the proposition. Syntactic valence on the other hand is the number of arguments that a verb can take at any given time. These arguments can be full NPs, affixes or even zero realization (if it has a referential value). In LuBukusu the RFM is characterized as a morphologically marked means through which valence is reduced. In other words it merges the entities that are involved in the action/ state of the verb, so that only one entity fulfills two semantic roles. The RFM specifies that the agent is the same person as the patient. As will become clear later, valence reduction by reflexivization is perhaps more clear with morphological and lexical means (where the verb is inherently reflexive without any overt realization). Because of the separation of syntactic and semantic valence, this study argues that the RFM increases valence semantically, but syntactically reduces the verb valence. Syntactic means reminiscent of English reflexives is possibly an extra argument (syntactically speaking). Since LuBukusu characterizes both morphological and lexical reflexives, then issues of valence need fresh theoretical recasting.

## 2.2 The Reciprocal Marker

The RCM is perhaps one of the most studied affixes in Bantu morphology (Kimenyi (1980), Bokamba (1981) Mchombo (1991), Nurse and & Philippson (2006)). In this section, I discuss the structure of LuBukusu RCM in terms of morphological realization and syntactic context.

### 2.2.1 Morphological realization

The reciprocal in LuBukusu, just as in most other Bantu languages (See especially Nurse and Philippson (2006) for a comparative study of the Bantu reciprocal) is clearly marked by the verbal suffix /*-an-*/.

**(26a)** Babaana ba-a-rem-an-a
Children SM-Tns-cut-RCM-fv
*The children cut each other*

**(26b)** Efwe khu-a-biyil-an-a
We SM-Tns-hate-RCM-fv
*We hate each other*

Notice that the RCM, unlike the RFM, is relatively morphologically fixed; only occurring as the suffix *–an-* in most contexts (Exceptions are considered later in this section). In order to explain why this is the case, we need to draw some lessons from the variability in the RFM. We noted that the RFM changes its form mainly due to the numerous variations of the neighbouring morphemes. In most cases, these affixes are tense markers, and since tense in LuBukusu varies considerably, then there is likely to be spill over effect of corresponding variation to adjacent affixes, the RFM included. The scenario is quite different when it comes to the RCM. Perhaps most important is the fact that its position is typically post verbal, therefore making it to interact more often with the final vowel, and since the final vowel is always realized as [*a*] in LuBukusu, an adjacent affix is also likely to acquire similar characteristics. It is also notable that most Bantu languages have the *–an* reciprocal marker as shown in Digo and Kiswahili in (27) below.

**(27a)** Watoto wa-li-pig-an-a **Kiswahili**
CL2child SM-Tns-hit-RCM-fv
*(The) children hit each other*

**(27b)** Anache a-pig-an-a **Digo**
CL2child SM-hit-RCM-fv
*(The) children hit each other*

**(27c)** Azimai wo-on-an-a **CiNsenga**
Women SM-see-RCM-fv
*Women see each other*

In some languages the reciprocal is realized by a related but slightly different form.In Luganda and Ci-Yao, for instance, the verbal suffix *-agan-* is used; in Kikongo, the dialect of Zoombo region of Northern Angola the morpheme *-azyan-*is used (Carter & Makoondekwa, 1987), and in Runyambo the reciprocal is realized by the morpheme *-angan-* (see Rugemalira 1993). In all these languages the suffix *-an-* remains, sometimes restricted to lexicalized forms whose roots are no longer attestable as independent verbs within the language. In Runyambo, the following reciprocals, derived with *-an-*, lack independently existing verb stems: *bag-an-a* 'share, divide up'; *bug-an-a* 'meet'; *fuka-an-a* 'wrestle, struggle'; *hak-an-a* 'argue'; *iw-an-a* 'fight'; *ing-an-a* 'be equal.' Reciprocals with the form *–an-* are also derivable from independently attested verb roots. These include forms such as *tong-a* 'demand payment' *tong-an-a* 'quarrel'; *tond-a* 'create' *tond-an-a* 'discriminate against (by origin)'; *nyw-a* 'drink' *nyw-an-a* 'become friends'; *jend-a* 'go, walk' *jend-an-a* 'go together.' In LuBukusu this includes forms such as *lom-a* 'speak' *lom-an-a* 'quarrel'; *ing-a* 'tighten' *ing-an-a* 'argue'. Perhaps a more detailed analysis may reveal a clear pattern in Bantu languages. Because of the evidence available, I argue that whether or not a given verb is inherently reciprocal will very much depend on the meaning of the verb involved. I refer to such verbs as those of 'friction' and 'convergence'. From the examples, they include; *meet, argue, quarrel, struggle,* and others. The regular formation of the reciprocal in Runyambo as, in the other

languages listed above, is through the derivation makes sense." ([Rugemalira, 1993 :150). This is shown in such derivations as *nob-angan-a* 'hate each other'; *jun-angan-a* 'help each other.

Even with the attestable differences in the form of the reciprocal affix in Bantu, the form *–an-* still stands out as common. with variations only brought about by addition of some extra material to the basic form. On such a basis, it is then viable to conclude that in LuBukusu, like in many other Bantu languages, the RCM is basically *–an-*. In case of variations, as in monosyllabic stems, addition of other material to the base is involved. I briefly examine the variation next.

### 2.2.1.1 Monosyllabic Stems

In LuBukusu, variation in the conventional *-an* occurs in a small number of verbs with a monosyllabic stem where the consonant [č] is inserted immediately before the RCM to interrupt the possibility of a hetero- morphemic cluster of VV forms (Mutonyi (2000)). Notice that the verb stem in the reciprocal form has a long vowel *ee/ ii* preceding the RCM. This is shown in the following set of examples:

| Verb Stem | RCM | Gloss |
|---|---|---|
| -r-a | -chan- | -reechana- 'put each other' |
| -sy-a | -chan- | -syeechana-'grind each other' |
| -ry-a | -chan- | -riichana-'fear each other |
| -w-a | -chan- | -weechana-'give each other' |
| -ly-a | -chan- | -liichana-'eat each other' |
| -chy-a | -chan- | -chiichana-'go for each other' |
| -kw-a | -chan- | -kwiichanakho-'fall on each other' |

**Table 7: RCM in Monosyllabic stems**

Further notice that the long vowel in the reciprocal verb is either [ii] or [ee] normally selected on the basis of how the basic verb stem patterns. There is enough evidence to lead to the following generalizations:

**Generalization 11**

In monosyllabic verbs with RCM, [ee] is the default form selected in all environments. [ii] is selected when the verb root ends in a glide that is not always part of the stem when the verb undergoes other derivational processes such as nominalization.

Consider the following data:

| **(28) Verb Stem** | **Nominalized Forms** |
|---|---|
| i) Sy-a 'grind' | buusye (flour), esyo (grinding stone) |
| ii) W-a 'give' | bahani (givers) |
| iii) Ry-a 'fear'-verb | buuri (fear-noun) |
| iv) Ly-a 'eat' | buliilo (where something is eaten) |

In (i) and (ii), the glide remains part of the derived forms indicating that when the RCM is added, the preceding vowel will be [ee]. On the other hand, in (iii) and (iv), the glide is lost during nominalization hence the selection of [ii] in reciprocal formation.

The form [-chan-] is the RCM variant because the insertion of a consonant before the conventional [–an] is an idiosyncratic property of reciprocal formation. Other verb extensions such as applicatives and causatives do not trigger similar patterning.[24] Consider (29) below:

[24] The intensive [ak] like the RCM also requires the addition of [ch] in monosyllabic stems. In the passive, such stems trigger the insertion of [bw]. See Mutonyi (2000) for details

| (29) Verb Stem | Applicative | Causative |
|---|---|---|
| i) Sy-a 'grind' | Sy-eel-a | Sy-eesy-a |
| ii) W-a 'give' | W-eel-a | W-eesy-a |
| iii) Ry-a 'fear'-verb | R-iir-a | R-iisy-a |
| iv) Ly-a 'eat' | L-iil-a | L-iisy-a |

More evidence to support the consonant insertion occurs in cases of verbal extension co-occurrences. In all cases where the RCM comes immediately after the verb stem, there is insertion of the consonant. However, when other extensions come first there is no such insertion. This is exemplified below:

**(30)** Sy-el-an-a Vs Syee-chan-il-a

**2.2.1.2 The Reciprocal Marker in Phrasal Constituents**

Like the RFM, the RCM also occurs in nominalized phrases where the form marking reciprocal still remains [an]. However, unlike the RFM, the productivity of the RCM with different noun classes is quite limited to only a few classes. Most specifically, it occurs with classes 5, 14 and 15 as shown below, respectively.

**(31a)** Lili-p-an-a lye babaana
CL5-fight-RCM-fv of CL2child
*The children's fight*

**(31b)** Bubu-p-an-i bwe babaana
CL14-fight-RCM-fv of CL2child
*The children's fight*

**(31c)** Khukhu-p-an-a khwe babaana
CL15-fight-RCM-fv of CL2child
*The children's fight*

All the other classes yield ungrammatical structures. Further note that the reduplication of the class prefix in (31) is not a productive process. Instead there is enough evidence to

suggest that it only occurs with the verb for fight '*ipa*' which can also occur as '*khupa*'. All the other verbs only involve addition of the noun class prefix with the RCM. Such flexibility also allows for a possibility of variation in (31a) and (31b) as shown in (32a) and (32b) below.

**(32a)** Likhu-p-an-a lye babaana
CL5-fight-RCM-fv of CL2child
*The children's fight*

**(32b)** Bukhu-p-an-i bwe babaana
CL14-fight-RCM-fv of CL2child
*The children's fight*

### 2.2.2 Syntactic Restrictions to the Reciprocal Marker

In this section, I examine the syntactic relations existing between the RCM and other constituents in a syntactic context. Most specifically, the following aspects are considered: a) Licensing Conditions; b) Constituent ordering; c) Grammatical Functions and Valence Operations

#### 2.2.2.1 Licensing Conditions

There are several conditions that must be met for the RCM to occur in the contexts in which it does. In this section I discuss some of these conditions.

##### 2.2.2.1.1 Transitivity, Locality and Agreement

Like the RFM, the RCM is also licensed by verb transitivity i.e. it only occurs in verbs that also allow an object, within their VP. The object may either surface as the OM or as a free nominal form. However both the OM and the nominal form never occur together.

**(33a)** Ba-a-ir-a omwaana
SM-Tns-kill-fv CL1child
*They killed the child*

**(33b)** Ba-a-ir-an-a
SM-Tns-kill-RCM-fv
*They killed each other*

**(33c)** *Ba-a-ir-an-a omwaana
SM-Tns-kill-RCM-fv CL1child
*They killed each other *the child*

In (33a), the verb '*ira*' is transitive and therefore requires an object, in this case represented by *omwaana.* Consequently the verb can take a RCM as seen in (33b). However both the object and RCM are not allowed in the same structure, hence the ungrammaticality of (33c). Syntactically speaking, the RCM is therefore equivalent to an internal argument. Its presence precludes the presence of any other constituent serving the same function. On the other hand, consider the intransitive verb in (34) below.

**(34)** ?Ba-tim-an-a
SM-ran-RCM-fv
**They ran each other*

Since the verb is intransitive, the use of a RCM yields ungrammaticality. There is a possibility, though, of the sentence being acceptable if one considers contextual interpretation, where the verb *–tima* may mean 'chase'. In this case it is possible for people to 'chase each other'. On the other hand if it means 'run', then one cannot conceptualize a situation where people can 'ran each other'. This means that the verb may either be interpreted transitively or intransitively.

In addition, if one of the valence increasing affixes such as causative or applicative is added, then the RCM and the OM can co-occur.

**(35)** Ba-tim-an-il-a
SM-ran-RCM-Appl-fv
They ran for each other

Because the applicative marker increases the arguments of a verb, its presence on an intransitive verb makes it transitive hence allowing a RCM.

In the literature, reciprocal elements, just like reflexives, are said to be subject to the binding relations i.e they select an antecedent in the local domain which is normally the minimal clause or NP containing the reciprocal and the accessible antecedent (Chomsky 1981). In Bantu, the reciprocal has received varied analyses. Mchombo (2004:13), for example, observes that the reciprocal in Bantu, unlike the reflexive, is not a nominal argument subject to principles of binding. Instead, 'it is a morpho-lexical process deriving reciprocal predicates apparently not susceptible to syntactic binding'. However, this is a view that turns out to be controversial. In a study of Malagasy, an Austronesian language, Keenan and Razafimamonjy (2001), (henceforth K&R), observe that the reciprocal is derived morphologically through the affixation of the reciprocal morpheme *–if-* to a transitive verb. The reciprocal *–if-*

> "...immediately precedes the active prefix *aN-*. Its presence excludes an overt accusative NP..."(Keenan & Razafimamonjy 2001: 41).

In this regard, the Malagasy reciprocal behaves in a manner comparable to that in LuBukusu (See section 2.1.1.2 above). However, K&R state that the position excluded by the presence of the reciprocal remains syntactically active, indicated by the empty category symbol *'e.'* They then claim that

> [It] is this empty position which corresponds to the presence of the reciprocal pronoun each other /one another in English and which determines one of the arguments, of the reciprocal relation used in semantic interpretation of Ss built from reciprocal verbs. (ibid.)

K&R make the further claim that

"...the NP which *–if-* requires to be plural is the *antecedent* of *–if-*. We also say that this antecedent *reciprocally binds* the reciprocal empty category (ec) licensed by *–if-*. The motivation for this terminology is that given an occurrence of *–if-*, the positions determined by its antecedent and its ec are those which determine the arguments of the reciprocal relation used to interpret the reciprocal expression." (ibid. 42).

Having adopted this stance, K&R pose a number of questions about the antecedent of the reciprocal in Malagasy. These relate to whether the antecedent of the reciprocal (a) always occurs external to the VP projected by the verb that the reciprocal morpheme *–if-* occurs in; (b) c-commands the reciprocal empty category; (c) occurs as an argument of the reciprocal predicate, within the same complete functional complex as the reciprocal verb; and, (d) locally binds the empty category licensed by the reciprocal morpheme. Their conclusion is that for Malagasy, answers to questions (a) (b) and (d) are negative. For (c) the answer is affirmative, that the agent phrases of the non-active verbs are arguments of the verb. The facts about Malagasy are, in relevant respects, comparable to Lubukusu. Note, however, that there is a measure of contradiction in the K& R account. On the one hand, the plural antecedent "reciprocally binds" the reciprocal empty category licensed by the reciprocal morpheme. On the other hand the antecedent does not either c-command or locally bind the empty category licensed by the reciprocal morpheme, as indicated by the negative answers to the questions (b) and (d) above. What then is to be made of the question whether any kind of binding is evident in reciprocal constructions? I turn to this question next.

The idea of reciprocal binding appears to be rooted in efforts to provide a coherent account of the reciprocal relation that is evident in the semantic interpretation of expressions built from reciprocal verbs. Consider the sentence below:

**(36)** Chingokho ne chinyuni cha-a-siim-an-a
CL10chicken and CL10birds SM-Tns-like-RCM-fv
*Chickens and birds like each other*

The sentence has as its primary reading the claim that the chickens like the birds and the birds like the chicken. Equally, the sentence conveys the reading that the chickens like each other and the birds like each other. The latter reading could be derived from sentence coordination, or by the use of the RFM. This additional interpretation will be set aside for now. Taking the standard approach to accounting for the initial interpretation, the assignment of semantic roles is done in the argument structure. A general constraint in the assignment of semantic roles is that each semantic role be assigned to a particular argument of the predicate and each argument be assigned a single semantic role. Naturally, if reciprocal verbs are derived lexically, and the reciprocal predicate is detransitivized, then the single argument and the antecedent must be assigned a single role. Yet, as the semantic interpretation indicates, the group denoted by the antecedent argument requires that the objects denoted participate both as agents and patients or beneficiaries. How can this be achieved? The simple clue lies, once again, in the treatment of reflexives that, apparently, resolve the apparent violation of this general constraint on semantic role assignment. A single entity participates as both agent and patient or beneficiary. The reflexive, as a pronominal argument, receives the semantic role of patient. However, as an anaphor, it is bound to the antecedent, effectively resolving the reading that the antecedent is construed as both agent and patient. In other words, in the discourse structure representation the two syntactic arguments are mapped onto the same entity (cf. Sells, Zaenen & Zec 1986). It is but a simple step to the speculation that reciprocal constructions must

equally involve binding, and that this binding relation is both local i.e. with a clausemate antecedent or 'long distance' i.e. with an antecedent in a higher clause via agreement marking on a host verb- the same conditioning as the RFM.

Agreement plays the crucial role of licensing relations between constituents in syntax. Like the RFM, the RCM does not display overt agreement features. However the role such features play still remains crucial to any theory that seeks to account for the licensing of sentential constituents such as reciprocal pronominals. In LuBukusu, the role of agreement in licensing reciprocal marking depends on the status of the RCM, much as the status of the RFM as an incorporated pronoun ensures licensing via agreement. I take the RCM in LuBukusu to be a pronominal argument for the same reasons the RFM is i.e. that i) it excludes a free NP in the VP internal position; ii) its selection is as a result of the properties of the V lexical head i.e. it only occurs in transitive verbs. Consider:

(37) Babaana ba-a-siim-an-a
CL2child SM-Tns-like-RCM-fv
*(The) children like each other*

In this structure, two categories of agreement, both of which license the RCM, can be identified: a) The agreement between the RCM and the subject NP in terms of number i.e the RCM always requires a plural antecedent; b) The agreement between the verb and the RCM.

In MP, type (a) leads to the tradional binding relation hence the correct interpretation. On the other hand, type (b) concerns agree relations between the verb and its complement. In this case the verb has interpretable case features and uninterpretable person and number features, while the RCM has uninterpretable case features and interpretable person and

number features. In the course of the derivation, the two sets of features agree, and convergence takes place.

### 2.2.2.2 Constituent Ordering

This refers to the order in which morphemes attach to the verb stem. The RCM always occurs as a suffix, but its position is not quite rigid. Whereas it is realized before all post verbal TAM markers, it may occur immediately after the verb stem or with an intervening derivational affix, most probably an applicative, an intensive marker or a causative. The set of data below helps to illustrate these facts.

| **Reciprocal** | **Applicative** | **Causative** | **Intensive** |
|---|---|---|---|
| ba-siim-an-a<br>They love each other | Ba-siim-il-an-a<br>Ba-siim-an-il-a 'They love each other with' | ba- siim-an-isy-a<br>ba-siim-isy-an-a They caused to love each other | ba-siim-an-ak-a<br>*ba-siim-ak-an-a They love each other intensively |

**Table 8: RCM and Verbal extensions**

Notice that the applicative and the causative markers can all occur before or after the RCM, with the possibility of a slight difference in meaning specifying whether or not the the subject is the direct beneficiary or cause of the action. Details are discussed in chapter three and four. The intensive marker on the other hand can only occur after the RCM mainly because the role of the intensifier is more peripheral to the verb i.e it is not part of the verb's subcategorization structure. The details concerning change of meaning as a result of verbal extension ordering, and the varied patterns of co-occureences are discussed in chapter three. The most important question, however, that may arise from such possibilities concerns an explanation as to why some elements are more closely bound to the verb than others. Baker (1985) proposes the Mirror Principle to account for the relations between morphological and syntactic derivations. This, he says should

directly reflect each other. He notes that there is a close parallelism between morphology and syntax i.e. morphological derivation reflects syntactic derivation and vice versa. If, for example, the morphological structure of a complex word is derived through head movement of the lexical root to the heads where the morphemes are base-generated, the Mirror Principle follows straight forwardly. The order of morphemes in a complex word reflects the natural syntactic embedding of the heads that correspond to those morphemes and the order of the constituents in the sentence. In short, there should be a correlation between the position of a morpheme and its position on a tree diagram (The Principle equates syntactic derivation to a tree diagram). The assumption is that those morphemes that are closer to the verb have the lowest projections in the tree, proceeding upwards in that order. The morphemes are heads of their respective projections. The implications of these facts are presented in details in chapter four and five.

#### 2.2.2.3 Grammatical Marking, Functions and Valence Operations

The question of whether the RCM is an anaphoric pronoun or clitic remains crucial in determining its grammatical status, and hence its interpretation. It is slightly different from the RFM. First it occurs as a suffix, and secondly it rarely occurs with a free NP. Its categorization must then also be different. There have been suggestions that it is derivational (e.g. Baker (1985) like most suffixes, and is therefore part of the lexical properties of the verb within which it is realized. In this study, the RCM and RFM assume the full status of lexical heads heading phrasal categories that may or may not have the specifier position filled. In addition, the RCM unlike other suffixes is not purely derivational. On the contrary it participates in syntactic relations just like inflectional affixes do. These issues will be discussed in details in Chapter five.

The verb suffixed with the RCM cannot take a direct object. In other words, the verb combines with transitive verbs to form surface intransitives, as illustrated in (37) above. In terms of meaning, the sentence shows that the NP subject is both agent and patient. This is because the action or activity is mutually done between the participants. According to Mchombo & Nguga, (1994:23):

> The usage of the reciprocal construction is that of ascribing the members of a group the property that they are involved in an activity such that each member is performing the action on others

Indeed, this is the commonest reading and it is clearest when the group consists of two members. Although the construction is syntactically intransitive, it is semantically transitive. Further consider (38) below.

**(38a)** *Omu-sooleeli ne omu-khaana ba-a-p-an-a Wanjala
CL1-boy and CL1-girl SM-Tns-beat-RCM-fv Wanjala.
*The boy and the girl beat each other.*

**(38b)** *Omu-sooleeli ne omu-khaana ba-a-mu-p-an-a
CL1-boy and CL1-girl SM-Tns-OM-beat-RCM-fv.
*The boy and the girl beat each other.*

In (38a), adding a direct object in a reciprocal construction is unacceptable, since the construction is syntactically intransitive. Further, (38b) illustrates that replacing the direct object with the OM is equally unacceptable mainly because such an OM makes the construction transitive, yet a reciprocal only occurs in surface intransitives. However, the construction is semantically transitive since two participants are doing some action on one another. The participants are both subjects (agents) and objects (patient) at the same time. Consequently, the subject must be plural. I examine the Agr-eene form next.

## 2.3 Agr-*eene* alone

In LuBukusu, as noted in section 2.1 above, a free pronominal NP labeled *Agr-eene*, is realized in different positions either optionally or obligatorily, but on all occasions it must have an antecedent that is either in the syntax or in the extended discourse. Significant questions arise concerning *Agr-eene*: a) How is it realized morphologically and which are the factors that affect this realization? b) Which is its relative position within a sentence? c) To what extend does the position affect its interpretation?

In this section, I discuss *Agr-eene* under two broad sections: Morphological structure and Syntactic properties with the aim of providing answers to these questions.

### 2.3.1 Morphological Structure of Agr-*eene*

The Agr-*eene* form is employed in a variety of ways. It can be used as an emphatic adverbial reflexive, it can serve to emphasize a reflexive reading, and in certain circumstances it can form a reflexive reading on its own. The agreement portion of the Agr-*eene* form matches the noun class and person of its antecedent (see table below) and there is evidence that *-eene* may mean 'own' or 'owner', a familiar type of anaphoric atom (see Safir, 1996, 2004). Consider the structure below:

**(39a)** Wekesa a-a-e-chuub-a omweene
Wekesa SM-Tns-RFM-encourage-fv Agr-own
*Wekesa encouraged himself*

**(39b)** Wekesa a-a-chuub-a omweene
Wekesa SM-Tns-encourage-fv Agr-owner
*Wekesa encouraged the owner*

In (39a) Agr-eene refers to emphatic 'own' while in (39b) it is a possessor. Altogether, the Agr-*eene* form consists of three morphemes, the agreement, the noun class marker, and the 'own' morpheme. Consider the following paradigm:

| Person | Singular | Plural |
|---|---|---|
| **First** | Samweene-Myself | Fwabeene-Ourselves |
| **Second** | Wamweene-Yourself | Mwabeene-Yourselves |
| | | |
| **Third** | Omweene-Himself/herself<br><br>Agr+-eene- Itself | Babeene-Themselves<br><br>Agr+eene-Non human |

**Table 9: Realization of Agr-*eene***

From the paradigm in table 9, the morphological realization of the Agr-*eene* form is determined by person, number and noun class features all of which are represented by a segment of the form. Given a form such as the first person singular *samweene*, a morphological segmentation will yield the following structure:

**(40)** Sa-mu-eene
1PSg-CL1-own
Myself

The first segment represents person and number i.e 1st person singular, the next segment is the class 1 affix while the last part means 'own' or when used in a possessive construction, 'owner'. The inherent structure of Agr-eene is then represented in the rule below:

**Generalization 12**

**PERSON + NUMBER- NOUN CLASS- OWN/OWNER**

Out of these slots, the last one is relatively fixed as *–eene*, while the rest vary depending

on the noun that they agree with. Consequently, Just as in most languages, I conclude that LuBukusu distinguishes three persons depending on the relation between the speaker and the interlocutors. Each of these persons has a distinct realization of Agr-eene. Again, conforming to the two number distinctions, the Agr-eene form varies according to number and class marking.

### 2.3.2 Syntactic Properties

A number of issues arise, related to contextual realization i.e. grammatical marking and functions, and licensing of the Agr-eene form. Some of these are discussed below in line with what was said about RFM and RCM.

#### 2.3.2.1 Licensing

The main licensing conditions for the Agr-eene form are transitivity, semantic role, clause typology and agreement.

##### 2.3.2.1.1 Transitivity

Transitivity is the requirement that whenever an Agr-eene form is used to implement emphatic reflexivity, then it must occur with a transitive verb. Other conditions apply elsewhere.

**(41a)** Wekesa$_i$ a- siim-a omweene$_j$
Wekesa SM-loves-fv Agr-own
*Wekesa loves himself'*

**(41b)** Wekesa$_i$ a- e-siim-a omweene$_i$
Wekesa SM-RFM-loves-fv Agr-own
*Wekesa loves himself*

In (41a), *omweene* cannot be construed with *Wekesa,* and therefore it performs the role of direct object. Because of this, this usage is only licensed in transitive verbs. In (41b),

*omweene* is coreferential with *Wekesa,* and must therefore occur in transitive configurations for emphatic reflexive implementation. Note that it is also possible to have Agr-eene in intransitive verbs. In such cases, the referent can only be the subject of the clause, with the Agr-eene form serving an emphatic role for the subject.

**(42)** Wekesa$_i$ a- a-tiim-a omweene$_i$
Wekesa SM-Tns-run-fv Agr-own
*Wekesa himself ran*

In this structure, Agr-eene and Wekesa are coreferential with the role of the former mainly being an adjunct.

**2.3.2.1.2 Semantic Role**

From the discussion above, Agr-eene distinguishes three main roles: a) Overt syntactic Reflexive; b) Emphatic pronominal; c) Independent pronoun. These roles determine, to a great extend, the context of occurrence. Also see the paradigm above for illustrations of how agreement features license the Agr-eene form.

**a) Reflexive Agr-eene**

Consider:

**(43)** Wekesa a-a-e-siim-a omweene
Wekesa SM-Tns-RFM-like-fv Agr-own
*Wekesa likes himself*

In this sentence *omweene* is reflexive because it cooccurs with the RFM to emphasize the reflexivity. It is not allowed to have an extra sentencial antecedent. In addition, its presence is optional because the RFM already indicates the reflexivity. Different issues arise for Agr-eene in (44).

**(44)** Wekesa a- a-kachul-il-a Wanjala khu omweene
Wekesa SM-Tns-tell-Appl-fv Wanjala on Agr-own(er)
*Wekesa talked to Wanjala about himself*

*Omweene* occuring as a complement of the PP may be interpreted as being reflexive. Its antecedent however varies depending on the speaker's intended meaning. It may refer to either Wekesa or Wanjala within the same sentence. In addition it may refer to an extra-sentencial antecedent in which case it becomes a possessive pronoun.

**b) Emphatic Pronominal**

This is perhaps the most common use of the Agr-eene form. Consider:

**(45a)** Wekesa omweene a-ch-il-e
Wekesa Agr-own SM-go-Tns-fv
*Wekesa himself went*

**(45b)** Wekesa a-ch-il-e omweene
Wekesa SM-go-Tns-fv Agr-own
*Wekesa went himself*

**(45c)** Omweene a-ch-il-e
Agr-own SM-go-Tns-fv
*He himself went*

**(45d)** A-ch-il-e omweene
SM-go-Tns-fv Agr-own
*He went himself*

In all these cases, the Agr-eene form helps to put emphasis on the antecedent. Note that this emphasis is contrastive i.e if the antecedent is X then the meaning derived is 'no one else but X'. In (45a) and (45b), the antecedent is undoubtedly *Wekesa*. What changes is only the position of the pronoun reminiscent of adverbials. Agr-eene in (45c) also has similar characteristics i.e marking contrastive focus. The antecedent in this case is an implied subject that has been elided but which is marked by the SM. In addition, this Agr-eene can itself be the subject, but in that case it will only be possessive, referring to

owner. (45d) is the same as (45b), with an omitted subject NP, which is the antecedent and the centre of contrastive focus.

In addition to the properties already discussed, different issues arise when Agr-eene occurs in a different clause as in (46) below.

**(46a)** Wekesa a-a-bool-el-a Wanjala a-li omweene a-chi-il-e
Wekesa SM-Tns-tell-Appl-a Wanjala Agr-that Agr-own SM-go-Tns-fv
Wekesa told Wanjala that *himself/he himself went

**(46b)** Wekesa a-a-bool-el-a Wanjala a-li a-chi-il-e omweene
Wekesa SM-Tns-tell-Appl-a Wanjala Agr-that SM-go-Tns-fv Agr-own
Wekesa told Wanjala that he himself went

In (46a), omweene in the lower subject position is interpreted as referring to Wekesa, Wanjala or an extra-sentential antecedent. The role remains contrastive, but if the antecedent is extra-sentential, then it may be possessive. In (46b), the possessive interpretation is ruled out.

**c). Possessive Form**

As seen in (46c) and (46a), Agr-eene can also occur as a possessive construction. In such cases, it is equivalent to English 'owner' or Kiswahili 'Mwenye'. For this role to be realized, the form must occur on its own in an argument position, either as subject or object.

**2.3.2.1.3 Clause Typoloy**

This deals with the type of clause in which the Agr-eene form occurs, and its relation to a potential antecedent. In LuBukusu Agr-*eene* can occur in both simple and complex clauses, either as subject, object or adverbial and can pick out any focused individual in

the domain of discourse, including those outside the sentence. It can have a non-c-commanding antecedent. Consider the following:

**(47a)** Billi a-a-bon-a omweene
Billi SM-Tns-see-fv Agr-own
*Bill saw the owner/*himself'*

**(47b)** Omweene a-a-bon-a
Agr-own SM-go-asp-fv
*He went*

**(47c)** Billi a-a-tim-a omweene
Billi SM-Tns-run-fv Agr-own
*Bill himself ran*

The first two simple clauses host the Agr-eene form that acts as a possessive pronominal with a potential antecedent contained in the discourse context. In (47a), the form occurs as an object while in (47b), it is the subject. Notice that the type of verb (whether transitive or intransitive) has an effect on the nature of the antecedent selected. In (47a), the transitive verb makes Agr-eene a pronoun with an extra sentential antecedent, whereas in (47c), the intransitive verb makes it emphatic. See discussion on (46) above for Agr-eene in complex clauses.

**2.3.2.1.4 Agreement**

Agreement properties are very important in licensing the occurrence of Agr-eene. In section 2.3.1 we noted that the antecedent usually determines the overt agreement features of person, number and noun class attached on the Agr-eene form. Thus in (60), 'kimisaala'as an antecedent is marked on Agr-eene.

**(48)** Kimisaala kimieene ky-a-fun-ikh-a
CL4tree Agr-own SM-Tns-break-Stat-fv
*The trees themselves broke*

Further, Agreement via agree relations in the computational process is crucial in ensuring grammaticality. Agr-eene in object position, for example is licensed when it merges with the verb head for feature checking. On the other hand, Agr-eene in subject position is licensed when it raises to [Spec, TP] position as an EPP requirement. Details are discussed in chapter five.

### 2.3.2.2 Grammatical Marking and Functions

Whether or not Agr-*eene* has the status of an argument or an adjunct appears to depend on its syntactic context. Agr-*eene* seems to have the status of an argument when it occurs in the absence of the RFM. In the latter case it does not appear to act as an anaphor, at least with respect to locality restrictions, rather it acts in the manner of what Safir (2004) calls an Unbounded Dependent-form (UD-form).

In terms of functions, different issues arise. For example, it is noticeable that Agr-eene in object position appears to act like a pronoun susceptible to Principle B effects. Consider:

**(49)** Wekesa a-p-a omweene
Wekesa SM-beat-fv Agr-own
*Wekesa beat *himself/ him*

If the verb is transitive, the use of Agr-*eene* as the object cannot be coconstrued with the subject, but if the verb is intransitive, then the reflexive is coconstrued with the subject, has adverbial force, and must be treated as an adjunct as in (50) below:

**(50)** Wekesa a-ch-a omweene
Wekesa SM- go-fv Agr-own
*?Wekesa went himself*

The distribution of Agr-*eene* alone also extends, as might be expected for a UD-form, to long distance anaphora, where it is either bound to a (non-local) argument in the sentence or is anaphoric to a discourse antecedent (or is focused). Thus, except where it co-occurs with the RFM or RCM, Agr-*eene* behaves much like a pronoun.

## 2.4 The Phrasal Reciprocal

Sometimes Agr-eene can be reduplicated in certain contexts to serve specific grammatical functions. Consider:

**(51a)** Babaana ba-a-p-an-a babeene khu beene
CL2child SM-Tns-fight-RCM-fv Agr-own on Agr-own
The children fought each other

**(51b)** Babaana ba-a-p-an-a babeene ne babeene
CL2child SM-Tns-fight-RCM-fv Agr-own with Agr-own
The children fought each other(and nobody else)

**(51c)** Babaana ba-a-tim-a * babeene khu beene/
CL2child SM-Tns-fight-RCM-fvAgr-own on Agr-own/
Babeene ne babeene
Agr-own with Agr-own
The children ran (and nobody else)

Two forms of reduplicated Agr-eene are noticeable: Agr-eene khu Agr-eene and Agr-eene with Agr-eene. The former occurs only with the RCM to emphasize the reciprocity, while the latter is freer i.e. it may occur with the RCM as in (51b) or with an intransitive verb as in (51c). It helps to mark contrastive focus on an antecedent. In this manner it acts more like the emphatic Agr-eene form discussed in section 2.3 above. For a better understanding of the nature of the reduplicated Agr-eene, I discuss it next based on its morphological realization and syntactic context of occurrence.

### 2.4.1 Morphological Realization

The phrasal reciprocal can be represented as *'x-eene khu/ne x-eene'* where *x* is an agreeing subject marker while *khu* is the preposition *on* (and occasionally other prepositions) and *ne* literally meaning *with*. This shows that the action is on each other for the former, and provides contrastive information focus for the latter. The phrasal reciprocal must show agreement features of the noun class of the antecedent. It is a sort of reduplication of the reflexive carrying the morphological features of number and person. As mentioned in the introduction to section 2.3, the *Agr-eene* form has three morphemes: person, number/class and OWN. *Babeene* is for example divided, the underlying structure will be *babeene khu babeene*. However to avoid repetition, the *ba-* 3rd person prefix, *ba-* number Agr-affix, and *–eene-* 'OWN'. When this is made a reciprocal second occurrence deletes the person affix to produce the structure; *babeene khu beene*. It would then be predictable that *fwabeene khu fwabeene* will become *fwabeene khu beene, omweene khu omweene* will be *omweene khu mweene*. The emphatic reduplicated Agr-eene does not however involve such reductions.

### 2.4.2 Syntactic Realization

The reduplicated Agr-eene occurs in different syntactic environments and serves varied grammatical functions. When we consider the data in (51) above, the following occurrence issues arise:

a) Two forms of reduplicated Agr-eene are realized: Agr-eene on Agr-eene (henceforth R1- Agr-eene) and Agr-eene ne Agr-eene (R2- Agr-eene).
b) R1-Agr-eene is only licensed in reciprocal constructions in conjunction with RCM to reiterate the reciprocity.

c) R2-Agr-eene occurs in any other environment that allows contrastive focus or emphasis on a given NP that functions as the antecedent.

d) R1-Agr-eene only occurs in underlying object position as an optional constituent following from the pro-drop characteristics of the language. Consequently, its antecedent is always the subject.

e) R2-Agr-eene has the characteristics of a pronoun, and therefore can occur in canonical pronoun positions, usually with either a sentential or an extra-sentential antecedent.

## 2.5 Conclusion

In this chapter, I examined the reflexive marker and the reciprocal marker giving their underlying forms as *–i-* and *–an* respectively with inherent variations mainly motivated by adjacent phonemes of other morphemes. The syntactic relations that the RFM and RCM enter into with other elements in a syntactic configuration were also discussed with great emphasis on the licensing conditions, constituent ordering, grammatical marking, and grammatical functions and valence operations. In addition, lexical *Agr-eene* form and the phrasal reciprocal usually represented as *Agr-ene* on *A*gr-*eene* were also discussed.

# Chapter Three

# Binding Relations

## 3.0 Introduction

This chapter focuses on the question of how the anaphoric elements discussed in chapter two are able to achieve coreference relations. I argue that the varied interaction patterns between the said elements with their antecedents on one hand and with each other, on the other, form a set of strategies for anaphora in LuBukusu. In order to capture all these facts, the chapter is structured as follows: section 3.1 deals with the interaction between individual anaphoric elements with their antecedents, in what I conveniently refer to as local binding relations. Section 3.2 examines non- local coreference relations triggered mainly by Agr-eene and the phrasal reciprocal. Section 3.3 deals with focus constructions as they interact with anaphora. Section 3.4 discusses anaphora in special verbs that do not require overt marking. Section 3.5 concludes the chapter.

## 3.1 Local Binding Relations

The interaction between each anaphoric element and its potential antecedent within the same clause has always been some kind of default basis in dealing with binding relations. In this section, I largely do the same by examining how the RFM, RCM, Agr-eene, and the phrasal reciprocal achieve coreference with clausemate antecedents.

### 3.1.1 Reflexive Binding

In the literature, (see especially, Black 1998, Cole et al, 2001, Reuland 2001, and Kayne 2005) one of the defining features of anaphors is accessibility to a subject that is context

specific. The key hypothesis for the LuBukusu facts is that the RFM obligatorily selects a clause mate antecedent i.e. the subject of the verb hosting the RFM. Consider the example below.

**(1)** Wekesa a- e-siim-a
Wekesa SM-RFM-likes-fv
*Wekesa likes himself*

In (1), the RFM is construed with *Wekesa* within the domain of a clause. In case of long distance relations, then the distant antecedent must share features with the SM on the host verb (see (2) below). In other words, the RFM anaphoric strategy is largely local. Consider the contrasting environments below:

**(2a)** Wekesa a- lom-a a-li a-e-siim-a
Wekesa SM-said-fv Agr-that SM-RFM-likes-fv
*Wekesa said that he likes himself*

**(2b)** Wekesa a- lom-a a-li Wanjala a-e-siim-a
Wekesa SM-said-fv Agr-that Wanjala SM-RFM likes-fv
*Wekesa said that Wanjala likes himself*

In (2a), construal is with an unspecified pronominal subject of the lower clause. There is also the possibility of the RFM referring to the matrix subject at least via proxy. This is achieved if the person referred to by the RFM is similar to the matrix subject. How then does one represent the two competing antecedents? Issues related to agreement again come into play. However, which ever way one looks at it, issues of semantic and syntactic representations need to be distinguished first in any adequate theory of construal. In this study, I develop a hypothesis, along the lines of Safir (2004) that is able to distinguish between strict local binding that is subject to binding principles with non-local binding controlled by logophoric or semantic factors. I defer detailed explanations

of logophoricity to section 3.2 for obvious reasons: it deals with non-local binding. Meanwhile, I concentrate on a brief justification for the hypothesis that the RFM is strictly local. In the literature, two tests have often been put forth to help in the characterization of the type of binding relations. These include c-command and the type of reading under VP ellipsis. Sentence (2b) best illustrates c-command relations. Notice that the RFM in the lower clause can only refer to *Wanjala* the subject of the same clause. It is ungrammatical for example to relate the RFM to the higher subject because this would contravene the traditional c-command relations. There is some kind of 'blocking effect' by intervening antecedents.[25] This is best illustrated in (3) where both the lower and higher subjects have different number features, so that the higher subject alone matches the SM of the lower verb.

**(3)** *Babaana ba-a- lom-a ba-li Wanjala ba-e-siim-a
CL2child SM-Tns-say-fv Agr-that Wanjala SM-RFM likes-fv
*The children said that Wanjala likes *themselves/ them*

The ungrammaticality is as a result of the SM in the lower verb being coreferential with the higher subject *babaana* in the process overlooking the clause mate antecedent. This contravenes the locality requirements. The lower subject must therefore agree with the lower SM in order to establish coreference. On the type of reading under VP ellipsis, consider the following:

**(4a)** Sikuku a-a-lom-a a-li a-a-e-siima. Sitawa yeesi
Sikuku SM-Tns-say-fv Agr-that SM-Tns-RFM-like. Sitawa also
*Sikuku said that he likes himself. Sitawa did too.*

**(4b)** Sikuku a-a-lom-a a-li a-mu-siim-a. Sitawa yeesi
Sikuku SM-Tns-say Agr-that SM-OM-like-fv. Sitawa also
*Sikuku said that he likes him. Sitawa did too.*

[25] The use of the term 'blocking effect' is not used in the sense of Cole et al (2001) to refer to the situation in languages such as Chinese where there is no overt morphological agreement between the subject and the verb such that the LD reflexive *Ziji* is blocked from being coindexed with a matrix subject if there are intervening subjects whose person features are different from those of the subject.

Notice that (4a) has a RFM as the object which in turn triggers a reflexive reading within the minimal clause domain, while (4b) has an OM that is not minimally bound. In the literature, a reflexive should result in both the 'sloppy' and 'strict' reading under VP ellipsis, while a pronominal only has the strict reading. This distinction is extended to the examples in (4a) and (4b) with the interpretations in (5a) and (5b) respectively.

**(5a) i)** Sitawa said that Sikuku likes himself.
**ii)** Sitawa said that she likes herself.

**(5b) (i)** *Sitawa said that Sikuku likes him.
**(ii)** Sitawa said that she likes him

2 while the second is the strict reading. It is therefore correctly predictable that in the reflexive form, both readings are available whereas only the strict reading is available for the sentence with the pronominal form. The same argument can be extended to the RCM as discussed in the next section.

### 3.1.2 Reciprocal Binding Relations

In LuBukusu, reciprocal marking is indicated by the RCM, and as noted for the RFM, the key hypothesis for LuBukusu reciprocal relations is that the RCM is always bound by the subject of the verb that is also its host. In other words, reciprocal binding occurs within the minimal clause that contains both the antecedent and the RCM. Consider the following sentences.

**(6a)** Basoleeli ba-a-siing-an-a
CL2boy SM-Tns-washed-RCM-fv
*(The) boys washed each other*

**(6b)** Basaani ba-a-chanu-an-a ka-machuune
CL2man SM-Tns-combed-RCM-fv CL6-hair
*(The) men combed each other's hair*

**(6c)** Bakhasi ba-a-bool-el-a basaani ba-siim-an-e
CL2woman SM-Tns-tell-Appl-fv CL2man SM-like-RCM-fv
*The women told the men to like each other*

The RCM is construed with the subject NP *Basoleeli* in (6a) while in (6b), the antecedent is the subject *Basaani*[7]. In (6c), the RCM embedded in the subordinate clause is construed with the subject of the same clause and not that of the matrix clause. In all the cases therefore, the antecedent is a c-commanding subject. More support for positing a local binding relation for the RCM comes from the availability of both the strict and sloppy readings under VP ellipsis. Consider the following sentence:

**(7)** Bakhasi ba-a-lom-a ba-li ba-siim-an-a. Basaani boosi.
CL2women SM-Tns-say-fv Agr-that SM-like-RCM-fv CL2men also
*The women said that they like each other. The men did so too.*

In this example, the sloppy and the strict readings are both available, and are represented in (8) below, respectively.

**(8) i)** The men said that the women like each other.
**ii)** The men said that they (men) like each other.

With the nature of the binding relations for the affixes already established, I turn to the lexical forms next.

### 3.1.3 Agr-eene and Binding Relations

The relation between Agr-eene and a potential antecedent is quite flexible. It can refer to an exclusively local antecedent, an extra- sentential one or both. The conditions under

which this happens are varied, and mainly depend on the role that Agr-eene plays in any given construction. As noted in section 2.3.2.1.2, Agr-eene can be used as: a) an emphatic reflexive mainly emphasizing a reflexive reading; b) A discourse dependant form mainly in long distance environments; c) The sole marker for a reflexive reading especially in oblique object positions where it serves as an argument.

Unlike the RCM and RFM, the distribution of Agr-eene according to whether or not it is locally bound or non-local, depends on the properties it has, and the nature of the context in which it occurs. In order to capture such intricate properties of Agr-eene, two main categories are identified: a) unbounded Agr-eene; and b) bound Agr-eene. I discuss the former category in section 3.2 since it involves non-local binding. Meanwhile, I examine bound Agr-eene because of its local orientation.

The most common type of Agr-eene that is strictly bound to an antecedent within the minimal clause is that which forms a reflexive reading in conjunction with the RFM as in (9) below.

**(9)** *Wekesa* a-a-*e*-siim-a *omweene*
Wekesa SM-Tns-RFM-like-fv Agr-own
*Wekesa likes himself*

Here, Agr-*eene* is strictly bound to the subject *Wekesa* as it enhances the reflexive reading that is indicated by the RFM. However, this Agr-*eene* is not an argument on its own, and is therefore analysed as an adjunct. Note that this conclusion is based on the fact that Agr-*eene* is optional, its position can change so that it does not necessarily follow the verb, and an adverb can be scrambled between it and the verb. Recall that these features are common to adjuncts.

Another type of locally bound Agr-eene occurs as an adjunct in constructions such as (10) where the verb is intransitive.

**(10)** Wekesa a-a-ch-a omweene
Wekesa SM-Tns-go-fv Agr-own
*Wekesa went himself*

Here, Agr-eene co-occurs with the subject of the intransitive verb, but it only helps to mark contrastive focus with the sense that 'it was *Wekesa* and nobody else'. This means that Agr-eene is bound by the subject, but in this case it is only analyzed as an adjunct.

The last type of bound Agr-eene is one that can form a reflexive reading with the subject or object of a clause when it occurs as a prepositional object of a PP. Consider (11) below:

**(11)** Wanjala a- bool-el-a ekholo embakha khu eng'eene
Wanjala SM-told-Ben-fv CL9clan CL9story on Agr-own
*Wanjala told the clan a story about themselves*

In (11), the Agr-*eene* form strictly selects the indirect object *ekholo* as the antecedent mainly because of the shared agreement features of CL9, a pointer to the importance of agreement in any attempt to account for the nature of anaphoric strategies.

There are additional contexts where Agr-*eene* can be bound locally. These include cases where the argument position anaphoric to the subject is embedded in the object, when the anaphoric argument position is in an adjunct, or when the anaphoric argument is the second object of a causativized verb (e.g., the y argument of 'c verb-cause x y') and is dependent on the causal subject. These cases are exemplified in (12a), (12b) and (12c) respectively.

**(12a)** Ba-khulundu ba-ulil-a chimbakha khu babeene
CL2-priest SM-heard-fv CL10-stories on Agr-own
*The priests heard stories about themselves*

**(12b)** Maria a-bon-a e-ndemu enyuma wo omweene
Mary SM-saw-fv CL3-snake behind of Agr-own
*Mary saw a snake behind herself*

**(12c)** Petero a-p-y-a Wanjala omweene
Petero SM-beat-Caus-fv Wanjala Agr-own
*Peter made Wanjala to beat himself*

In (12a), Agr-*eene* is embedded in the object NP as a post modifier to the head noun. In (12b), the adjunct *enyuma wo omweene* contains a PP complement which in turn contains Agr-*eene* as the complement to the PP. *Omweene* in (12c) corresponds to the direct object of the causativized verb, and may be construed with the subject or with an extra-sentential antecedent when Agr-eene is owner.

Agr-eene within a PP also allows a split antecedent as illustrated in the following sentence.

**(13)** Wekesa a-a-kachul-il-a Marko khu beene
Wekesa SM-Tns-tell-Ben-fv about themselves
*Wekesa told Marko about themselves*

Both the subject *Wekesa* and the object *Marko* form the combined antecedent for Agr-eene.

All in all, local binding relations by Agr-eene are quite complex, and must be understood in the context of a host of other factors, some of which include its role as an adjunct or an argument, and the type of verb that it co-occurs with.

### 3.1.4 The Phrasal Reciprocal and Binding Relations

The distribution of the phrasal reciprocal is more restricted than that of Agr-eene. Nevertheless its role is similar to that of Agr-*eene* in so far as it occurs optionally in sentences where it corresponds to an argument that has been bound by the RCM. However, unlike Agr-eene, the phrasal reciprocal only selects a local antecedent. The following generalization helps to capture binding relations by the phrasal reciprocal.

> **Generalization 13**
>
> The phrasal reciprocal always renders a reciprocal reading in conjunction with the RCM, and it does not succeed in fostering an emphatic reciprocal interpretation with a clause mate antecedent if an RCM is absent.

This generalization is illustrated below.

**(14a)** Basoleeli ba-a-siing-an-a babeene khu beene
CL2boy SM-Tns-washed-RCM-fv Agr-own on Agr-own
*The boys washed each other*

**(14b)** *Basoleeli ba-a-siing-a babeene khu beene
CL2boy SM-Tns-washed-fv Agr-own on Agr-own
*The boys washed each other*

**(14c)** Ba-khulundu ba-ulil-a chimbakha khu babeene ne babeene
CL2-priest SM-heard-fv CL4-stories on Agr-own on Agr-own
*The Priests heard stories about each other*

**(14d)** ? Ba-khulundu ba-ulil-a chimbakha khu babeene khu beene
CL2-priest SM-heard-fv CL4-stories on Agr-own on Agr-own
*The Priests heard stories about each other*

(14a) is perhaps the most preferred usage where the phrasal reciprocal serves to enhance the reciprocal reading conventionally marked by the RCM. Its antecedent is therefore the

local subject *basoleeli.* In (14b), the phrasal reciprocal cannot form a reciprocal reading with the local subject when in direct object position.[26]

On the other hand, the phrasal reciprocal in (14c) occurs within the PP adjunct, and is construed with the subject. Notice that the phrase has a '*ne*' preposition. When *'khu'* is used in (14d), the degree of acceptability reduces pointing to a possible difference in interpretation between the two forms of the phrasal reciprocal, one with *'ne'* and the other with *'khu'*.

Again, like the Agr-eene form, the phrasal reciprocal allows a split antecedent and can have an emphatic reading when used with intransitive verbs. Consider the following examples.

**(15a)** Wekesa a-a-bool-el-a Wanjala khu-kachul-a khu babeene ne babeene
Wekesa SM-Tns-tell Wanjala to-talk-fv on Agr-own with Agr-own
*Wekesa told Wanjala to talk about each other*

**(15b)** Basoleeli ba-a-ch-a babeene ne babeene
CL2boy SM-Tns-go-fv Agr-own with Agr-own
*The boys went (and nobody else)*

In (15a), the antecedent to the phrasal reciprocal consists of a combination of both *Wekesa* and *Wanjala* even though they occur in different syntactic positions. In (15b), the phrase marks contrastive focus.

The facts on the phrasal reciprocal reveal that it can occur either as an optional adjunct or as an argument especially in PP complement positions. In addition, it helps to put focus

[26] What looks like an exception to this requirement is in cases where the verb is inherently reciprocal i.e. reciprocal is licensed by a lexically incorporated RCM as in;
Bakhasi ba-a-kanan-a babeene khu beene
CL2woman SM-Tns-meet-fv Agr-own on Agr-own
*The women met each other*

on its antecedent in relation to the action or state expressed in the proposition. Lastly, its interpretation is always with a clausemate antecedent.

Note that in our discussion of local binding relations in this section, we only focused on the relations where the anaphoric element is bound by a subject antecedent in what is conventionally known as subject orientation. However, there are cases where binding is by an element that is not in the canonical subject position, hence putting to question the subject orientation hypothesis. I examine these exceptions next.

### 3.1.5 Exceptions to Subject Orientation

The notion of subject becomes unclear when one considers exceptional case marking (ECM) verbs commonly occurring as epistemic and perception verbs on one hand, and causativized constructions, on the other.

#### 3.1.5.1 Epistemic and Perception Verbs

In LuBukusu, some perception verbs, and some epistemic verbs participate in a construction that is traditionally called 'object raising' or 'Exceptional Casemarking' ((ECM) construction) in the literature of generative grammar (See particularly Haegeman (1991) for a detailed analysis of ECM constructions). In such constructions, the subject of a non-finite clausal complement behaves more like a direct object syntactically, even though it is thematically selected by the complement verb and not by the perception or epistemic verb. The relevant fact in LuBukusu is that the argument corresponding to the clausal complement subject of the embedded verb can appear as an OM, RFM or RCM on the matrix perception or epistemic verb, even though that argument is only selected by the subordinate verb. Examples of this kind with an epistemic verb are illustrated below.

**(16a)** Yohana a-a-e-bukul-a omweene khu-b-a omumiliyu
Yohana SM-Tns-RFM-consider-fv Agr-own to-be-fv CL1-smart
*John considers himself (to be) smart.*

**(16b)** Ba-saani ba-a-bukul-an-a ba-lwaal-a
CL2-men SM-Tns-consider-RCM-fv SM-sick-fv
*The men considered each other sick.*

**(16c)** Ba-saani ba- ba-ikomb-ang-a ba-fw-e
CL2-men SM-OM-expect-Asp-fv SM-die-fv
*The men expect them to die.*

Notice that (16a) is an infinitive complement with the reflexive raised from canonical object position to lower subject position. The lower clause in (16b) has an empty subject position whose agreement features are marked on the SM. Similarly, (16c) also shows agreement with the OM attached to the epistemic verb, on the subordinate verb. It is however (16a) and (16b) that best illustrate the contradiction to the subject orientation hypothesis. The RFM in the former appears on the epistemic verb corresponding to the subject of the complement clause *omweene* and the matrix subject *Yohana*. Similar issues arise for the latter where the RCM on the matrix verb is bound by the two subjects. Such constructions must however not be likened to object control predicates illustrated in (17) below.

**(17)** Yohana a-a-bool-el-a Wanjala khu-tim-a
Yohana SM-Tns-tell-Appl-fv Wanjala to-run-fv
*Yohana told Wanjala to run*

From this example, two ways by which ECM varies from object control are identified. First, object control verbs involve two sets of selectional restrictions, one set assigned to the structural direct object of the matrix verb, and a separate set assigned to the complement clause subject, where the latter is analyzed as a silent argument, PRO, in Principles and Parameters approaches. In (17), *Wanjala* is the direct object of the higher

clause which in turn controls the empty (PRO) lower subject position, so that a coreference chain is formed between the direct object and PRO.

The second difference is where the complement of the epistemic verb in ECM constructions bears subject agreement corresponding to the noun class of the subject of the matrix clause, which would be consistent with the assumption that the object anaphor on the matrix verb corresponds to the subordinate subject, and the object anaphor is in turn bound by the matrix subject. Further still, there are cases where the epistemic or perception verb licenses construal between a RFM on the lower verb with the matrix subject. This is made possible because the lower verb bears agreement features of the subject, and not the object. Compare (18) and (19) below.

**(18)** Wekesa a-a- bool-el-a Wanjala a-e-siim-e
Wekesa SM-Tns-told-Ben-fv Wanjala SM-RFM-likes-fv
*Wekesa told Wanjala to like himself*

Here, the main challenge is whether the RFM is coreferential with the object *Wanjala* in contravention of the subject orientation requirement. The answer lies in what we consider the antecedent to be. It is true that the RFM can be, and is indeed construed with the object, but the clause containing the RFM has to be subjunctive On the other hand, the matrix subject is ruled out as an antecedent mainly because of the nature of the verb. If the subject must be an antecedent, then it requires an epistemic or perception verb such as *'reeba'* (ask), or *'saba'* (request). This is shown in (19) below.

**(19)** Babaana ba-a- reeb-a Wanjala ba-e-siim-e
Cl2 children SM-Tns-told-Ben-fv Wanjala SM-RFM-likes-fv
*The children asked Wanjala to allow them like themselves*

In this case, the object *Wanjala* is ruled out as a possible antecedent mainly because of the mismatch of agreement features between the object and the SM on the lower verb, made possible by the nature of the higher verb. Note that *ba-* is a plural agreement affix.

### 3.1.5.2 Causatives and Subject Orientation

In causativized verbs, the cause of the event described by the verb is added as an arguments is the antecedent. This is especially so in cases where a transitive verb is causativized. If we use symbols to represent the basic structure of transitive verbs as V(x y) where x is an agent and y a patient, then the causativized version V+ Causative (c x y) allows for two possible subjects; both the cause c and the agent x. Consider the following:

**(20a)** Bakhasi ba-a-e-nyw-esy-a Kamabeele
CL2woman SM-Tns-RFM-drink-Caus-fv CL5milk
*The women made themselves to drink milk*

**(20b)** Bakhasi ba-a-nyw-esy-an-a Kamabeele
CL2woman SM-Tns-drink-Caus-RCM-fv CL5milk
*The women made themselves to drink milk*

Notice that these sentences involve only one verb root where the RFM/RCM represent the x argument which ideally does not behave like a subject. The c argument is represented by *Bakhasi* while *kamabeele* is the y argument. In this case x is anteceded by c. The subject orientation hypothesis therefore takes a much more broad definition of subject.

A different scenario arises if the object y is what the RFM/ RCM correspond to. The antecedent will be the OM x argument as shown in (21) below.

**(21a)** Wekesa a-a-mu-i-siim-isy-a
Wekesa SM-Tns-OM-RFM-like-Caus-fv
*Wekesa made him to like himself*

**(21b)** Wekesa a-a-ba- siim-isy-an-a
Wekesa SM-Tns-OM-like-Caus-RCM-fv
*Wekesa made them to like each other*

In (21a) the OM is the agent of the action in the verb whereas *Wekesa* is the cause. The RFM is the patient and selects the x argument as the antecedent and not the c argument. In (21b), the OM attached to the verb is the antecedent. A related case is shown in (22) where the causative construction demotes the force subject to adjunct position and leaves the c argument as the antecedent of the RFM.

**(22)** Wekesa a-a-e-siim-isy-a khu baana
Wekesa SM-Tns-RFM-like-Caus-fv on CL2child
*Wekesa made the children to like *himself/him*
*'Wekesa made himself be liked by the children'*

Here, the cause *Wekesa* is the only available antecedent because the agent appears in an adjunct position as a complement of the preposition in the PP.[27]

From the foregoing discussion, it is evident that the definition of subject as the antecedent must be broadened to include both the agent and cause arguments in causative constructions.

### 3.1.5.3 Comitative Constructions

Comitatives are constructions where a post verbal '*ne*' phrase contains an argument that is semantically interpreted as part of the subject. In these cases a singular structural subject can support a reciprocal reading. Consider the following example:

[27] The realization of *babaana* as *baana* is due to a process of affix deletion motivated by the addition of the preposition *khu*. Note that *baana* is the root form.



UNIVERSITY OF NAIROBI
LIBRARY

**(23)** Wekesa ba-a- p-an-a ne Wanjala
Wekesa SM-Tns-fight-RCM-fv with Wanjala
*Wekesa and Wanjala fought each other*

In this example, *Wekesa* is the structural subject but must combine with *Wanjala* contained in the PP to form an antecedent for the RCM. The indication of this plural antecedent is indicated by the SM on the verb.[28]

Note that the subject+comitative phrase antecedents should not be confused with the usual 'split antecedent' phenomena, where more than one thematic argument can count as an antecedent. For example, split antecedents are permitted for Agr-*eene*, but not for RFM or RCM.

**(24)** Wekesa a-a-kachul-il-a Marko khu beene
Wekesa SM-Tns-tell-Ben-fv Marko on Agr-own
*Wekesa told Marko about themselves*

It might appear in some circumstances that the RFM or RCM on a subordinate verb can take a split antecedent where the absent subject of the subordinate verb corresponds to more than one antecedent in the higher clause, as in the examples below.

**(25a)** Wanjusi a-boleel-a omukhasi o-wewe a-li ya-enyekh-a
Wanjusi SM-told-fv CL1wife Agr-his Agr-that it-supposed-fv
Ba-e-siim-e babeene
SM-RFM-like-fv Agr-own
*Wanjusi told his wife that they are supposed to love themselves*

---

[28] The RFM is also possible in comitative constructions of this kind, but the comitative does not add to the subject to create a plural antecedent when agreement is singular. Rather there are two varieties singular reflexive readings, one which adds to the object of the verb and the other which is interpreted as instrumental.

a) Wanjala a-e-siim-a ne Wafula
Wanjala SM-RFM-like-fv with Wafula
'Wanjala and Wafula like themselves' or
'Wanjala likes himself and he also likes Wafula'

b) Wanjala a-e-khiing-a ne omwaana
Wanjala SM-RFM-defended-fv with CL1chilld
'Wanjala shielded himself with a child' or 'Wanjala and the child defended

**(25b)** Wanjusi a-boleel-a omukhasi o-wewe a-li ya-enyekh-a
Wanjusi SM-told-fv CL1wife Agr-his Agr-that it-supposed-fv
Ba-siim-an-e babeene khu beene
SM-like-RCM-fv Agr-own on Agr-own
*Wanjusi told his wife that they are supposed to like each other*

This appears, however, to be a case of pro-drop, where a null pronominal subject is picking up the split antecedent, and the null subject pronoun is inducing agreement on the subordinate SM, with the result that the RFM and RCM are in fact locally anteceded by a plural SM in the lower clause. The fact that the subject of the lower clause could be any human plural in the discourse, not necessarily the antecedents in the matrix clause, provides additional evidence for the view that the null subject acts like an ordinary pronoun.

Returning now to antecedency in the comitative construction, one might try to argue that the real antecedent for the RCM is the agreement marker (SM) on the verb. The idea is that whenever the class of the SM does not match that of the structural subject, it is the agreement morpheme that determines the acceptable interpretations for the reciprocals. The comitative construction in LuBukusu thus raises interesting questions about how antecedents are calculated when the unique structural antecedent available does not appear to do the job.

## 3.2 Non- Local Binding Relations

We have already noted that when the anaphoric elements occur in argument positions in a clause, they often trigger local binding susceptible to principle A effects. This is especially true for the affixal markers. In addition, when the affixes co-occur with free lexical anaphoric forms, the antecedent of both is also local, but the free forms play an

emphatic role. Since, the affixes and the phrasal reciprocal are always local, then non-local relations will only involve Agr-eene. In LuBukusu, two categories of non-local Agr-eene are identified. Consider the following example:

**(26)** Wamalwa a- siim-a omweene
Wamalwa SM-likes-fv Agr-own
*Wamalwa likes him*

The Agr-*eene* form is in the object position of a transitive verb and behaves like a pronoun in the sense that it is not bound by any NP within the clause. It is therefore subject to principle B effects.[29]

Further still, there exists a form of Agr-*eene* that is in an argument position, but is unbounded. I refer to it as a focus costruction determined by discourse factors. It is able to pick out any focused entity in the domain of discourse either sentence internally or externally. Consider (27) below:

**(27)** Wamalwa a-a-kachul- il- a Wanjala khu omweene
Wamalwa SM-Tns-spoke-App-fv Wanjala on Agr-own
*Wamalwa talked to Wanjala about himself*

In this example, the Agr-*eene* form either has an extra- sentential antecedent or can select either *Wamalwa* or *Wanjala*. This means that there is some kind of competition between

[29] What looks like an exception to (26) is where the subject is either a 1st or 2nd person pronoun with an agreeing Agr-eene form in the object position as shown in (a) and (b) below:

a) Ese na-a-siim-a samweene
I SM-Tns-like-fv Agr-own
I like *me/myself

b) Ewe wa-a-siim-a wamweene
you SM-Tns-like-fv Agr-own
you like *you/yourself

In both sentences, the Agr-eene form can be interpreted as being coreferential to the clausemate subject even when there is no RFM. Why this succeeds is mainly attributed to the referential specifity of 1st and 2nd person pronouns.

the different possible antecedents. How does a native speaker then tell the intended antecedent in communication? I content that the answer to such a question largely depends on logophoric factors which have to do with point of view and shared knowledge of the context, together with syntactic factors relating to c-command. The representation of the varied interpretations provide an interesting analysis that may open new ground in the way anaphoric relations are computed in language. I defer this question to chapter four. I turn to an important part of anaphora: Focus.

## 3.3 Focus and Anaphora

Focus is a term used in linguistics to refer to emphasis placed on elements perviously mentioned in a sentence or in a larger discourse context by way of giving new information. In language, the emphasis may be by means of prosodic effects such as intonation or stress, or by giving an extra element such as a word or affix. Sometimes, the emphasis may be contrastive by implying the exclusion of other elements that may have occupied the focal position. In this section I concentrate on focus triggered by the presence of anaphoric elements mostly in conjunction with each other, sometimes creating quite complex combinations. Recall that in section 3.1 we noted cases where the Agr-eene form is focal on its own when it occurs in intrastive constructions. I discuss the most common combinations next.

### 3.3.1 RCM + Agr-eene + Phrasal Reciprocal

The RCM can combine with the lexical anaphoric forms to create various patterns each with a specific point of emphasis. Consider the following examples;

**(28a)** Bakhasi babeene ba-a-yeet-an-a babeene khu beene
Cl2 women Agr-own SM-Tns-help-RCM-fv Agr-own on own
*Women themselves helped each other*

**(28b)** Bakhasi ba-a-yeet-an-a babeene khu beene, babeene
Cl2 women SM-Tns-help-RCM-fv Agr-own on own Agr-own
*Women themselves helped each other*

**(28c)** Bakhasi babeene ne babeene ba-a-yeet-an-a babeene khu beene
Cl2 women Agr-own on Agr-own SM-Tns-help-RCM-fv Agr-own on own
*Women themselves helped each other*

**(28d)** Bakhasi ba-a-yeet-an-a babeene khu beene, babeene ne babeene
Cl2 women SM-Tns-help-RCM-fv Agr-own on ownAgr-eene on Agr-eene
*Women themselves helped each other*

In (28a), Agr-eene occurs with the subject *Bakhasi* in the subject position. In this case it helps to put emphasis to the head noun in a kind of contrastive focus. In other words, the interpretation is such that it was the women, and nobody else that helped each other. A related interpretation is where the women helped themselves without anybody else's assistance.The phrasal reciprocal on the other hand helps to enhance the reciprocal reading by putting focus on the participant in the proposition that receives the reciprocal action. The same relation is evident in (28b) only that here, the contrastive Agr-*eene* is in the final position. In (28c) and (28d), the phrasal reciprocal with *'khu'* is predictably focal, while the one with *'ne'* is contrastive only that the intensity of emphasis is further enhanced by the apparent reduplication of Agr-eene. From this discussion, a more specific generalization regarding the RCM + Agr-eene + Phrasal reciprocal co-occurrence is made in (2) below:

**Generalization 14**

The phrasal reciprocal is always a focus construction enhancing reciprocal reading whenever it co-occurs with the RCM. Agr-*eene* or a phrasal reciprocal with '*ne'* can only add contrastive focus when added to the clause.

### 3.3.2 RFM + Agr-eene + Phrasal Reciprocal

The RFM can also form complex co-occurrence possibilities with varied interpretations. Consider the following:

**(29a)** Bakhasi babeene ne babeene ba-e-yeet-a babeene
Cl2 women Agr-own on Agr-own SM-help-fv Agr-own
*Women themselves helped themeselves*

**(29b)** Bakhasi ba-e-yeet-a babeene, babeene ne babeene
Cl2 women SM-RFM-help-fv Agr-own Agr-own on Agr-own
*Women themselves helped themselves*

**(29c)** Bakhasi babeene ba-e-yeet-a babeene
Cl2 women Agr-own SM-RFM-help-fv Agr-own
*Women themselves helped themselves*

**(29d)** Bakhasi ba-e-yeet-a babeene, babeene
Cl2 women SM-RFM-help-fv Agr-own Agr-own
*Women themselves helped themselves*

Similar issues arise for the reflexive. The RFM is always the reflexive marker, but the reflexivity can be enhanced by an Agr-*eene* form that immediately follows the verb. This is shown in all the examples. In addition, there is a possibility of Agr-*eene* forms serving contrastive focus either within the subject NP or at the end, as shown in (29a) and (29b) for the former, and in (29b) and (29d) for the latter. We can capture these possibilities in (3) below:

**Generalization 15**

Agr-eene is always a focus construction enhancing reflexive reading whenever it co-occurs with the RFM. Agr-eene or a phrasal reciprocal with '*ne*' can only add contrastive focus when added to the clause.

### 3.3.3 RFM + RCM

Co-occurrence between RFM and RCM in a single clause would appear to be unexpected, since both markers may be thought to be competing to represent the single direct object argument slot. However, there are a number of cases where these affixes do co-occur. This co-occurrence appears to be productive, which suggests that at least one of the affixes does not require it to represent an argument of the verb stem on which it occurs. Consider the following examples:

**(30a)** Ba-khasi ba-e-yeet-an-a
CL2-woman SM-RFM-helped-RCM-fv
*(The) women helped themselves/each other*

**(30b)** Ba-khasi ba-e-fumy-an-a
CL2-woman SM-RFM-praised-RCM-fv
*(The) women praised each other*

In order to understand the difference between the two sentences we take the group *Bakhasi* to be an entity consisting of members x, y, z...n, such that, the RCM without a RFM will always have the possibility of say x helping y, y helping z, and so on, but there is no possibility of say x helping x. Slightly different is the case in (30) where apart from x helping y, there is also the possibility of x helping x. In addition, the RFM may serve an emphatic role expressing the idea that it was the women themselves who helped each other. The antecedent for both elements remains the subject. These finer grained judgments raise interesting questions for the semantic construction of reflexive and reciprocal meanings. Why would the fully distributed reciprocal predictably trump the reflexive and make its role secondary? Which is the argument structure for the two anaphoric elements since they appear to be competing for the same argument position? Such questions are quite justified especially in light the relation between transitivity and

the distribution of the RFM and the RCM. It seems that co-occurrence between RFM and RCM in a single clause would appear to be unexpected, since both markers may be thought to be competing to represent the single direct object argument slot.How then does such a conflict exist in the same environment? Perhaps the only explanation would be that the two roles play out in some kind of hierarchy beginning with the most economical i.e reciprocal marking then followed by the reflexive that only reiterates the boundaries within which the former unfolds. The role of the RFM is therefore more pragmatically determined. It is similar to the role played by contrastive Agr-eene in a reciprocal construction.

### 3.3.4 RFM + Oblique Argument

The RFM can cooccur with Agr-eene that is an oblique argument. Cosider the example below:

**(31)** Yohana a-e-lom-a khu omweene
Yohana SM-RFM-spoke -fv on Agr-own
*John spoke about himself*

Without the RFM, Agr-eene may have an extra-sentential antecedent. However its presence ensures coreference with the subject antecedent. Its role is hence disambiguating. The following important questions about the nature of RFM/ Agr-eene cooccurrence arise: i) If the co-occurrence is optional, are the anaphoric relations formed by the the independent forms in complementary distribution? ii) If it is obligatory, what difference does it make in terms of variation in interpretation? In order to answer these questions, consider the following:

**(32a)** Yohana a-e-lom-a khu omweene
Yohana SM-RFM-spoke -fv on Agr-own
*Yohana spoke about himself*

**(32b)** Yohana a-lom-a khu omweene
Yohana SM-spoke -fv on Agr-own
*Yohana spoke about him*

**(32c)** Yohana a-e-lom-a
Yohana SM-RFM-told -fv
*Yohana told himself*

**(32d)** Yohana a-e-lom-a kho
Yohana SM-RFM-spoke -fv on
*Yohana spoke about himself*

In (32a), the RFM and Agr-*eene* cooccur, only Agr-*eene* occurs in (32b), while in (32c) and (32d), the RFM is realized alone, but with the latter having a latent NP in the complement position of the PP. In order to determine whether both the RFM and Agr-*eene* are obligatory or not, we need first to identify the argument structure of the verb – *loma.* In LuBukusu, this verb is either ditransitive or intransitive, but with varied meaning reminiscent of English 'speak' and 'talk'. Since the ditransitive meaning is intended in (32a), then an agent and a goal are required for structure completeness.

This requirement is fulfilled with the presence of *Yohana* and Agr-*eene*. The RFM only ensures reflexivity with the subject. In this sense both are obligatory but if subject reflexivity is not the issue, then Agr-*eene* alone is sufficient in (32b), in which case its reference will be with an external antecedent much like a pronoun is in similar contexts. The situation is slightly different in (32c). Without a VP internal object, the verb has to change meaning. It then follows that the RFM and Agr-*eene* may co-occur in ditransitive verbs where both are obligatory, but with the latter functioning as an oblique object. If it

is in VP internal position, the RFM alone is obligatory. Agr-*eene* will then only be emphatic. Without the RFM in the ditransitive context, Agr-*eene* has an extra-sentential antecedent.

There are cases where the RFM/ Agr-eene co occurrence is unacceptable, especially in contexts where the direct object refers to a different entity other than the subject. Consider the example below.

**(33a)** *Maria a-a-e-bon-a e-ndemu enyuma wo omweene
Mary SM-RFM-saw-fv CL9-snake behind of Agr-own
*Mary saw a snake behind her*

**(33b)** Maria a-a-bon-a e-ndemu enyuma wo omweene
Mary SM-Tns-see-fv CL9-snake behind of Agr-own
*Mary saw a snake behind her*

In (33a), the presence of the RFM makes it ungrammatical because the direct object *'endemu'* is different from the subject antecedent. When the RFM is removed in (33b), the sentence becomes grammatical. Here, Agr-*eene* is embedded in the sentential adjunct as a complement of the PP. It is hence obligatory.

### 3.3.5 Multiple Cooccurrences in Causative/ Applicative Constructions

There are quite productive cases where more than one RCM is possible, only on condition that there is always a causative or applicative affix.

**(34)** Ba-baana ba-a-p-an-il-an-a kumukaati
CL2-child SM-Tns-fight-RCM-Appl-RCM-fv CL3-bread
*Children fought for bread for each other*

Once again, questions arise as to what effect the doubled RCM could have on the argument structure of the verb. One strategy would be to regard cases like these as

morphophonological reduplications unrelated to argument structure, but there is a semantic effect in these instances (focus), which probably requires further scrutiny. The fact that the RCM permits multiple occurrences on a single verb stem, but the RFM does not, suggests the RCM is more likely to be the affix with a less consistent effect on verb valence. The most logical explanation therefore is that of semantic expansion via applicative/ causative extensions. Consider:

**(35a)** Babaana ba- a-p-an-il-a kumukaati
CL2children SM-Tns-fight-RCM-Appl-fv CL3bread
The children fought (each other or other people) for/ with bread

**(35b)** Babaana ba- a-p-an-il-an-a kumukaati
CL2children SM-Tns-fight-RCM-Appl-RCM-fv CL3bread
*The children fought (each other or other people) for bread for each other*

**(35c)** Emuuna ya-a- siim-an-y-a chisaang'i
CL9Squirrel SM-Tns-fight-RCM-Caus-fv CL10animals
*The squirrel caused the animals to like (each other)*

**(35d)** Emuuna ya-a- siim-an-isy-an-a chisaang'i
CL9Squirrel SM-Tns-fight-RCM-Caus-RCM-fv CL10animals
*The squirrel caused the animals to like each other*

In (35a) and (35b), the initial RCM is the conventional one i.e. corresponding to the subject or some other people not mentioned in the sentence. The second RCM in (35b) is motivated by the applicative to show the beneficiary as the children (not anybody else). In fact this RCM is in complementary distribution with a free NP so long as it specifies a beneficiary. It is therefore possible to specify such a beneficiary by adding an NP before *kumukaati* as in (36) below.

**(36)** Babaana ba- a-p-an-il-a basakhulu kumukaati
CL2children SM-Tns-fight-RCM-Appl-fv CL2old man CL3bread
*The children fought (each other or other people) for bread for old men*

In (35c) and (35d), the story is similar. In the former, the RCM corresponds to the object. The animals are the ones made to like. This is the same argument referred to by the initial RCM in (35d), with the second clearly specifying the object of the liking. It was each other and not anybody else. If one wants to show the object of the liking as somebody else other than the animals, then an NP can be put after *chisaang'i.*

**(37)** Emuuna ya-a- siim-an-isy-a chisaang'i ne babandu
CL9squirrel SM-Tns-fight-RCM-Caus-fv CL10animals with CL2person
*The squirrel caused the animals to like people*

Perhaps, also, an explanation for Appl/ Caus Vs single RCM ordering can be derived from such explanations. When an RCM precedes an applicative, the sense of 'verb- for oneself or others/with' is expressed. On the other hand, if the applicative comes first, 'verb- for benefit of subject' sense is expressed. When the causative precedes the RCM, the sense of 'verb- cause subject' is implied. When the order is changed, the causation affects persons other than the subject.

## 3.4 Verb-specific anaphoric Relations

There are limited classes of cases of lexical reflexives and lexical reciprocals, that is to say, cases where a verb can be understood to be either reflexive or reciprocal in the absence of either the RFM or RCM. I consider each of these next.

### 3.4.1 Verb Specific Reflexive Relations

These verbs are understood reflexively, yet they do not have an overt marker for such reflexivity. Consider the following:

**(38a)** Wekesa a-a-bek-a
Wekesa SM-Tns-cut-fv
*Wekesa shaved his hair*

**(38b)** Wekesa a-a-saab-a
Wekesa SM-Tns-wash-fv
*Wekesa washed himself*

In (38a) the verb for shave has an inherent reference to the subject of the sentence implying that the shaving was done to *Wekesa.* In (38b), the verb is also reflexive.

### 3.4.2 Verb Specific Reciprocal Relation

There are also some verbs that are lexically reciprocal, and these have an RCM-like morpheme that is attached to the verb root where the root has no meaning on its own, or has a different meaning. This is exemplified in (42) below.

**(39a)** Wekesa a-a-ingan-a ne mawe
Wekesa SM-Tns-argue-fv with his mother
*Wekesa argued with his mother*

**(39b)** Bakhasi ba-a-kanan-a
CL2woman SM-Tns-meet-fv
*The women met each other*

The verb *inga*, in (39a), literally means 'stretch' while *ingana* means 'argue'. The verb -*akanan*- 'meet', is an example of a verb that has no independent meaning if the apparent RCM -*an*- is subtracted from it. Also, no RCM can occur with such verbs except the one that already exists in the root.

## 3.5 Conclusion

This chapter focused on the question of how the anaphoric elements discussed in chapter two are able to achieve binding relations. It was noted that the relations are either local or non-local. The RFM and the RCM, for example are exclusively local i.e. they are bound by a c-commanding clause mate subject. I addition, they trigger both the sloppy and strict

reading under VP ellipsis. On the other hand, the binding relations involving Agr-*eene* and the phrasal reciprocal are quite flexible with the former having greater flexibility. For example it can refer to an exclusively local antecedent or an extra-sentential one. The former is, for example, possible when it cooccurrs with the RFM, or is in a PP complement position. In other contexts, its reference is long distant. The phrasal reciprocal on the other hand, supports a reciprocal reading in conjunction with the RCM. In other environments, it is a focus construction. There are however exceptions to the subject orientation hypothesis for local binding specifically in ECM, causative and comitative constructions which prompt a redefinition of the concept of 'subject'. Focus constructions are also discussed. The basic idea is that co-occurrence of anaphoric elements creates emphatic or contrastive focus. For example RCM and RFM cooccurrence leads to a reciprocal reading with a reflexive emphasis. Lastly, inherently reflexive and reciprocal verbs are discussed.

# Chapter Four

# Theoretical Analysis of Anaphoric Elements

## 4.0 Introduction

This chapter deals with the representation of LuBukusu anaphoric elements within the Minimalist framework. In order to set the stage for the analysis, a recap of the procedure of sentence analysis within MP is given in section 4.1. In section 4.2 the affixal markers, which include the RFM and RCM are examined in details with their varied representations given in justified tree diagrams. A similar analysis is extended to the non-affixal markers in section 4.3, more specifically Agr-eene and the phrasal reciprocal. The chapter conclusions are given in section 4.4.

### 4.1 The Minimalist Program Revisited

The basic principles of MP are discussed in chapter one. In this section, the focus narrows down to the actual analysis of sentences, culminating into how anaphora is dealt with in MP. To set the stage for the analysis, I revisit some of the key assumptions of MP, below.

It has already been noted that Phrase structure representations in minimalism are built up step-by-step as set forth by the Bare Phrase Structure (BPS) model, which, among other things, has the following basic properties:

(1) Its structure is derivational. That is, it is built from the bottom up, bit by bit.

(2) It does not have a preconceived structure.

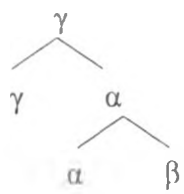
(3) It only has binary branching.

(4) It does not distinguish between a "head" and a "terminal".

(5) It operates with two basic operations, Merge and Move.

According to Chomsky, Merge is a function that takes two objects and merges them into an unordered set with a label. The label identifies the properties of the phrase. This is shown in the abstract structure below.

(1)



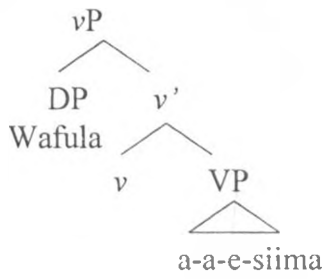

Here, γ corresponds to a general label specifying the identity of the whole structure. As noted in section 1.9.3, this is some kind of verb phrase (known as light *v*P), a functional category present to implement movement of the external argument to [Spec, *v*P] position. In order to set the stage for a more detailed analysis, I use sentence (2) below to illustrate what is expected.

**(2)** Wafula a-a-e-siim-a
Wafula SM-Tns-RFM-like-fv
*Wafula likes himself*

In minimalism, structure building proceeds step by step, beginning with the initial step that leads to the formation of the lower VP via merging the V and the object in transitive constructions, and subsequently followed by merging the external argument in its the

position i.e. the [Spec, *v*P] position, licensed by the *v* head.[30] This is represented in the following structure

(3)

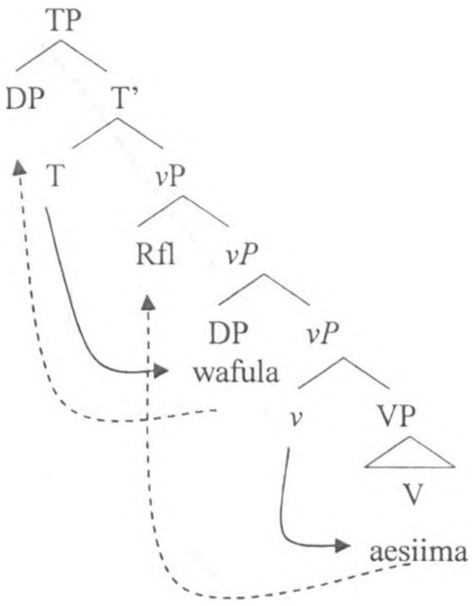



The next step involves merging and movement operations aimed at assigning case. Note that *v* is merged with an accusative case feature that requires a DP with which it matches via a form of agree relation. In the structure, there is no object DP in the canonical VP position, instead the RFM serves the role of the object, and must therefore be analysed as such. In the initial merge, the subject DP is excluded, first, because it is in its θ- position, and, secondly, the *v* only searches down the tree for a DP that it can agree with, and it only finds the RFM which in turn raises to [Spec, *v*P].

The last step involves merging of T which has a case feature that also requires phrasal movement of a DP lacking such a feature. In this case the object cannot raise to T because it no longer has an uninterpretable case feature (its case was checked via *v* in the preceding merge derivation). This scenario then leaves the subject as the only DP that can be attracted by T. These facts are represented in the structure below:

[30] The Predicate Internal Subject Hypothesis (PISH) requires the external argument to be generated in the specifier position of the v head with which they enter into a theta relation.

(4)




Note that the Minimal Link Condition (MLC) licenses the raising of the subject DP across the object because they are in the same minimal domain and also since the object's case is already checked it is no longer visible for further checking. With this background in mind, I now turn to the representation of the anaphoric elements in LuBukusu.

## 4.2 Analysis of the Affixal Markers

As noted in the introduction, the affixal anaphoric elements in LuBukusu include the RFM and the RCM. I examine each in turn.

### 4.2.1 The Reflexive

In the representation of affixes, it is important to remember that some of these affixes represent complete arguments with case and theta roles, and must hence be fully represented as such. The RFM is one such affix (see section 2.1.2.4 on the grammatical role of the RFM). How then is such crucial information represented in this model? As

stated in the aforementioned MP claims, categorical information such as this is either functional or lexical. Given that the latter is already represented, then it follows that the affixes require functional nodes that will copy the inherent information carried by such affixes in the syntax (semantic and agreement properties), in movement procedures aimed at implementing agreement of interpretable and uninterpretable features in the configuration. In representing this information, the starting point should at least be a structure that reflects the Bantu morpheme ordering template, as discussed in Baker (1985). In what follows, I review some of Baker's key assumptions and use them to give an initial representation of the RFM.

According to Baker's Mirror Principle, morphological derivations must directly reflect syntactic derivations, and vice versa. This kind of thinking stems from certain key assumptions about both morphology and syntax. For morphology, no distinction is made between derivation and inflection. In addition, morphological processes apply to a given form one at a time, in a cyclic manner working from the inside, systematically moving outwards. This means that affixes that are closer to the verb occur before those that are further. On the other hand, syntax is the level of description where semantic/ thematic relationships are explicitly represented. Consider the structures in (5) below for illustration.

**(5a)** Wekesa a- mu-siim-a
Wekesa SM-OM-like-fv
*Wekesa likes him*

**(5b)** Wekesa a- siim-a Wanjala
Wekesa SM-like-fv Wanjala
*Wekesa likes Wanjala*

One thing is clear from the two sentences: The OM and the NP 'Wanjala' have the same

semantic relationship to the verb i.e. they have a similar theta role and case marking. They are both semantic objects. Let me assume that since LuBukusu has an SVO word order, then the basic representation of these sentences would be as in (6).

**(6)** NP1 VERB NP2

Where NP1 and NP2 correspond to the subject and object respectively. Beyond such a semantic level, syntax also can be identified with the surface level of description related to what actually is said in terms of position and content. The OM in (5a) is therefore a surface prefix. The surface level representation for (5a) and (5b) would then take the form in (7a) and (7b) respectively.

**(7a)** NP1 NP2 VERB

**(7b)** NP1 VERB NP2

This line of argument perfectly corresponds to the two levels of LF and PF proposed in MP. Further, the representation of any structure at the semantic level need not correspond to the same structure at the syntactic level. Whereas the semantic level in (6) captures both sentences, that of (7) shows a clear distinction between the two levels of representation.

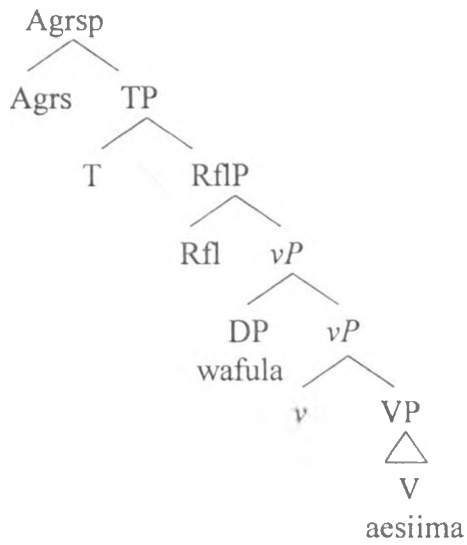
Because syntactic derivations match the morphological ordering, an overview of the affix order in LuBukusu is crucial as a basis for the proposed analysis.

In LuBukusu, and indeed in most Bantu languages, the VP is represented by the following template where parentheses indicate optionality.

**(8)** VP→ (Neg) SM T (A) (M) (OM) V (Extensions) FV

Consequently, a potential representation for (1) is shown in (9) below, excluding the semantic position of the object DP. Note however that in subsequent representations, I largely reject the Agr approach for reasons that become clear in the course of the discussion.

**(9)**




Going up from the VP, a functional category RflP is created to check the features of the RFM. It does not, in this case, take the affix because it occurs substantively elsewhere i.e as an affix attached to V. Again in conformity with the VP template, T is selected above the ReflP to represent tense realized as –a-, but which is assimilated in the RFM in a phonological process. Finally, the AgrsP is inserted corresponding to the external argument.

Whereas the minimalist idea that functional heads are created to check the features on lexical items in the course of the computation is sustained in (9), a number of issues seem

to call for a different approach. Naturally before we propose such an approach there is need to highlight the shortcomings of the above analysis.

Perhaps most crucial, and relating to methodology is the apparent congestion of the structure with the introduction of new labels, AgrsP and ReflP that, in a way reduplicate the functions played by [Spec, TP] and [Spec, *v*P]. Recall that the external argument surfaces in [Spec *v*P] as its theta position and moves to [Spec TP] attracted by the T head for feature checking. On the other hand, the internal argument is attracted by *v* to outer [Spec *v*P] for accusative case checking. The proposal of extra functional categories is not only redundant, but also less economical.

Secondly, and more empirical are the contradictions and the evidence inherent in the data. As mentioned in chapter 2, the RFM is licensed in environments where the OM is excluded, implying that the two affixes play similar roles; as objects of transitive verbs. In addition, recall that in LuBukusu, the occurrence of the OM/RFM precludes the occurrence of the lexical object in the VP internal position. This means that whenever such an NP occurs in the same clause as the OM/ RFM, then the NP is not in an argument position, and is hence motivated probably by semantic factors, and therefore its landing site for syntactic derivation should be motivated by such semantic reasons. In addition, the OM and RFM should have similar representations that can capture the common language facts. The representation in (9) fails to do this.

The structure however makes an important distinction regarding the RFM in relation to a possible Agr-eene DP in post verbal position. Notice that there is no category for objects to check the uninterpretable case features associated with object DPs because case has

already been assigned to the RFM through agree relations with the *v*. Therefore, Agr-eene is simply adjoined to the VP as an adjunct or emphatic category. With these preliminaries in mind, I turn to the reflexive in clausal domains.

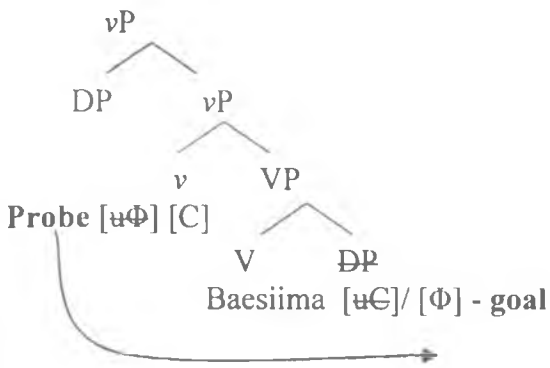
#### 4.2.1.1 The Reflexive in a Clausal Domain

I extend the analysis of the RFM with the most commonly occurring context of declarative clauses. Consider the sentence below:

(10). Babaana ba-e- siim-a
CL2children SM-RFM-loves-fv
*The children love themselves*

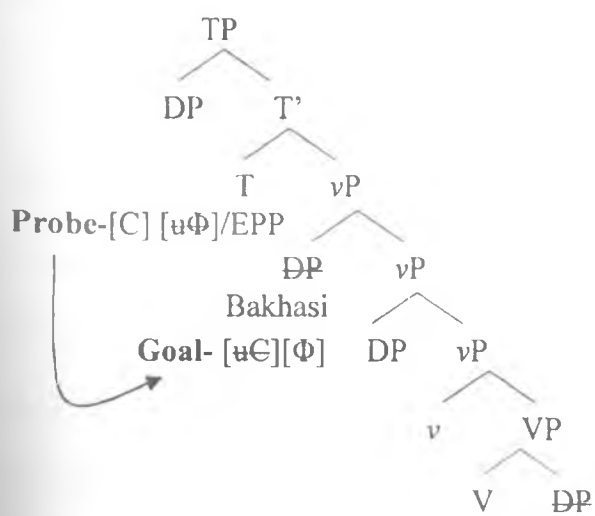
The first step in the derivation is to merge the RFM, present in the internal VP position, with the verb *basiima* not forgetting that the object DP has an uninterpretable case feature but interpretable Φ-features. In the second step, V is merged to the structure and since it lacks values for its Φ-features, it acts as a probe and searches for an active goal to match. Such a goal is found in the object DP which is active by virtue of having [uC]. This DP then moves to [Spec *v*P] to satisfy the case checking requirement that it is checked under specifier- head relations. The two match and agree hence becoming inactive. This is illustrated below:

(11)




The next step will, predictably, involve the merging of the external argument leading to the formation of the first phase. Again, we notice that the subject DP lacks a value for its case feature but has valued Φ-features. The said DP thus remains an active goal for a searching probe. Such a probe is found after merging T to the first phase. Note that T has unvalued Φ-features and thus becomes the probe that finds a matching goal in the active subject DP. The two match and agree leaving them with no unvalued features. However the subject DP must move to the [Spec,TP] position due to an EPP feature i.e. the need to check unvalued case features, which find value via T. These steps are illustrated below:

(12)



Note that in the last step the computation is complete and all the unvalued features have found value and are therefore interpretable, and hence are deleted. Also note that the motivation for the varied structural notations is entirely empirical (Adger 2008).

How do we account for the binding relations? The general assumption has always been that pronouns and reflexives are distinct categories often appearing in complementary distribution and are distinguished by pairs of opposing features identified as [+anaphor] for reflexives and [+pronominal] for pronouns. Covert categories PRO and pro also exist, and each is also identified according to the presence or absence of the stated features. See Chomsky (1982) for detailed discussions of such distinctions. In more recent literature, however, +/- anaphor is not a primitive feature and must therefore not be used as a distinguishing feature for pronouns and anaphors. Instead, as Zwart (2002) argues, the lexicon has only one type of pronoun, a root pronoun. It then follows that the difference between a pronoun and an anaphor is syntactic in the sense that the anaphor, unlike the pronoun, lacks values for certain features, whose identity has been a subject of debate recently.

Reinhart and Reuland (1993) for example claim that anaphors lack a feature R, for referential independence while pronouns have the feature. Further, they claim that an NP is +R iff it carries full specification for Φ-features and structural case. Following from this argument, reflexives cannot therefore lack a Φ-feature – it must be something else. Let us see how we can work out what the feature really is. Given a sentence;

(12) Wekesa a- e- bon- a
Wekesa SM-RFM-saw-fv
*Wekesa saw himself*

The derivation will proceed in such a way that the initial merge will involve the transitive verb *bon-a* with a functional category (a form of DP) in the VP internal position to implement the initial merge derivation. Recall that the creation of such a functional category is in line with the assumption that languages are uniform in their core level of representation. The idea is that a DP category can always be selected if the subcategorization of the V allows, and when this is the case then its normal position is VP internal in SVO languages. This proposition is in tandem with MP thinking that a DP is itself a complement of a phase head from which it inherits the ability to agree with its complement and assign case. The DP is marked as [uC] [Φ] and therefore it becomes an active goal for a probing *v* that has no value for Φ-features. The next step will involve the merging of the external argument leading to the formation of the first phase. As noted above, the subject DP lacks a value for its case feature but has valued Φ-features. The said DP thus remains an active goal for a searching probe- T that is merged to the first phase. Further, after the computation, the subject DP must move to the [Spec:TP] position due to an EPP feature i.e. the need to check unvalued case features, which find value via T. The computation is shown in (13) below with a little simplification.

**(12)** [TP Wekesa$_i$ T[*v*P t$_j$[*v*P t$_i$ *v*[VP V aebona DP$_j$]]]]

Note the use of subscripts is exceptionally to indicate the positions of copies left by moved categories. I also leave issues of how the RFM ends up on the verb open, but following Baker (1988), I argue for an independent morphological process of incorporation (See chapter 2 for details on how incorporation works out in LuBukusu).

Further. following from Reinhart and Reuland (1993), the unvalued feature of the DP probably has something to do with reference. However, this cannot be the whole story because anaphors can be bound by non-referential expressions, as in (14) where the interrogative element *naanu* is the only accessible antecedent.

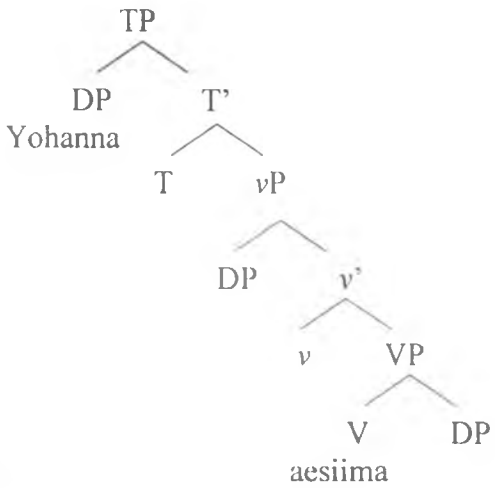
(14) Naanu o- wa- e- bon- a ?
Who SM-Tns-RFM-saw-fv
*'who saw himself/herself?'*

Within MP, an economy condition could be the reason for having an element with unvalued features. According to Chomsky (2001 b), a feature without a value is more economical than a feature with a value. The economy here is assumed to arise from the fact that an unvalued element normally gets value in the syntactic derivation. I will refer to this feature as F in line with Heinat (2006). Recall again that in GB, NPs in a given sentence are either bound or free. They are bound if they are coindexed and one C-commands the other. In MP, coindexing is replaced by an agree relation between an antecedent and a pronominal expression (incorporated or free). In other words, if the pronominal and the antecedent can enter an agree/checking relation then they are coreferential. Deductively, if there is no agree relation the pronominal expression and the 'antecedent' are not coreferential. If we consider the example above, the copy of the reflexive has an unvalued feature [F] necessitating the formation of an agree relation within the syntactic derivation with the external DP that has value for its referential feature. Nevertheless, the difference between anaphors and pronouns mainly arises from the fact that the latter's [uF] gets value from the syntactic derivation. Let me consider how this works out in the following sentence:

(15) Yohanna a-e-siim-a
Yohanna SM-RFM-like-fv
*Yohanna likes himself*

Putting all the stages aside, the computation process will yield the representation in (16) below.

(16)

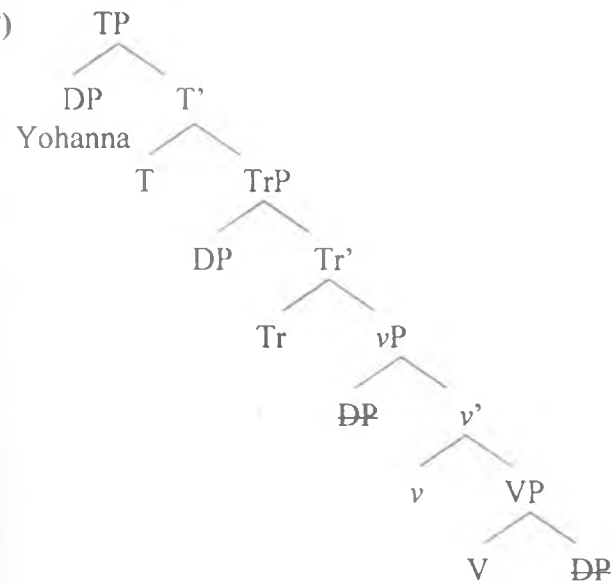



Notice that I have intentionally left out the inner [Spec *v*P] for the external argument, but it should still be at the back of our mind that the external DP raises from this position to [Spec TP] because of feature checking.

Further, in order to account for the presence of the RFM within the VP, I consider the traditional movement analysis (Kayne 1975, 1989) which assumes that affixes are generated VP internally and later on in the derivation adjoin to the inflectional head. In this kind of analysis, the affix adjoins to *v* leaving its copy in the VP internal position. Note that this kind of analysis can be extended to the OM because of the inherent asymmetry of the syntactic functions of the two affixes. In other words, all transitive constructions have the internal argument merging with the head of inflection (*v*).

With the foregoing argument in mind, it is possible then to draw a clear distinction between transitive constructions with affixes, and those with full NPs, using an analysis where the features of the object DP (whether lexical or affixed) are checked on the head of a TrP (Transitive phrase) and a distinction made between the two forms via an EPP feature such that if the object DP is lexical then it raises to [Spec, TrP], conversely, if it is an affix then no such raising takes place. Consequently, the structure in (16) will take the form in (17) below:

(17)




Note that this analysis also receives support from the fact that D- Structure in MP is no longer formulable and therefore it is no longer a requirement for arguments to be merged in thematic positions. This means that affixes/clitics are arguments which are merged in the position where they surface while their nominal features can attract aspectual features subject to ordinary movement. In other words, given the copy theory of movement, there are certain circumstances in which a moved element may be pronounced in a position lower than the one in which it has raised because independent requirements of the

morphology or phonology. In LuBukusu for example, it is a requirement that affixes join with the verb stem at PF.[31]

After the syntactic requirements are satisfied, there remains the task of accounting for the binding relations between the RFM and its antecedent that will differentiate it from the non dependant pronominals.

Recall the assumption that all pronominal expressions have an uninterpretable referential feature [uF] that must find value perhaps at LF or at the conceptual- intentional interface (where LF touches other cognitive systems) as a condition for referential completeness (as opposed to well formedness). Such a distinction is made when one considers the English structures in (18) below.

**(18a)** *John* Killed *himself*

**(18b)** John killed *him*

**(18c)** **John* killed *him*

**(18d)** *John killed *himself*

Following Safir (2005), I use the italics to indicate co reference. In this case nothing in the syntax makes (18c) and (18d) ungrammatical except that the coreferential indexing is unacceptable in contravention of the conventional binding conditions. On the other hand, (18a) and (18b) are well formed because of both syntactic and coreference issues. *John* and *himself* refer to the same entity in (a), while *him* is free in (b). If we follow the proposal that all pronominal expressions have uninterpretable referential features which must find value in a context either defined by syntactic or other factors, then the RFM

[31] In the present analysis, all projections of v head are a kind of vP. We therefore use vP instead of TrP, but the underlying idea is still sustained

will defer from an OM on the basis of feature valuation. Whereas the former is valued in a syntactic configuration, the latter requires no such valuation. A puzzle however remains as to the status of the unchecked features of the free pronominal. Will it not contravene the MP requirement that all unvalued features must be checked for the derivation to 'converge' at spell- out. I will suggest here that [uF] on pronominals comes in two or even in three forms, with each form specifying the context of valuation. The first type values the RFM in a local environment when it attaches to the *v* head specified by the antecedent. The second has value outside the sentence, while the third caters for the valuation of the UD- forms. In this study [uF] at the end of the computation will imply that the lexical head containing this feature is a free pronominal, while the absence means coreference. This kind of thinking provides adequate grounds for a reformulation of the BT principles based on feature checking.

**(19) Binding Theory (Preliminary version)**

*Principle A:* If α is an anaphor it must have a valued F at the end of the computation.

*Principle B:* If α is a pronoun it must have uF at the end of the computation.

*Principle C:* If α is an R- expression it has no F.

Notice that principles A and B correctly predict the difference between anaphors and pronouns in terms of feature valuation, a key component of MP. The binding domain has however not been specified, but the assumption is that in a phase based theory such as the one I am using in this study, feature checking takes place within a phase domain, and therefore valuation of F naturally implies that the probe must be a c- commanding antecedent. Further, the principles as now formulated, will exclude the generation of

structures such as (18c) and (18d) because, the pronoun *him* will have uF in the former, while the anaphor *himself* will have F.

Viewing binding relations within a minimalist framework also has an additional theory internal advantage. Indices are excluded as a core representation of binding relations in conformity to the Inclusiveness Condition (Chomsky 1995:228) stated as follows;

> A perfect Language should meet the condition of inclusiveness: any structure formed by the computation [...] is constituted of elements already present in the lexical items selected for [the numeration] N; no new objects are added in the course of the computation apart from rearrangement of lexical properties (in particular no indices [...]).

I set aside more comments on BT principles for now, but I turn to them in the last section of this chapter. Meanwhile I examine a different context of RFM occurrence- The complex noun phrase.

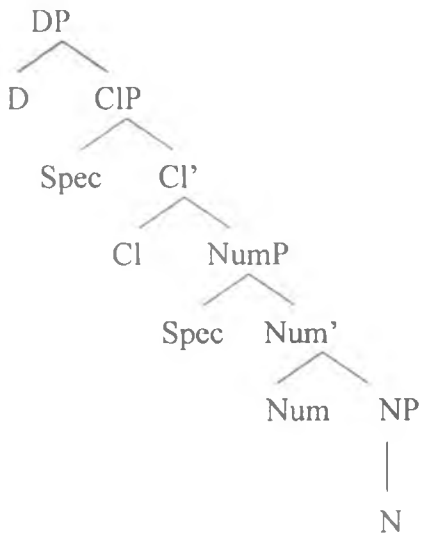
#### 4.2.1.2 The Reflexive Marker in Complex Noun Phrases

The context of occurrence and the licensing condition were discussed in chapter two. The purpose of this section is to extend the proposal of representation to account for the RFM in complex NP contexts. To do this properly, there is need to understand the general representation of NPs in the syntax of Minimalism. Certain assumptions emerge about the general structure of NPs: (a) Movement is only constrained to the left in a conventional 'bottom- up' configuration; (b) Only one adjunction is allowed per head or maximal projection; (c) There are spell-out rules converting features like class and number to phonetic form realized at PF.

The implication of such rules is that an NP ideally follows similar computation as a

clause: A 'bottom up' approach, binary branching and derivation by rule. Carsten (1991) suggests that the main features of an NP, class and number are represented by functional projections only there to implement movement for purposes of feature checking. A typical NP in LuBukusu would then contain the following structure:

(20)

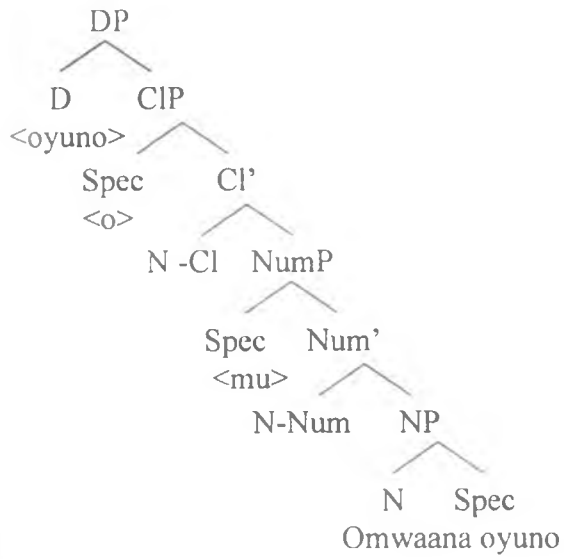



The main motivation for proposing such a structure is both empirical and theoretical. For the former, issues of word order come into play, while the latter considers theoretical assumptions in MP. It has been suggested in the literature (see mainly Abney 1987, Tang 1999, Alexiadou 2001) that the maximal projection for a typical noun phrase is a DP category whose complement is the NP, with a host of parametric functional categories determined by respective languages. The most crucial of such categories are Number (Num) and Class (Cl), proposed to check the respective features of the noun before spell-out. Consider a simple NP in (21):

(21) O-mu-aana o-yu-no
CL1-Num-child spec-Agr-this
*This child*

The basic derivation should merge the N to the determiner, to give rise to a DP, in line with Chomsky's Bare Phrase Structure (Chomsky, 1995). Further, the features on both the head and the determiner require functional categories in the syntax to check them in the course of the derivation. I propose that these categories are ClP and NumP, which are incidentally marked on N. Given one of the facts stated above on movement being only leftward, the underlying positions of these categories are therefore to the left of the NP, whose head is used to check the features on the respective FC heads. The implication here will be that the DP will be separated from the NP by the FCs, mainly because the features on the D head are checked via a Spec- head relation i.e by movement of the spec of NP to D. All these processes will then lead to the derivation of (21) as (22) below:

(22)



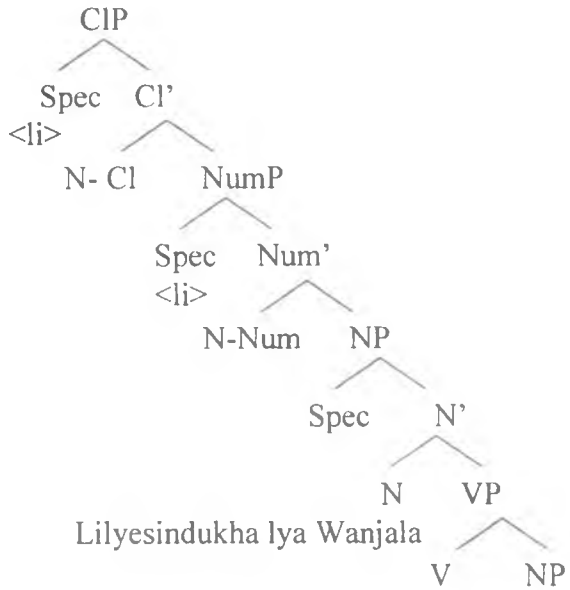

Notice that in LuBukusu, the affix features are already assigned to the noun when it enters the numeration, and hence the noun never moves out of NP in the overt component. The movement is only covert, hence the brackets on the features. This is crucial because the word order can then be easily derived at PF. In fact according to the

Linear Correspondence Axiom (Kayne 1994, Chomsky 1995), linear order now only applies at PF. With this structure in mind, I now turn to derived nouns that also host the RFM.

As already seen, the context in which a nominal self reflexive occurs in LuBukusu always contains a derived noun as the head with a non- nominal category whose head should be adjoined to the N head. Following Kayne (1994), such adjunction should be to the left. This is because the head will project the rest of the morphology from the right and in the process allow the nominalization process to take place from the left. The structure of the complex nominal in (23) below is shown in (24).

**(23)** Li-ly-e- siim-a lya Wanjala
Cl3-Num-RFM-like-fv of Wanjala
*Wanjala's self liking*

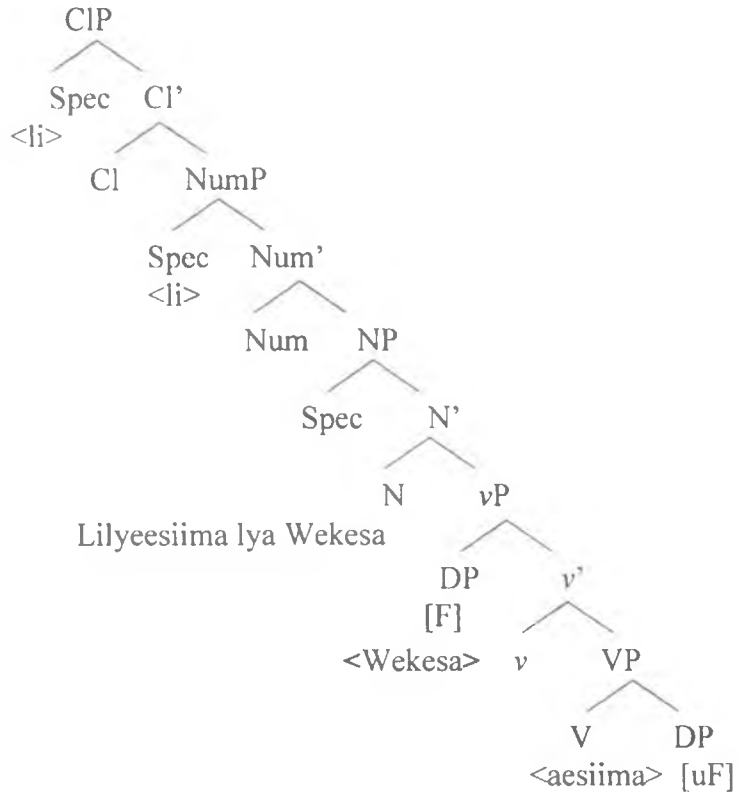
**(24)**




The process that leads to the formation of the complex NP is quite complex. However there is no doubt that the derivation starts in the NP. Since the whole structure is defined

according to the head N, then the first step would be to determine the derivation of this head. Notice that there is a non- nominal category labeled V attached to the head that appears to provide the base for nominalization, and since the V lacks the number and class features associated with the noun, it moves left ward to the N to receive these features and is hence deverbalised. The resultant noun then projects like any other noun. Notable is also the fact that the verb contains an RFM that is unvalued for coreference. This is the most significant point to consider in this study. The usual question of how coreference is implemented, specifically within the NP, comes to mind. I propose a procedure that is based on two crucial steps in the process of nominalization: a) pre nominaliztion, and b) post nominalization. The former deals with a verbal context that is responsible for licensing the RFM via coreference, while the latter is concerned with licensing features within the surface NP. The first implication of this analysis is the unification of the domain of coreference in both clausal and phrasal categories where the domain is standardly a *v*P, that has both the anaphor and its antecedent. The details of this kind of procedure for the sentence in (23) are shown in the representation in (25) below.

(25)




All the details aside, the most important claim is that deverbalized nouns are underlyingly pure clauses with a normal clause derivation, only this time the procedure takes place at the semantic level before deverbalization. Notice that the uninterpretable F feature on the lower DP finds value via coreference with the external argument of the v head. The only surface realization is on the N after nominalization has taken place. The main advantage for this kind of analysis obviously lies in the unification of the context in which binding relations occur, in both clausal and phrasal contexts. In other words, the vP remains the uniform domain allowing the application of Principle A of BT as reformulated in (19) above. The same kind of analysis can be extended to the analysis of the RFM in infinitival constructions. In the next section, I briefly examine how this analysis is implemented.

### 4.2.1.3 The Reflexive Marker in Infinitival Constructions

A similar analysis can be extended to Infinitival constructions where the infinitival *khu* is analysed as a class 15 noun class marker with a reduplicated number affix occurring next to the RFM. A representation of (26) is shown in (27).

(26) khu- khw-e- siing-a khwa Wanjala
Cl15-Num-RFM-wash-fv of Wanjala
*'Wanjala's washing of himself'*

(27) [ClP <khu> Cl[NumP<khu>Num[NP [N'khukhwesiinga[*v*P
<Wanjala>[*v' v*<aesiinga>[VP V NP]]]]]]]

In summary, the prenominalization step involves the verbal structure that has a verb which licenses deverbalization, and hence helps to implement coreference relations between the RFM and its antecedent, prior to the surface realization after nominalization takes place with the merging of the verb with N.

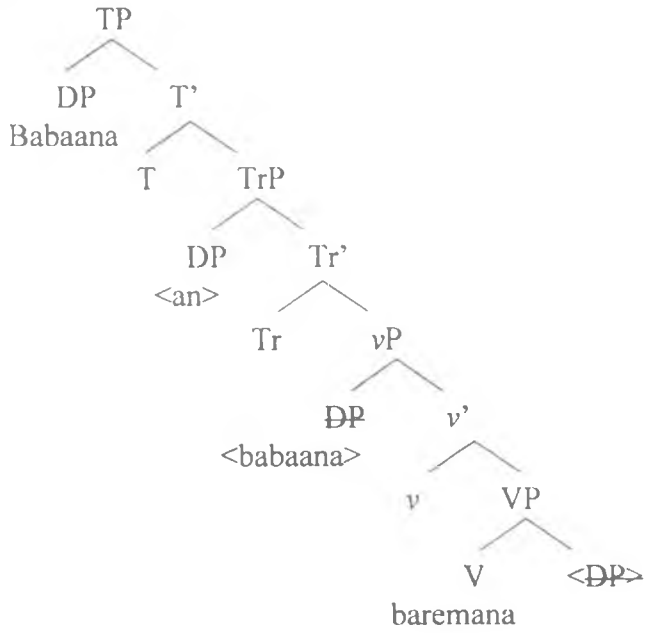
Having made clear generalization about the representation of the RFM, I next turn to the RCM, which essentially patterns in the same way as the RFM, and therefore should be subject to the same kind of analysis.

### 4.2.2 The Reciprocal

In order to understand the representation of the RCM, I recap several generalizations discussed in more details in Chapter two. First, recall that the RCM is relatively fixed in terms of morphophonological representation. It consistently occurs as a suffix –an, not only in LuBukusu but also in various other Bantu languages. Secondly, there is a variant form –chan- occurring mainly in verbs with monosyllabic stems ostensibly to guard against a possibility of heteromorphemic clusters of sounds. Thirdly, the RCM can establish relations in both clausal and phrasal contexts all of which involve a form of

check the features of the moved NP. The computation for (28) will therefore take the form in (29) below.

(29)




Recall that since D- Structure in MP is no longer formulable, it is no longer a requirement for arguments to be merged in thematic positions. This means that because affixes/ clitics are arguments which are merged in the position where they surface and their nominal features can attract aspectual features subject to ordinary movement, the RCM in (29) begins from the internal VP position fully specified with nominal features which include interpretable person and number features and uninterpretable case. So when the Tr head searches down for a matching head it only finds this argument. They both merge and the uninterpretable features are valued, with the semantic DP moving to the [Spec TrP] position. However, given the copy theory of movement, there are certain circumstances in which a moved element may be pronounced in a position lower than the one in which it has raised because of independent requirements of the morphology or phonology. In this case the RCM merges with the verb at PF, and this is the position where it is

pronounced. The next step involves the valuation of [uF] attached to the RCM. The only accessible subject with this feature is the subject DP *babaana* which then becomes the antecedent. This is in line with the assumption that all pronominal expressions have an uninterpretable referential feature [uF] that must find value perhaps at LF as a condition for referential completeness. The application of Principle A of BT is correctly predicted.

Notice that this sort of analysis is similar to the one proposed for the RFM, because as you may recall, we noted that both the RFM and RCM motivate the same kind of interpretations where the internal argument is merged, in terms of reference, to the external argument. I return to the comparison between these two elements in section 5.1 Presently, I examine contexts where the RCM contains what appears to be an object

In LuBukusu, the RCM can correspond to an inalienable possessor in a gestural or 'quasi-reciprocal' construction as in (30) below.

**(30)** Babaana ba-a-rem-an-a kimikhono
Children SM-Tns-cut-RCM-fv CL6hand
*The children cut each other's hands*

In this example, the RCM motivates what looks like an indirect object corresponding to the possessum of the subject, and which appears in the post verbal position. The issues that arise here involve assumptions about how lexical semantic arguments are projected onto syntactic structures. One thing is clear: it is not obvious that the RCM can be seen as a detransitivizing element if it can occur with what appears structurally to be an indirect object. It is possible, perhaps, to argue that the same kind of representation as in (29) applies to (30) where the RCM triggers a direct object in internal VP position, which then raises to TrP motivated by feature checking. Subsequently, the referential features are

checked on the subject antecedent.

The status of *kimikhono* (though already characterized as an object) is a subject of debate. This question is addressed in details in chapter two. The conclusion drawn is that this is some kind of indirect object. The most relevant question however is the relation between the RFM and such a DP. Strictly speaking, it is the possessum that suffers the action of the verb, and hence should be the direct object, but this is only a part of a possessor, whose reference is similar to that of the subject. It is therefore possible to have an NP *kimikhono kyewe* (his hands) forming the direct object, reminiscent of the English structure *The children cut each other's hands* where the reciprocal is bound by the subject DP. The RCM also occurs in phrasal constructions. I examine this context next.

#### 4.2.2.2 The Reciprocal Marker in Complex Noun Phrases

Recall that like the RFM, the RCM also occurs in nominalized phrases where the form marking reciprocal still remains [an]. As noted earlier, the productivity of the RCM with different noun classes is quite limited to only a few classes mainly classes 5, 14 and 15 as exemplified below respectively.

**(31a)** Lili-p-an-a lye babaana
CL5-fight-RCM-fv of CL2child
*The children's fight of each other*

**(31b)** Bubu-p-an-i bwe babaana
CL14-fight-RCM-fv of CL2child
*The children's fight of each other*

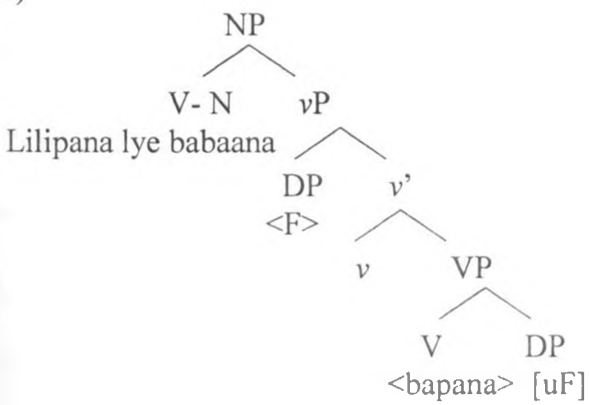
**(31c).** Khukhu-p-an-a
CL15-fight-RCM-fv of CL2child
*Fighting of each other*

As was the case with the RFM, the purpose of this section is to extend the proposal of NP

representation to account for the RCM in complex NP contexts as well. Certain assumptions have already been highlighted about the general structure of NPs in LuBukusu: (a) The NP is analyzed as a DP projection with intermediate functional categories there to check nominal features such as number, and class terminating with the N head; (b)Movement is only constrained to the left in a conventional 'bottom- up' configuration; (c) Only binary branching is allowed; (d) Spell-out rules convert affixes into PF features realized in a position below the position of merge. All these characteristics are exemplified in (22) above, and so I will not repeat them here. Instead I focus on nominalization that contains the RCM.

Setting aside the full representation of categories within the DP complex, I solely focus on the derivation that computes the binding domain for the RCM. The most important claim is that deverbalized nouns are underlyingly pure clauses with a normal clause derivation, only this time the procedure takes place at the semantic level before deverbalization. The representation of (31a) will then take the form of (32) below:

(32)




In this structure, the deverbalized noun starts out as a V contained in the VP, and fully specified for theta role and case features. Because of this, the usual issues of FCs headed by *v* that implements feature checking. Notice that the uninterpretable F feature on the

lower DP finds value via coreference with the external argument of the v head. The only surface realization is on the N after nominalization has taken place.

The representation of (31b) and (31c) will follow a similar pattern. The only problem is encountered when it comes to checking of uF in (31c). Notice that there is no overt agent of the reciprocal structure. This problem can however be overcome if one argues that the agent is interpreted as being arbitrary or indefinite. It is this arbitrary subject that acts as the antecedent to the RCM, and by extension helps to value the uF on the RCM.

In summary, the representation of both the RFM and the RCM assumes similar derivations whether in a clausal or phrasal complex. This is made possible by both theory internal and empirical factors. For the former, a unified analysis greatly boosts the descriptive powers of any theory, while for the latter, the two affixes seem to both trigger the presence of an internal argument which helps to explain binding relations in terms of a common feature checking configuration hence replacing the old principles of the Binding theory. I turn to the non- affixal markers next.

### 4.3 The Non-Affixal Anaphoric Elements

As discussed in chapters two and three, certain other elements function anaphorically in LuBukusu yet they are not affixes. These forms, which I call non-affixal anaphoric elements, normally constitute Agr-eene and the phrasal reciprocal, each with its own patterning and representation. I take each of them at a time.

#### 4.3.1 Representation of Agr-eene

Just to recap: LuBukusu, as noted in section 2.3 has a free pronominal NP labeled *Agr-*

*eene*, and is realized in different positions either optionally or obligatorily, but on all occasions it must have an antecedent that is either in the syntax or in the extended discourse. I propose that the form that Agr-eene takes in relation to whether it is optional (in which case it is an adjunct), or obligatory (where it is a complement) greatly determines its theoretical representation within MP.

Further recall that the Agr-*eene* form is employed in a variety of ways. It can be used as an emphatic adverbial reflexive, it can serve to emphasize a reflexive reading, and in certain circumstances it can form a reflexive reading on its own, or still be used as the principal argument literally expressing possession. I examine each of these categories next, with a view to determining the inherent differences in terms of their respective representation within a minimalist framework.

#### 4.3.1.1 Emphatic Adverbial Agr-eene

We noted in chapter two that this is perhaps the most common use of the Agr-eene form. Consider some of the examples repeated here as (33).

**(33a)** Wekesa omweene a-ch-il-e
Wekesa Agr-own SM-go-Tns-fv
*Wekesa himself went*

**(33b)** Wekesa a-ch-il-e omweene
Wekesa SM-go-Tns-fv Agr-own
*Wekesa went himself*

**(33c)** Omweene a-ch-il-e
Agr-own SM-go-Tns-fv
*He himself went*

In all cases, the Agr-eene form helps to put some kind of contrastive focus on the antecedent *Wekesa* in (33a) and (33b), while in (33c) the emphasis is with a missing NP

marked by the SM on the verb. In this last case, the reference of the Agr-eene may or may not be possessive. If it is possessive then the form is itself an argument.

Further, the adverbial usage is realized when Agr-eene co occurs with the RFM where the usage is meant to emphasize the reflexive reading. However the focus in this case is not contrastive. Consider:

**(34)** Wekesa a- e-siim-a omweene
Wekesa SM-RFM-loves-fv Agr-own
*Wekesa loves himself*

The Agr-eene form here helps to enhance the reflexivity introduced by the RFM. The two anaphoric elements are both dependant on the subject antecedent.

On the basis of these facts, the question that needs consideration concerns the role of MP assumptions in mapping such dependence relations motivated by Agr-eene forms. The first step towards answering this question is to characterize Agr-eene as used here among the syntactic categories. In the literature, three categories; Specifier, head and complement form the core of syntactic analysis. However, because Agr-eene discussed here is an adverbial then it can only be categorized as an adjunct. Hornstein et al (2005:195) while commenting on the status of adjuncts note:

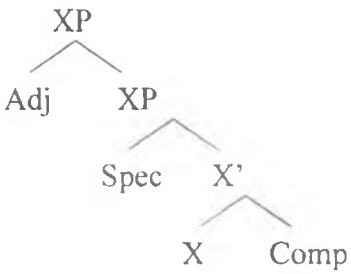
> How to deal with adjuncts is a vexed problem within generative grammar, one that has never been adequately resolved. The properties of adjuncts are quite different from those of complements and specifiers.

This problem lies in some of the key characteristics of adjuncts, for example adjuncts appear to have different case requirements, come in different category types and are semantically interpreted as conjuncts. It is also not clear what features are checked under merger by adjunction, and lastly, what type of relation exists between them and the

elements that they modify. I do not attempt to resolve these gaps, but rather to suggest how Agr-eene adjunct ought to be represented in MP, and hope that other linguists are motivated to focus more on the general phenomenon of adjunction with a view to shading more light on its place in MP framework.

Haegeman (1994) notes that adjuncts are like being in a balcony where you feel you are both inside and outside the house. What this metaphor implies is that adjuncts should be included in the projection of categories, but hold a more peripheral relation, and may just be exempted from the goings on in the 'house'. Because of this characteristic, the representation of adjuncts takes the form in (35) below.

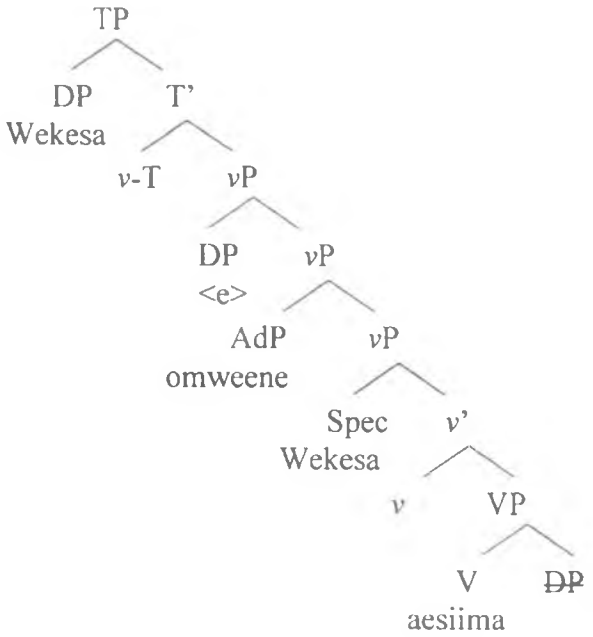
**(35)**




The basic requirement has been that because of the more distant relation that an adjunct has to the rest of the projection, it should not change the label and bar level of the general structure, but should also be a constituent of the lager structure. This basic contradiction does not help matters at all. I will retain this structure henceforth. Notice that one crucial advantage is that it does not change the label XP. Also an adjunct should be a phrasal projection whose label is specified by the type of adjunct in question.

Let me actualize this representation with a concrete illustration with example (34) which derives the structure in (36) below.

(36)




A significant distinction is made regarding Agr-eene occurring as an adjunct with the RFM which serves as the canonical object. This distinction lies in the representation obtained in the tree diagram. Because Agr-eene is an adjunct, nothing predicts that it must be bound in any given syntactic domain. On the contrary, the internal argument realized as RFM must be bound, by a c-commanding antecedent. Indeed this relation is achieved when the object argument moves to the outer [Spec *v*P] (analyzed as [Spec TrP] in the earlier analysis). The post verbal position is not specified because adjuncts are assumed to be quite flexible, and therefore do not determine word order.

In addition, the dependence status of Agr-eene vis-à-vis the binding of RFM is explained in syntactic terms. For the former containement configurations are required, while dominance deals with the latter. When an element is dependent, it requires an antecedent either sentence internally or externally. All these characteristics explain the relative flexibility of Agr-eene as discussed in details in chapter two. Similar arguments can be

extended to the representation of the phrasal reciprocal that is also largely adverbial. I discuss the details in section 4.3.2 below. Meanwhile, I complete this section with an analysis of Agr-eene in argument positions, which include the reflexive form in oblique object positions, and the possessive form.

#### 4.3.1.2 Agr-eene in Argument Positions

As already noted, Agr-eene can occur as a possessive construction in varied argument positions, or as an oblique object where the PP containing it is either an adjunct or a complement. These occurrences are illustrated in (36) below.

**(37a)** Yohana a-e-lom-a khu omweene
Yohana SM-RFM-spoke -fv on Agr-own
*John spoke about himself*

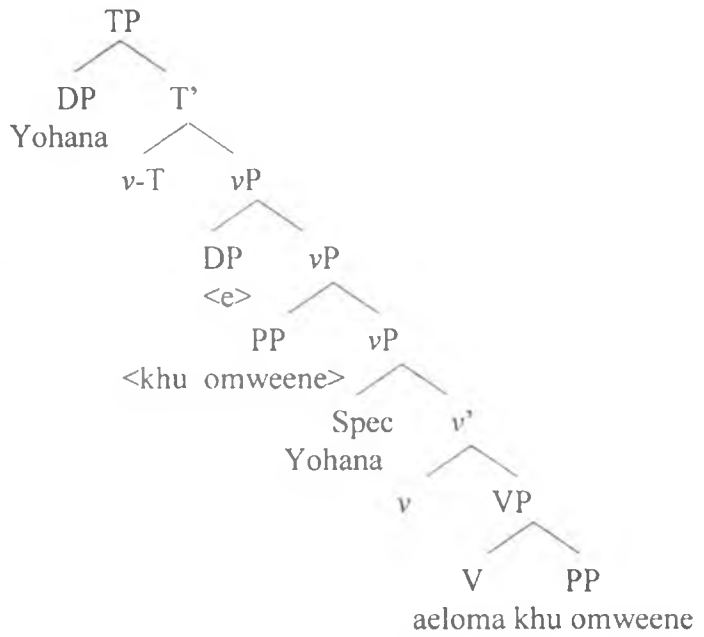
**(37b)** Yohana a-lom-a khu omweene
Yohana SM-spoke -fv on Agr-own
*John spoke about him*

**(37c)** Omweene a-lom-a khu Yohana
Agr-own SM-spoke -fv on Yohana
*The ownwer spoke about John*

In (a), Agr-eene is only interpreted as being co referential to the subject *Yohana* and must therefore be subject to locality constraints. In (b) the absence of the RFM make Agr-eene unbounded in addition to referring to an extra sentential possessor. (c) is an example of possessive Agr-eene in subject position. In terms of representation, (a) is of more interest because of the need to establish the binding domain. The last two follow the representations already discussed for both internal and external arguments respectively. The implication of this sort of approach is the clear distinction between various types of Agr-eene forms. I illustrate how (37a) is derived in a minimalist configuration.

Naturally, the starting point is to point out the differences between this sort of structure and those already discussed in order to determine the best way forward. Notice that there is co occurrence between the RFM and Agr-eene, but unlike in the earlier cases, Agr-eene is obligatory and occurs as an oblique object of a PP complement. The implication is that there are two internal arguments, one represented by the RFM and the other by the PP. The representation is shown in (38) below.

(38)




The binding of the RFM is by the c- commanding antecedent whereas that of Agr-eene is by the RFM, and by extension the subject. In other words, the occurrence of the RFM licenses a bound Agr-eene in oblique object position. It is also important to note that the movement of the PP complement is licensed by the need to check dative case.

From the foregoing, Agr-eene is classified as a dependant form which can either be bound or unbounded. It is bound if it occurs as an oblique object in an argument position

corresponding to the RFM. I conclude that it is the former that is subject to local binding relations. I examine the phrasal reciprocal next.

### 4.3.2 The Representation of the Phrasal Reciprocal

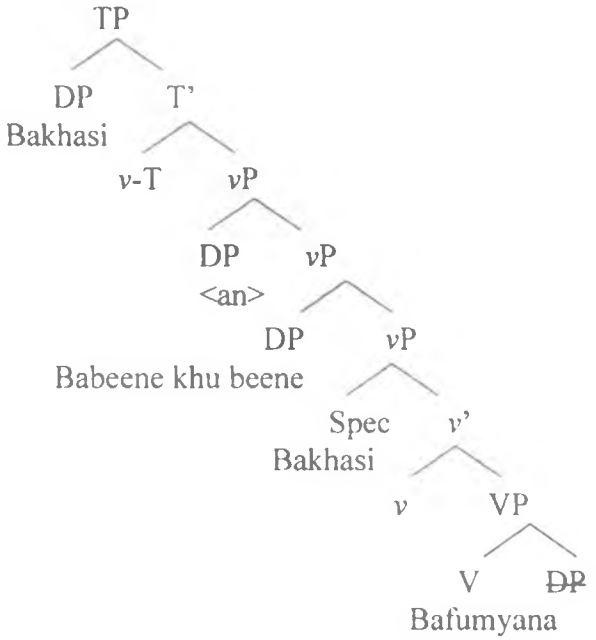
The phrasal reciprocal in LuBukusu is less constrained than the Agr-eene form. The following characteristics summarize the general occurrence of the phrasal reciprocal.

a) Two forms are realized: Agr-eene on Agr-eene (henceforth R1- Agr-eene) and Agr-eene with Agr-eene (R2- Agr-eene).
b) R1-Agr-eene is only licensed in reciprocal constructions in conjunction with RCM to reiterate the reciprocity.
c) R2-Agr-eene occurs in any other environment that allows contrastive focus or emphasis on a given NP that functions as the antecedent.
d) R2-Agr-eene has the characteristics of a pronoun, and therefore can occur in canonical pronoun positions, usually with either a sentential or an extra-sentential antecedent.

Note that in all these cases the phrasal reciprocal is adverbial and hence not subject to binding relations. If we follow a similar argument as that used for Agr-eene, then we may argue that the phrasal reciprocal is only a dependent form, whose antecedent is obtained from within the sentence or from outside. It is also susceptible to an analysis that treats it as an adjunct. Consider the sentence in (39), represented in (40).

**(39)** Ba-khasi ba-fumy-an-a babeene khu babeene
CL2-woman SM-praised-RCM-fv Agr-own on Agr-own
*The women praised each other*

(40)




Because the phrasal reciprocal is an adjunct, nothing predicts that it must be bound in any given syntactic domain. The only prediction relates to it being a dependant form whose interpretation is supplied by the subject. The only element whose referential features are interpreted is the RCM.

## 4.4 Conclusion

This chapter deals with how MP is able to account for anaphora facts. The basic model used is DbP. For all the anaphoric elements, it is assumed that their interpretation takes place in the semantic component (LF). The whole computation is induced by phases and regulated by the twin operations of merge and move. The structure presupposed is the *v*P hypothesis which is headed by a *v* functional category that implements the move operation licensed by the presence of uninterpretable features on both the probe and the goal. The RFM and RCM are licensed to move to spec *v*P position because being active goals to a searching *v* probe. They both have uninterpretable case features. In addition,

they value their uninterpretable R feature on their antecedents satisfying binding relation. As a result, BT principles are reformulated along these lines as follows:

*Principle A:* If α is an anaphor it must have a valued F at the end of the computation.

*Principle B:* If α is a pronoun it must have uF at the end of the computation.

*Principle C:* If α is an R- expression it has no F.

In addition, this section identifies two forms of non- affixal markers based on the syntactic binding. Except for Agr-eene in oblique argument positions, all the other forms are analyzed as adjuncts which are only present to do what adjuncts do best: Modify given syntactic elements, in this case NPs. This position is predictable from the MP representations that specify a local binding relation for the bound Agr-eene alone excluding the others.

# Chapter Five

# Co-occurrence and Anaphora

## 5.0 Introduction

As noted in chapter 3, the single anaphoric elements often co-occur with each other and with the anaphoric lexical elements forming quite complex anaphoric patterns. This ability to co-occur is licensed by different reasons which include: a) the need for emphasis; b) The use of verbal extensions such as causative and applicative markers. This chapter focuses on the theoretical representation of anaphoric relations that involve co-occurrences. Recall that in 4.3.1 and 4.3.2 I argued that whenever the non-affixal markers co-occur with the RFM or RCM, the former will always be emphatic, and are hence realized as adjuncts whose representation has already been dealt with. However, the scenario is quite different when the affixal markers co-occur, with or without the non-affixal markers. As shown in 3.3, RFM/RCM co-occurrence triggres an apparent competition for a single argument slot, since the two elements are analyzed as objects of a transitive verb. In this chapter, I show that this competition is only apparent, since one of the affixes only helps to reiterate the other in an affix referential hierarchy, and must therefore be analyzed as an adjunct. I do this in section 5.1. Further, I show that if RFM/ RCM co-occurrence is also licensed by argument increasing strategies, there is no argument slot competition, and the two represent different arguments, with independent representation in MP. I extend the same argument, in 5.2, to cases of multiple RCM or RFM co-occurrences in which case some of the reduplicated forms are adjuncts, and are therefore analyzed as focus elements. In addition, I develop a representation of lexical anaphoric elements in focus positions. The chapter conclusions are given in section 5.3.

## 5.1 Cooccurrence of RFM and RCM

I have already noted that the RFM and RCM can occur together due to the need to create emphasis or as arguments licensed by the presence of argument increasing affixes. I argue that such licensing factors play a great role in determining the subsequent representations of the resultant structures. Consequently, two categories of RFM/RCM co-occurrences emerge, and they form the basis for the remaining part of our discussion in this section.

### 5.1.1 RFM/RCM Co-occurrence without Applicative/ Causative

When the RFM co-occurs with the RCM for purposes of creating focus, the structure will primarily have a reciprocal reading in addition to a more secondary reflexive reading. Consider the following example:

**(1)** Ba-khasi ba-a-e-yeet-an-a
CL2-woman SM-Tns-RFM-help-RCM-fv
*(The) women helped each other (themselves)*

In this sentence, the interpretation is primarily reciprocal. The role of the RFM is to put focus on *bakhasi* to mark the boundary of reciprocal marking, in an interpretation represented as 'the women themselves helped each other'. In other words, this looks like a case of focus by means of an affix. In KiSwahili, the use of the RFM is unacceptable, and instead such emphasis is achieved by a free lexical NP equivalent to Agr-eene as shown in (2) below:

**(2a)** Wanawake wenyewe wa-li-saidi-an-a **Kiswahili**
CL2women Agr-own SM-Tns-help-RCM-fv
*The women themselves helped each other.*

**(2b)** *Wanawake wa-li-ji-saidi-an-a
CL2women SM-Tns-RFM-help-RCM-fv
*The women themselves helped each other*

Notice that whereas (2a) is acceptable, the addition of the RFM in (2b) makes it ungrammatical. The most important question however is how MP assumptions can be used to represent sentences with RFM/RCM co-occurrence.

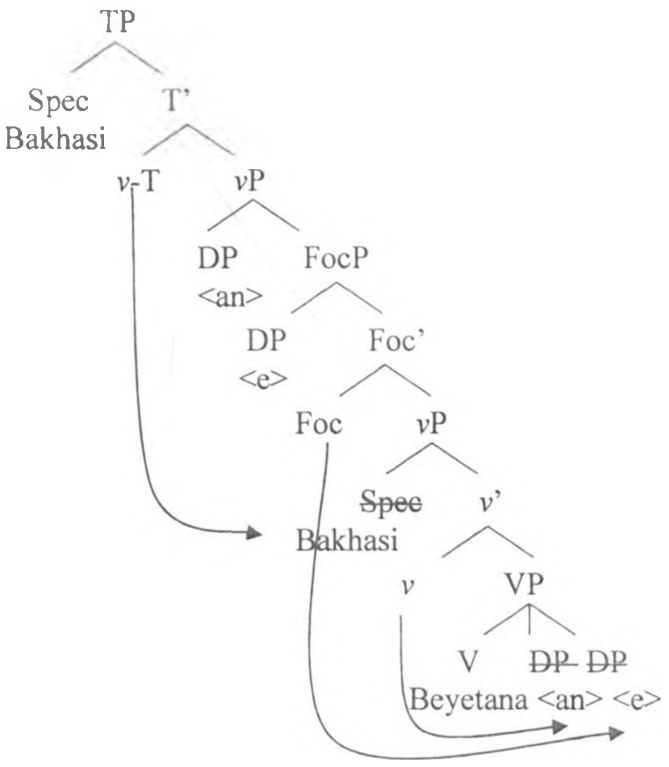
In order to address this question, we need to classify the respective affixes adequately. I argue that such a classification will rely on the generalization stated in (1) below:

**Generalization 16**

Whenever the RFM and RCM co-occur in the same clause the reciprocal reading supersedes the reflexive reading, and the RFM is demoted to adjunct position leaving the RCM as the canonical object argument.

As a consequence of this generalization, the representation of (1) proceeds along the lines of other transitive constructions, as shown in (3) below.

(3)




Notice that the VP internal argument is represented by the RCM, that moves to the outer *v*P for case checking. Its uF is valued by the c-commanding DP in [Spec TP] position. Notable is also the fact that the RFM is selected as an adjunct, whose position is intermediary between the internal and external arguments. The positional occurrence is however not accidental. Recall that one of the big language facts is that of relating structure to meaning. Adjuncts by their nature are used in language to provide information that has a semantic effect on a syntactic representation. Consequently, the position of the adjunct- between the RCM and the subject- is suitable as it relates, by way of emphasis, the reciprocal reading to the participants involved. This explains why this element is analyzed as a focus construction whose head probes for a matching goal that subsequently moves to [spec Foc] position once checking has taken place.

### 5.1.2 RFM/RCM Cooccurrence with Applicatives/ Causatives

As attested in most studies on Bantu verbal morphology, (see Mchombo 2004, Mutonyi 2000) Bantu provides means for adding extra arguments to the argument structure of a verb. These include the use of applicatives and causatives. Of importance to this study is the fact that both the RFM and RCM can occur as arguments in a clausal structure containing either an applicative or a causative. This possibility presents us with a new challenge: that of explaining the representation of the double argument constructions. Before this challenge is addressed, I briefly review the applicative and causative constructions in LuBukusu, and how they license RFM/ RCM anaphoric elements in argument positions.

The applicative in LuBukusu is marked by the suffixation of *–il-* or *–el-* with the consequence that a new argument is introduced into the syntactic construction. Consider the following examples:

**(4a)** Omwaana a-kha-teekh-a kamapwondi
CL1child SM-Tns-cook-fv CL6potato
*The child is cooking potatoes*

**(4b)** Omwaana a-kha-teekh-el-a Wekesa kamapwondi
CL1child SM-Tns-cook-Appl-fv Wekesa CL6potato
*The child is cooking potatoes for Wekesa*

Notice that (4a) is a simple transitive verb marked with a direct object *kamapwondi*. However with the addition of the applicative marker *–el-* in (4b), an extra argument (*Wekesa*) is added. I will use the term 'applied argument' to refer to the arguments introduced by applicatives. The applied argument is semantically related to the action of the verb by means of the applicative element. This relation is captured by the following

generalization:

**Generalization 17**

If an applicative *x* introduces an applied argument *y* then the applicative interpretation is such that verb-*x* *y*.

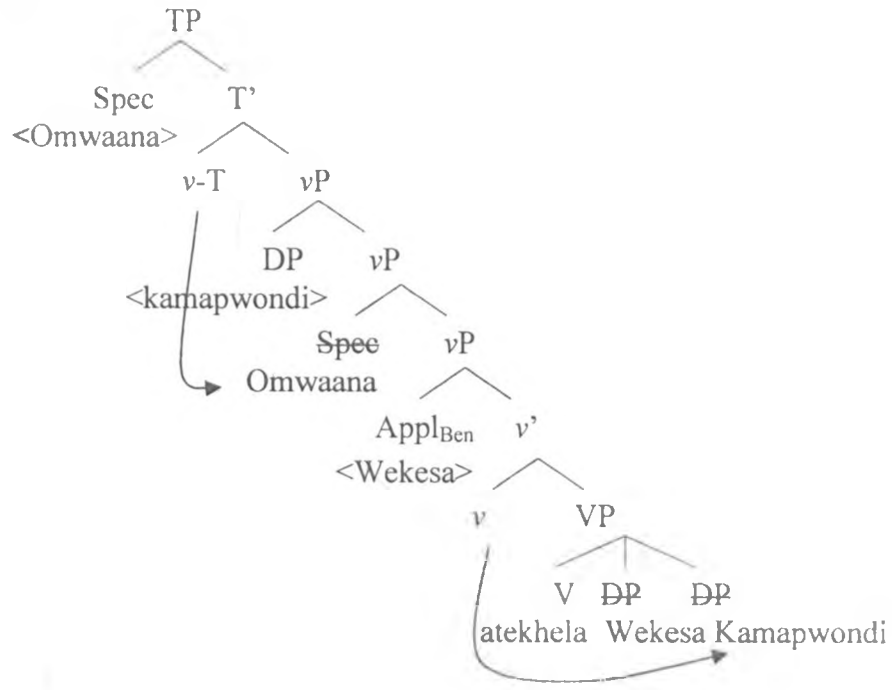
The first implication of this rule is that the applied argument in LuBukusu has a more distant relation to the direct object than to the verb. Pylkkanen (2002) refers to such applicatives as 'high' as opposed to 'low' applicatives instantiated in languages such as English double object constructions. Compare (4b) above with (6) below:

**(6)** Peter baked John a cake

Notice that the applied argument *John*, is directly related to the direct object *cake* in the sense that as a result of baking, *John* is in the possession of *a cake*. It cannot possibly mean that the cake was baked on behalf of John. On the contrary, both readings are available for the LuBukusu example. This kind of distinction is quite important for our purposes because it is ideally based on the relative position of the applied argument in relation to the verb in terms of the available semantic interpretation. The idea is simple: High applicatives appear next to the external argument because the applicative in this case relates the verb to the subject to express the 'on behalf of' notion. On the other hand, low applicatives occur next to the internal argument because of the possession interpretation between the object and applied argument. Note that the interpretation of such a hierarchy rests on the fact that external arguments are higher than internal arguments. This form of structural representation has direct consequences on the way MP represents applicative structures. I suggest a representation that is not quite different from what we have seen so far. Consequently, example (4b) will have the structure in (7)

below:

(7)




The derivation of this structure proceeds through the conventional computation process where the *v* head of the functional categories acts as a probe motivated by uninterpretable nominal features. It searches down the structure for a matching goal. Two DPs are active by virtue of having uninterpretable inherent case features. They consequently move to the [spec vP] positions. The applied argument moves below the external argument to be next to the verb in order to relate the action of the verb to the subject i.e. to show that potatoes were cooked by the children on behalf of Wekesa. On the other hand, the direct object moves to the outer vP, its normal position in transitive verbs.. Further notice that the applicative head is benefactive. This specification recognizes the fact that in most languages, the applicative has varied interpretations such as instrumental, locative and circumstantial. I suggest that the choice of the type of applicative used is a language

specific parameter.

With this background in mind, we can now examine RCM/ RFM co-occurrence in applicative constructions. Consider the examples below:

**(8a)** Babaana ba-e-siim-il-an-a
CL2child SM-RFM-like-Appl-RCM-fv
*The children love themselves for each other*

**(8b)** Babaana ba-siim-an-il-a kamapwondi
CL2child SM-like-RCM-Appl-fv CL6potato
*The children love each other with potatoes*

**(8c)** Babaana ba-e-siim-an-il-a
CL2child SM-like-RCM-Appl-fv
*The children love each other for themselves*

In (8a), the RFM is the direct object while the RCM is the benefactive applied argument. Notice the order of the applicative and the RCM which only allows a reading where the reciprocal argument is the applied object. In (8b) the RCM comes before the applicative and is therefore the direct object with an independent applied argument *kamapwondi*. The same thing applies to (8c) where the RFM is the applied argument. This sort of data leads naturally leads to the following generalization:

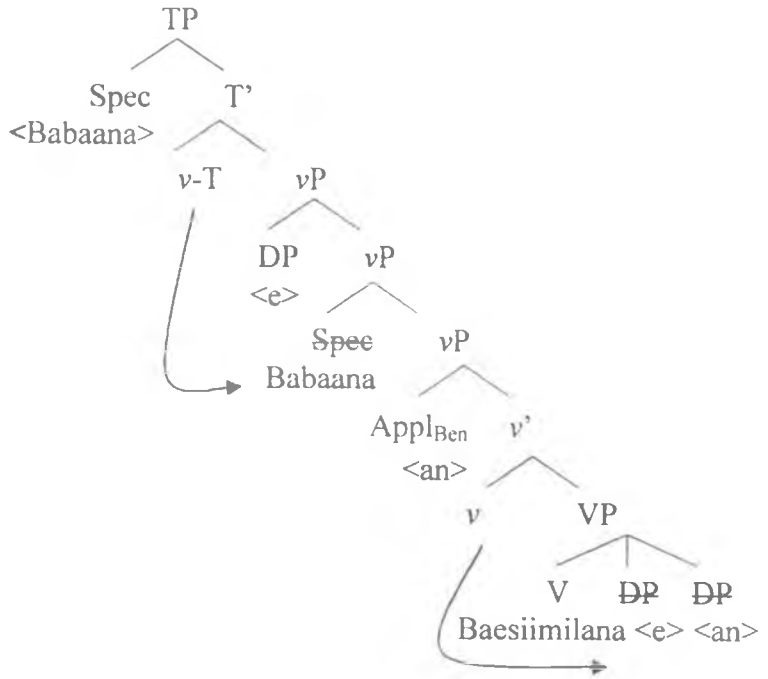
**Genaralization 18**

In applicative constructions with RCM, the RCM/applicative order identifies the status of the arguments as direct or applied object. If the applicative precedes the RCM, the RCM will always be the applied argument. If the order is reversed, the RCM is the direct object.

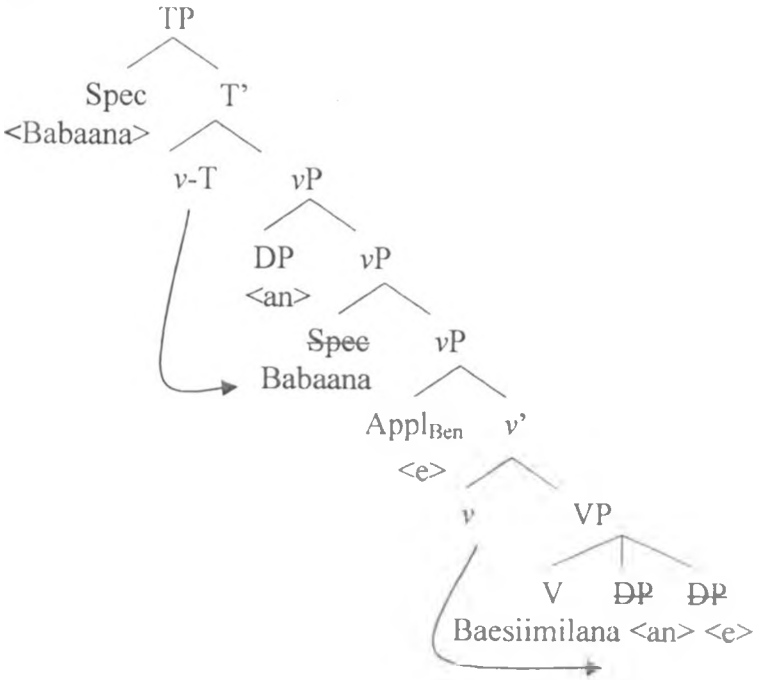
The implication of this generalization on the status of the RFM, when it co-occurs with the RCM is such that it will be a direct object if the applicative immediately follows the verb as in (8a), or an applied argument if the RCM immediately follows the verb as in (8c). These differences can be captured in the contrasting representations for (8a) and

(8b) in (9a) and (9b) respectively.

**(9a)**




**(9b)**



In these structures, the *v* in each tree searches down the structure for a matching goal. Two DPs are active by virtue of having uninterpretable case features. They, consequently move to the [spec *v*P] positions. The applied argument (*an* in (9a) and *e* in (9b)) moves below the external argument for the semantic reasons explained above, while the direct object moves to the outer *v*P. This means that the computation of (8b) also proceeds in the same manner, with the RCM functioning as the direct object and the NP *kamapwondi* as the applied object. I extend similar arguments to the causative constructions.

To recap, causative verbs universally seem to involve a causer argument that is normally absent from the syntax of non-causative verbs. This implies that causative verbs are derived by the addition of a head adding a causer argument to the meaning of the verb. Such a head is traditionally referred to as the causative marker realized differently in various languages. In LuBukusu, the causative marker is realized by the forms *–isy-, -esy-, or -y-,* the selection of any of them in a syntactic context being determined by phonological factors affecting neighbouring sounds.[32]

The examples in (10) help to illustrate causativization in LuBukusu, and also form the basis upon which we examine the representation of the RFM/ RCM co-occurrence in causative constructions.

**(10a)** Babaana ba-a-e-siim-an-isy-a
CL2child SM-Tns-RFM-like-RCM-Caus-fv
*The children made themselves like each other*

**(10b)** Babaana ba-a-e-siim-isy-an-a
CL2child SM-Tns-RFM-like-Caus-RCM-fv
*The children made each other like themselves*

[32] (See Mutonyi (2002) for a detailed discussion of the distribution of causative markers in LuBukusu).

In these structures, the *v* in each tree searches down the structure for a matching goal. Two DPs are active by virtue of having uninterpretable case features. They, consequently move to the [spec *v*P] positions. The applied argument (*an* in (9a) and *e* in (9b)) moves below the external argument for the semantic reasons explained above, while the direct object moves to the outer *v*P. This means that the computation of (8b) also proceeds in the same manner, with the RCM functioning as the direct object and the NP *kamapwondi* as the applied object. I extend similar arguments to the causative constructions.

To recap, causative verbs universally seem to involve a causer argument that is normally absent from the syntax of non-causative verbs. This implies that causative verbs are derived by the addition of a head adding a causer argument to the meaning of the verb. Such a head is traditionally referred to as the causative marker realized differently in various languages. In LuBukusu, the causative marker is realized by the forms *–isy-, -esy-, or -y-,* the selection of any of them in a syntactic context being determined by phonological factors affecting neighbouring sounds.[32]

The examples in (10) help to illustrate causativization in LuBukusu, and also form the basis upon which we examine the representation of the RFM/ RCM co-occurrence in causative constructions.

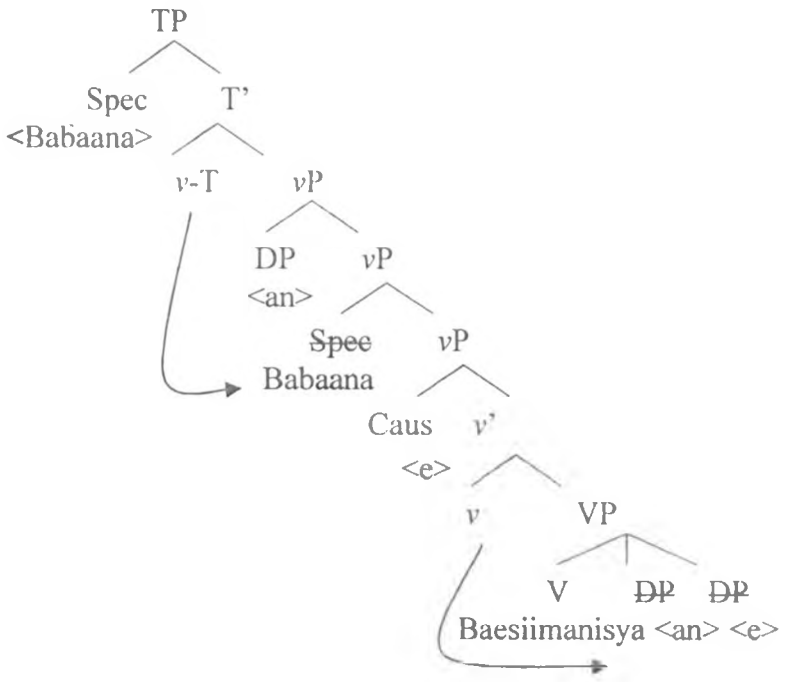
**(10a)** Babaana ba-a-e-siim-an-isy-a
CL2child SM-Tns-RFM-like-RCM-Caus-fv
*The children made themselves like each other*

**(10b)** Babaana ba-a-e-siim-isy-an-a
CL2child SM-Tns-RFM-like-Caus-RCM-fv
*The children made each other like themselves*

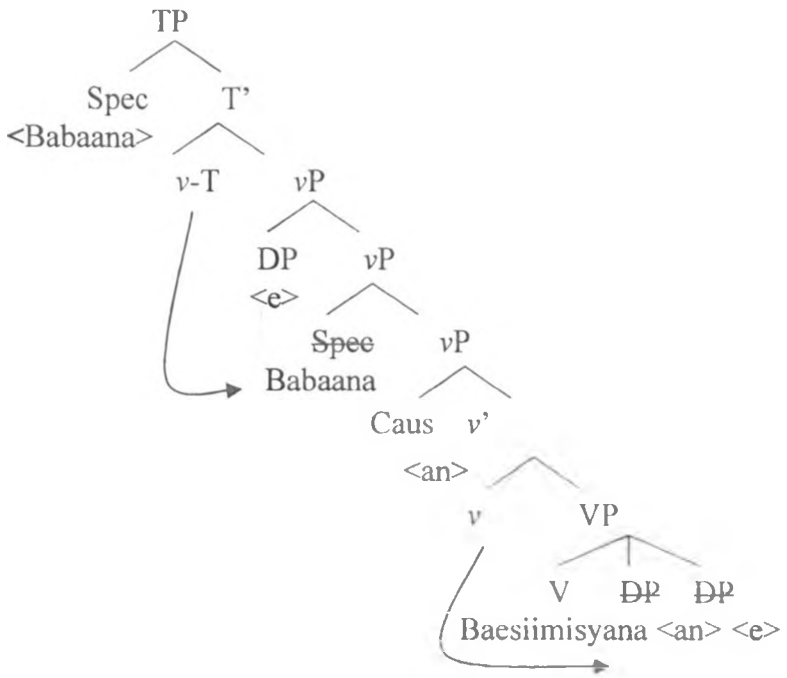
[32] (See Mutonyi (2002) for a detailed discussion of the distribution of causative markers in LuBukusu).

It goes without say that the RCM is the direct object in (10a) leaving the RFM as the cause argument. The reverse is the case in (10b) where the RFM is the direct object and the RCM, the cause argument. Again, the order of the causative vis-à-vis the RCM plays an important role in determining the argument classification of each of the anaphoric elements. When the RCM precedes the causative, the reciprocal argument is the direct object. Conversely, if the RCM follows the causative, the reciprocal is the cause argument leaving the reflexive to be the direct object. These facts can easily be captured in contrasting representation for (10a) and (10b) in (11a) and (11b) respectively.

**(11a)**



**(11b)**




The motivation for such structures has already been discussed. Perhaps what needs justification is the position of the causative argument. Notice that it occurs between the external argument and the *v*. The reason for this arises from its semantic role of relating the syntactic subject to the verb. The force of causation is derived from the external argument, and is subsequently linked to the verb by means of the causative. Again the implication for this reasoning is that, like the applicative, the cause argument also has a relatively distant relation to the direct object. In addition, the representation also captures an important fact about the cause argument: It is a kind of subject since it is infact the agent of the action of the verb, but in this case it is not the structural subject, but a thematic subject. Naturally, its position should be lower than that of the external argument.

## 5.2 Representation of Focus anaphoric elements

Both the affixal anaphoric elements and the free forms are realized in focus positions. In section 5.1.1 above, I examined how the RFM can occur as a focus element in conjunction with a RCM in contexts where there is no causative/ applicative. This focus element will therefore be excluded from this section. In chapter 3, I discussed productive cases of multiple occurrences of RCM and RFM motivated by causative and applicative markers, but in which at least one of the reduplicated affix is a focus element in the RCM construction, but both are arguments in RFM construction. The RCM case is discussed in this section. In addition, I also look at cases where Agr-eene and the phrasal reciprocal can add focus when combined with structures having RFM+ Agr-eene or RCM+ phrasal reciprocal forming quite complex patterns of focus.

### 5.2.1 Representation of RFM/RCM Reduplication

As a starting point, consider the following sentences:

**(12a)** Ba-baana ba-a-p-an-il-an-a kumukaati
CL2-child SM-Tns-fight-RCM-Appl-RCM-fv CL3-bread
*Children fought each other/other people for bread for each other*

**(12b)** Emuuna ya-a- siim-an-isy-an-a chisaang'i
CL9Squirrel SM-Tns-fight-RCM-Caus-RCM-fv CL10animals
*The squirrel caused the animals to like each other on each others effort*

Strictly speaking, the argument structure of these sentences is still similar to that already discussed for applicative and causative structures. In (12a), the RCM is the direct object while *kumukaati* is the applied argument. In (12b), the RCM is still the direct object with *chisaang'i* as the cause argument. The most logical question would then be concerned with the status of the extra RCM. It seems that extra information is added. In (12a), this

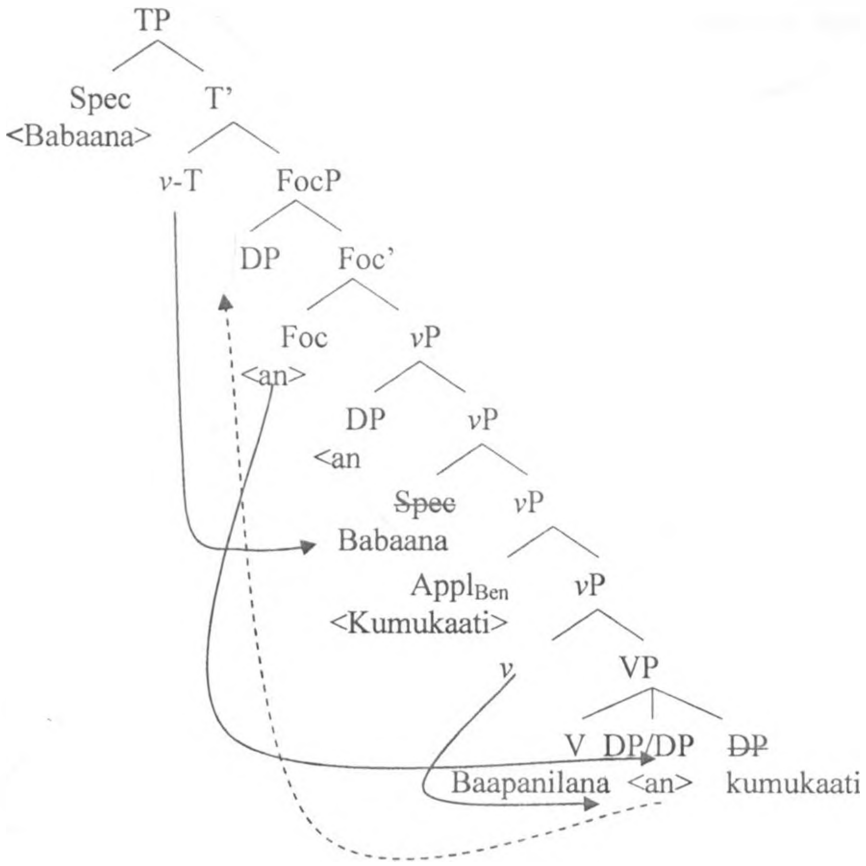
information reiterates the beneficiary of the applied argument. It is therefore a focus element. On the other hand, the cause argument in (12b) is focused as the force behind the action of the verb. It is also emphatic. This conclusion means that the extra RCM is an adjunct and must be analyzed as such. However, before this is done, I exemplify RFM reduplication in (13) below:

**(13a)** Babaana ba-a-e-i-khup-isy-a (khwisisi)
CL2child SM-Tns-RFM-RFM-hit-Caus-fv CL wall
*The children caused themselves to hit themselves against the wall*

**(13b)** Bakhasi ba-a-e-i-ng'on-el-a
CL2child SM-Tns-RFM-RFM-groom-Appl-fv
*The women groomed themselves for themselves*

Unlike the RCM, the two RFMs represent two arguments where the first is the direct object and the other is the cause/ applied argument. The representation should therefore be similar to that of double object constructions exemplified in 5.1 above. Going back to the double RCM, a model representation for such constructions is shown in the analysis of (12a) in (14) below:

**(14)**




This structure correctly predicts that *kumukaati* is the applied argument licensed by applicative feature checking via *v*. The double RCMs form the direct object and adjunct whose features are checked on their respective heads. The direct object moves to the outer *v*P for case checking. The adjunct provides additional information which specifies the beneficiary of 'bread'. The extra information is adjoined to a FocP after moving from the VP internal position motivated by the need to check focus.

It is predictable that a similar representation will still apply to the reduplicated RCM in causative constructions. The only difference will be on the labeling of causative instead of applicative.

### 5. 2.2 Agr-eene and Phrasal Reciprocal in Focus Positions

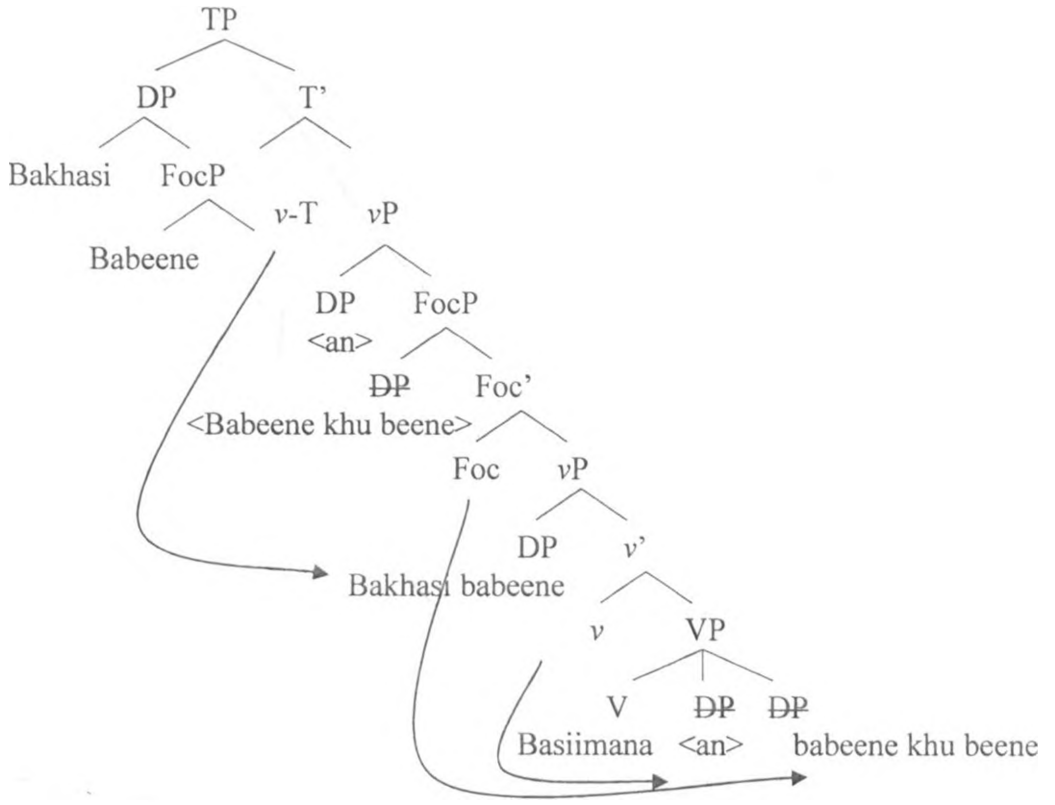
The role of Agr-eene and the phrasal reciprocal as focus elements has already been highlighted especially in contexts where they co-occur with the RFM/RCM respectively. In chapter 4, they are both analyzed as focus elements enhancing either the reflexive or reciprocal readings. In this section, I examine the representation of the two elements when added to structures already containing RFM + Agr-eene or RCM + phrasal reciprocal, forming quite complex patterns of focus whose details are discussed in section 3.3. One thing however emerges: whatever the pattern, Agr-eene and the phrasal reciprocal add an extra focus position which is contrastive i.e. it helps to specify the antecedent in contrast to other elements that might have occupied the focus position. Consider the following sentences:

**(15a)** Bakhasi babeene ba-a-yeet-an-a babeene khu beene
Cl2 women Agr-own SM-Tns-help-RCM-fv Agr-own on own
*Women themselves helped each other*

**(15b)** Bakhasi ba-e-yeet-a babeene, babeene
Cl2 women SM-RFM-help-fv Agr-own Agr-own
*Women themselves helped themselves*

Recall that the phrasal reciprocal in (15a) enhances the reciprocal marking whereas the first Agr-eene in (15b) enhances reflexive reading. The extra Agr-eene in both sentences provides an additional focus position that is assentially contrastive. How then do we represent this form in MP. The starting point is to determine its status within the syntactic construction. In (15a) the extra Agr-eene is part of the subject DP, and must therefore be represented in this position. The following structure is then derived.

(16)




Notice that the computation of this structure proceeds in the same way as all the others except that Agr-eene is in a focus position within the subject DP, and that its interpretation is determined by the c-commanding head of the DP, fulfilling the binding relations. I extend the same argument to (15b), where the extra Agr-eene is analysed as a focus element that only appears in the clause final position but its focus feature is checked under the focus head embedded in the subject DP position.

**5.3 Conclusion**

In this chapter, I have discussed various co-occurrence relations between the affixal and non- affixal markers while noting that the ability to co-occur is licensed by different reasons: a) the need for emphasis where the affixal and non –affixal markers pattern

differently to enhance the reciprocal or reflexive reading or create contrastive focus; b) Verbal extensions such as causative and applicative markers which add extra arguments to the argument structure of a verb where some of the added arguments could be the RFM, RCM, Agr-eene or the phrasal reciprocal. I have given evidence to show that the competition for the same argument position is only apparent, since one of the affixes only helps to reiterate the other in an affix referential hierarchy, and must therefore be analyzed as an adjunct. Further, I have exemplified the fact that RFM/ RCM co-occurrence is also licensed by argument increasing strategies, in which case there is no argument slot competition, and the two represent different arguments, with independent representation in MP. I extended the same argument to cases of multiple RCM or RFM co-occurrences where it is noted that for the RCM one of them is always an adjunct while the RFMs are arguments. I also examined the representation Agr-eene in extra focus position within the subject DP and concluded that the whole phrase containing the subject moves conventionally to spec TP position, but splits to create a focus head to check the focus features associated with Agr-eene. This is still the case even when Agr-eene is in a clause final position.

# Chapter Six

# Conclusions and Recommendations

## 6.0 Introduction

This study set out to investigate the question of how NPs in a syntactic construction are interpreted as being co-referential with other NPs in the same construction or in the general discourse structure. The main language of investigation was LuBukusu. Key puzzles that constituted the problem of the study were identified. For example, it was noted that the RFM optionally occurs with an *Agr-eene* form which raises interesting questions about how such cooccurrences should be syntactically interpreted and theoretically analyzed, and whether morphology plays any general role in the mapping of syntactic processes. In order to do this effectively, hypotheses were formulated. These are repeated below:

LuBukusu anaphoric relations motivate specific structural and discoursal patterning that can be accounted for by the available models of syntactic analysis like Minimalism. More specifically:

i) All the anaphoric elements are realized in a structurally similar way.
ii) The anaphoric elements are construed within the minimum clause in which they occur.
iii) Anaphoric relations are constrained by both syntactic and discoursal factors.
iv) The arguments of the Minimalist Program can capture all the anaphora facts.

To test these hypotheses, first there was need to identify the types of NPs that made up the class of anaphoric elements while describing their form and occurrence context. In addition, the minimalist program provided the framework for formulating ways of representing the resultant anaphoric relations. More specifically the study focused on Derivation by Phase model of MP to provide explanations and constraints for varied patterns of anaphora resulting from either the anaphoric NPs occurring independently or in conjunction with other forms. Data were primarily provided by LuBukusu, a Bantu language spoken in western parts of Kenya. I give the conclusions in 6.1 and the recommendations in 6.2.

## 6.1 Conclusions

On the question of the types of anaphoric elements, four types were identified, and were grouped in two broad categories which include the affixal markers (RFM and RCM), and the non-affixal markers (Agr-eene and phrasal reciprocal). These were analyzed in terms of morphological composition and syntactic context. For example, it was noted that the RFM is realized as *–i-,-e-* or *ye* in specified environments represented by the following simple rules:

a)...[*a*]/*RFM* → [ *e*]

b) .../*RFM* → *i*/ *G*/*e*

These specify the context as being made up of a morpheme immediately preceding the RFM and whose final vowel is *–a-*, consequently leading to the realization of the RFM as *–e-*. In (b), the position preceding the RFM has an unspecified value i.e. it can be any of the other four vowels: *e, i, o,* or *u* in which case, *-i-* is selected or *–e-* so long as it is

preceded by a glide or a consonant. On the other hand, the RCM is realized generally as –*an*, but which can also be –*chan*, a phonologically conditioned allomorph realized in combination with monosyllabic verb stems.

Further, the agreement portion of Agr-*eene* matches the noun class and person of its antecedent and -*eene* means 'own' or 'owner'. The phrasal reciprocal can be represented as *'x-eene khu/ne x-eene'* where *x* is an agreeing subject marker while *khu* is the preposition *on* and *ne* literally meaning *with.*

On the question of syntactic context, the study noted that apart from the general clausal context, both the RFM and RCM can occur within a deverbalized noun. The RFM is, for example, possible in all deverbal nouns ending in a patient nominalizer [e], partially possible in those ending in [o], and totally excluded in those ending in agentive [i], and [fu]. The RCM is quite unproductive in deverbal nouns as it occurs only with noun classes 5, 14 and 15. In addition, the RFM and RCM are licensed by verb transitivity. The most common idea being that if a verb can take an overt object or an OM, then that verb will also host the RFM or the RCM, in which case there cannot also be an OM on the verb in question unless it is licensed by one of the valence increasing devices such as applicative or causative. By extension, the two affixes occur in a strictly local environment with c-commanding antecedent in the minimal clause or NP.

On the syntactic constraints, the grammaticalization of the RFM and the RCM appearing in the position of an incorporated pronominal argument, supports their treatment as a pronominal arguments whose construal is determined by principles of syntactic binding

According to the present analysis, incorporated affixes interpreted as incorporated urguments such as the RFM and RCM enter into the derivation with feature specifications, normally marking agreement and they are these features that license their representation in a syntactic construction. This kind of treatment arises from the fact that both the RFM and RCM are parallel to the OM in terms of shared characteristics (identity in occurrence position (for the RFM) and conditioned complementarity with free NPs). This kind of treatment is also because of the separation of syntactic and semantic valence, where contra many Bantuists, the RFM and RCM are seen as increasing valence semantically, but syntactically merge both the agent and patient roles of the verb. The two therefore assume the full status of lexical heads heading phrasal categories participating in the syntactic derivation.

On the non-affixal markers, Agr-eene distinguishes three main roles: a) enhancing the reflexive reading; b) focus element; c) Independent pronoun. These roles determine, to a great extend, the context of occurrence. In order to enhance a reflexive reading, Agr-eene must occur with RFM. In other cases it requires an antecedent within the same sentence or the extended discourse context, in which case it adds new information in terms of focus. When it occurs independently as an argument, more often it marks possession. The phrasal reciprocal realizes two forms, characterized as R1 and R2 in this study. R1 is only licensed in conjuction with RCM to enhance reciprocal reading while R2 occurs in focus positions, and occasionally as a pronoun especially in oblique object positions.

On the question of how the anaphoric elements are able to achieve binding relations, it was noted that the relations are either local or non-local. The RFM and the RCM, for

example are exclusively local i.e. they are bound by a c-commanding clause mate subject. In addition, they trigger both the sloppy and strict reading under VP ellipsis. On the other hand, the binding relations involving Agr-eene and the phrasal reciprocal are quite flexible with the former having greater flexibility. For example it can refer to an exclusively local antecedent or an extra-sentential one. The former is, for example, possible when it co-occurrs with the RFM, or is in a PP complement position. In other contexts, its reference is long distant. The phrasal reciprocal, supports a reciprocal reading in conjunction with the RCM. In other environments, it is emphatic. There are however exceptions to the subject orientation hypothesis for local binding specifically in ECM, causative and commitative constructions which prompt a redefinition of the concept of 'subject'. Focus constructions are also discussed. The basic idea is that co-occurrence of anaphoric elements creates emphatic or contrastive focus. For example RCM and RFM co-occurrence leads to a reciprocal reading with a reflexive emphasis. Lastly, inherently reflexive and reciprocal verbs are discussed.

With the structural facts in place, MP comes in to account for such facts. The basic model used is DbP. For all the anaphoric elements, it is assumed that their interpretation takes place in the semantic component (LF). The whole computation is induced by phases and regulated by the twin operations of merge and move. The structure presupposed is the *v*P hypothesis which is headed by a *v* functional category that implements the move operation licensed by the presence of uninterpretable features on both the probe and the goal. The RFM and RCM are licensed to move to spec *v*P position because being active goals to a searching *v* probe. They both have uninterpretable case features. In addition,

they value their uninterpretable R feature on their antecedents satisfying binding relation. As a result, BT principles are reformulated along these lines as follows:

*Principle A:* If α is an anaphor it must have a valued F at the end of the computation.
*Principle B:* If α is a pronoun it must have uF at the end of the computation.
*Principle C:* If α is an R- expression it has no F.

A similar analysis is extended to Agr-eene and the phrasal reciprocal. When they are arguments, case determines movement, but when they are focus elements, they move to check the focus feature.

Finally, various co-occurrence relations between the affixal and non- affixal markers are discussed. The conclusion is that the ability to co-occur is licensed by different reasons: a) the need for focus where the affixal and non –affixal markers pattern differently to enhance the reciprocal or reflexive reading or create contrastive focus; b) Verbal extensions such as causative and applicative markers which add extra arguments to the argument structure of a verb where some of the added arguments could be the RFM, RCM, Agr-eene or the phrasal reciprocal. Further, the study exemplified the fact that RFM/ RCM co-occurrence is also licensed by argument increasing strategies, in which case there is no argument slot competition, and the two represent different arguments, with independent representation in MP. The same arguments are extended to cases of multiple RCM or RFM co-occurrences where it is noted that for the RCM one of them is always an adjunct while the RFMs are arguments. The thesis also examined the representation of Agr-eene in extra focus position within the subject DP and concluded

that the whole phrase containing the subject moves conventionally to spec TP position, but splits to create a focus head to check the focus features associated with Agr-eene. This is still the case even when Agr-eene is in a clause final position.

## 6.2 Recommendations

The study has noted that the whole phenomenon of anaphora covers both syntactic and discourse oriented anaphora. The focus of this study was on the former. I recommend a more inclusive analysis to capture even more intricate properties of anaphora in discourse.

The focus was mainly on the analysis of data from LuBukusu. I recommend a more comparative approach with other languages either within Bantu or from other language families.

The theoretical framework employed is mainly based on the generative framework. An analysis based on other theories may provide details of anaphora never discussed before.

# BIBLIOGRAPHY


Abdulaziz, M. (1996) *Transitivity in Swahili.* Koln: Koppe.

Abney, S. (1987) The Noun Phrase in its Sentential Aspect. M.I.T. Ph.D Dissertation.

Adger, D. (2008) A Minimalist Theory of Feature Structure. Seminar Paper, Queen Mary University of London. May 2008.

Akinbiyi, M.A and Adesola, O. eds. (2003) *Proceedings of the 4th World Congress of African Linguistics New Brunswick.* Koln: Rudiger Koppe Verlag.

Alexiadou, A. (2001) *Functional Structure in Nominals: Nominalization and Egartivity.* Amsterdam: John Benjamins.

Anderson, S.R. (2004) Some Points of Agreement. JHU IGERT Workshop. January 2004.

Austen, C. (1974) Aspects of Bukusu Syntax and Phonology. Ph.D dissertation, Indiana University.

Baker, M. (1985) The Mirror Principle and Morphosyntactic Explanation. *Linguistic-Inquiry* 16.3: 373-415.

Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University of Chicago Press.

Baker, M. (2003). Agreement, dislocation, and partial configurationality. In *Formal approaches to function in grammar*: In honor of Eloise Jelinek, eds. Andrew Carnie, Heidi Harley and MaryAnn Willie, 107-132. Amsterdam: John Benjamins.Chicago: University of Chicago Press.

Baybee, J.L. (1985) *Morphology: A study of the Relation between Meaning and Form.* Amsterdam: John Benjamins Publishing Company.

Black, C.A. (1998) A Step-by-Step Introduction to the Government and Binding Theory of Syntax. Botes on Linguistics 73.

Bokamba, E.G. (1981) Aspects of Bantu Syntax. Preliminary edition. Department of Linguistics, University of Illinois, Urbana-Champaign, Illinois

Bresnan, J. (1997) Mixed Categories as head sharing constructions. In: M. Butt and T.H King eds. *Proceedings of the LFG97 Conference.* Stanford on line CSLI Publ

Bresnan, J. (2001). *Lexical Functional Syntax*. Oxford: Blackwell Publishers.

Bresnan, J. and Mchombo, S.A. (1987). Topic, Pronoun and Agreement in Chichewa. *Language* 63(4): 741-782.

Bresnan, J. and Mchombo, S.A. (1995). The Lexical Integrity Principle: Evidence from Bantu. *Natural Language & Linguistic Theory* 13: 181-254..

Bresnan, J. and Moshi, L. (1990) Object Asymmetries in Comparative Bantu Syntax. *Linguistic Inquiry* 21:147-185.

Burzio, L. (1996) The Role of the antecedent in anaphoric relations. In *Current issues in comparative Grammar*. ed. Robert Freidin, 1-45. Dordrecht:Kluwer

Carstens, V. (1991) The Morphology and Syntax of Determiner Phrases in Kiswahili, University of California at Los Angeles.

Carter, H. and Makoondekwa, J. (1987). *Kongo Language Course: Maloongi Makikoongo. A Course in the Dialect of Zoombo, Northern Angola*. Madison: African Studies Program, University of Wisconsin.

Chimbutane, F. (2002). Grammatical Functions in Changana: Types, Properties and Function Alternations. M.Phil dissertation, Australian National University.

Chomsky, N. (1973) 'Conditions on Transformations', in S.R. Anderson and P. Kiparsky (eds) *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht, Foris Publications.

Chomsky (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Linguistic Inquiry Monograph six. Cambridge: The MIT Press.

Chomsky (1986). *Barriers.* Linguistic Inquiry Monograph Thirteen. Cambridge, MA and London: The MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N. (2000). Minimalist Inquiries: The Framework. *Step by Step: Essays on Minimalist Syntax in honor of Howard Lasnik.* Martin, D. Michaels and J. Uriagereka (eds). Cambridge, Mass: MIT Press.

Chomsky, N. (2001a). *On Nature and Language*. Adriana Belletti and Luigi Rizzi, eds. Cambridge: Cambridge University Press.

Chomsky, N (2001b). Derivation by Phase. In: Kenstovicz, M. ed. *Ken Hale: A Life in Language*. Cambridge, Mass: MIT Press. 1-54.

Chomsky, N. and Lasnik, H. (1993) Principles and Parameters Theory. In Jacobs, J, A von Stechow, W Sternefeld, and T.Venneman (1993). *Syntax. An International Handbook of Contemporary Research.* Berlin: Walter de Gruyter.

Cole,P., Hermon, G., Huang, C-T.J. (2001) Long distance Reflexives: The State of the Art. In: *Long distance Reflexives. Syntax and Semantics* Vol 33. New york: Academic Press.

Cole, P. and Hermon, G. (1998) Long distance Reflexives in Singapore Malay: an apparent typological anomally. *Linguistic Typology* 2(1).

Corbett, G. and Al Mtenje (1987) Gender agreement in Chichewa. *Studies in African Linguistics, 18. 1-38*

Creissels, D. and Godard, D. (2005) Tswana Infinitives as a Mixed Category. *Proceedings of the HPSG05 Conference.* Stanford on line CSLI Publ.

Dalrymple, M, M Kanazawa, Yookyung K, S. Mchombo & Stanley, P. (1998). Reciprocal Expressions and the Concept of Reciprocity. *Linguistics & Philosophy 21*: 159-210

Dalrymple, M, S. Mchombo, and Stanley P. (1994). Semantic Similarities and Syntactic Contrasts between Chichewa and English Reciprocals. *Linguistic Inquiry* **25**(1): 145-163.

De Blois, K.F. (1975) *Bukusu Generative Phonology and Aspects of Bantu Structure.* Koninklyk Museum voor midden Afrika, Tervuren, Belgie.

Deen, K. (2004) Let's Build Some Agreement: The status of the Subject Prefix in Nairobi Swahili. Paper presented at the 35th Annual Conference of African Linguistics, Harvard University.

De Jong, J.J. (1996) The case of bound pronouns in peripheral Romance. Ph.D dissertation, Goningen University.

Demuth, K and M. Johnson (1989). Interaction Between Discourse Functions and Agreement in Setawana. *Journal of African Languages and Linguistics 11*: 21-35.

Demuth, K. (2000) Bantu noun class systems: Loanword and acquisition. Evidence for semantic productivity. In: Senft, G.ed. *Systems of Nominal Classification.* Cambridge:Cambridge University Press. 270-297.

Diercks, M. (2006) Object agreement in Swahili and the interpretation of objects. Ms Georgetown University.
Diercks, M (2010) Agreement with Subjects in LuBukusu. PhD Dissertation. Georgetown University.
Diercks, M and Sikuku, J (In Preparation) Object Marking in LuBukusu.
Dlayedwa, N C. (2002). Valency-Reducing processes in Xhosa. Doctoral dissertation, University of Essex, UK.
Du Plessis, J. A. and M. Visser (1992). *Xhosa Syntax*. Pretoria, Via Afrika.
Falk, Yehuda N. (2001). *Lexical Functional Grammar. An Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI Publications.
Epstein, S.D. and Seely, T.D. (2002) Rule application as cycles in a level free syntax. In: Epstein, S.D. and Seely, T.D. eds. *Derivation and Explanation in the Minimalist Program.* Oxford: Blackwell.
Fedders, A. and Salvadori, C. (1980) *People and Cultures of Kenya.* Nairobi: Tranafrica.
Featherston, S. (2007) Data in generative grammar: the stick and the carrot. *Theoretical Linguistics* 33.3: 269-318.
Fretheim, T and Hellan, L (1982) *Papers from the Sixth ScandnavianConference of Linguistics.* Trondheim: Tapir
Gelderen, E. (2000) *A History of English Reflexive Pronouns*. Amsterdam: John Bejamins.
Giorgi, A. (1984) Towards a theory of long distance anaphora: a GB approach. *Linguistic Review* 3:307-359.
Goldsmith, J and F. Sabimana (1985). The Kirundi Verb. Unpublished paper, The University of Chicago.
Gueron, J., Obenauer, H-G, and Pollock, J-Y (1985) *Grammatical Representation.* Dotrecht: Foris
Guthrie, M. (1962). The Status of Radical Extensions in Bantu Languages. *Journal of African Languages 1*: 202-220.
Guthrie, M. (1967-71). *Comparative Bantu*. Farnsborough: Gregg International Publishers.
Haegeman, L. (1992) *Theory and Description in Generative Grammar: A case Study in West Flemish.* Cambridge: Cambridge University Press.
Haegeman, L. (1994) *Introduction to Government and Binding Theory.* 2nd ed. Oxford: Blackwell.
Hagege, C. (1974) Les pronoms Logophoriques. In : *Bulletin de la Societe de Linguistique de.*
Hellan, L. and Christensen, K.K (1986) *Topics in Scandnavian Syntax.* Dotrecht: Reidel.
Hornstein, N., Nunes, J. and Grohmann, K. (2005) *Understanding Minimalism.* Cambridge: Cambridge University Press.
Horrocks, G. (1987) *Generative Grammar.* London : Longman.
Hyman, L. (1991). Conceptual Issues in the Comparative Study of the Bantu Verb Stem. In: Salikoko M. S. and L. Moshi. eds. *Topics in African Linguistics.* Amsterdam/Philadelphia: John Benjamins Publishing Co.: 3-34.
Hyman, L. and S. Mchombo (1992). Morphotactic Constraints in the Chichewa Verb Stem. Papers from the Parasession on the Place of Morphology in a Grammar,

Berkeley, CA, Berkeley Linguistic Society.
Kayne, R.S. (1994) (1994) *The Antisymmetry of Syntax.* Cambridge, Massachusetts: MIT Press.
Kayne, R.S. (2005) *Movement and Silence.* Oxford: Oxford University Press.
Keach, C. N. (1995). Subject and Object Markers as Agreement and Pronoun Incorporation in Swahili. In: A. Akinlabi. ed. *Theoretical Approaches to African Linguistics 1*. Trenton, NJ, Africa World Press, Inc. 109-116.
Keenan, E.L. and J. P Razafimamonjy (2001) Reciprocals in Malagasy. In: Harold Torrence ed. *UCLA Working Papers in African Linguistics* 1. UCLA Department of Linguistics Number 6. 42-89.
Kimenyi, A (1980) *A Relational Grammar of Kinyarwanda.* Berkeley: University of California Press.
Kinyalolo, K. (1993) The logophoric pronoun emi in Fon as an LF operator/ anaphor. *Proceedings of NELS,* Vol 23: 223-237.
Koster, J. (1985) Reflexives in Dutch. In: Gueron, Obenauer, and Pollock (1985) *Grammatical Representation.*Dotrecht: Foris. 141-168.
Koster, J. (1987) *Domains and Dynasties.* Dotrecht: Foris.
Koster, J. and Reuland, E. eds. (1991) *Long distance Anaphora.* Cambridge: Cambridge University Press.
Kraal, P.J. (2005) Agrammar of Makonde. Ph.D. dissertation, Leiden University.
Jelinek, E (1984). Empty Categories, Case, and Configurationality. *Natural Language and Linguistic Theory* 2:39-76.
Letsholo, R. 2002. Syntactic Domains in Ikalanga. Doctoral dissertation. University of Michigan.
Maling, J. (1982) Non clause bounded reflexives in Icelandic. In: Fretheim and Hellan (1982) *Papers from the Sixth ScandnavianConference of Linguistics.* Trondheim: Tapir. 90-106.
Maling, J. (1986) Clause bounded reflexives in modern Icelandic. In: Hellan and Christensen (1986) *Topics in Scandnavian Syntax.* Dotrecht: Reidel.
Malouf, (2000) *Computational Linguistics in the Netherlands.* Amsterdam:Rodolph
Marantz, A. (1993) Implications of asymmetries in double object constructions. In: Mchombo, S, ed. (1993) *Theoretical Aspects of Bantu Grammar*. Stanford, CA: CSLI Publications. 113-151.
Markman, V.G. (2009) On the parametric variation of case and agreement. *Natural Language Linguistic Theory 27*. 379-426.
Marlo, M.R. (2003) Prefixal reduplication in Lusaamia: evidence from morphology. In: Akinbiyi, M.A and Adesola, O. eds. (2003) *Proceedings of the 4th World Congress of African Linguistics New Brunswick.* Koln: Rudiger Koppe Verlag.
Masayo, I, Wechsler, S and Zec, D. eds. (1986) *Working Papers in Grammatical Theory and Discourse Structure: Interraction of Morphology, Syntax and Discourse.* Stanford: CSLI Publications.
Matsinhe, S (1994). The Status of Verbal Affixes in Bantu Languages with Special Reference to Tsonga: Problems and Possibilities. *South African Journal of African Languages 14. No. 4*: 163-176.
Mchombo, S, ed. (1993) *Theoretical Aspects of Bantu Grammar*. Stanford, CA: CSLI Publications

Mchombo, S. A. (1993). A Formal Analysis of the Stative Construction in Bantu. *Journal of African Languages and Linguistics 14*: 5-28.
Mchombo, S. A. (1993). On the Binding of the Reflexive and the Reciprocal in Chichewa. In Mchombo, S. ed. (1993) *Theoretical Aspects of Bantu Grammar*. Stanford, CA: CSLI Publications. 181-208.
Mchombo, S.A. (2002). Affixes, Clitics and Bantu morphosyntax. In A. Mengistu and P. Collins (eds.) *Language Universals and Variation.* Westport, CT: Greenwood Press. 182-213
Mchombo, Sam. (2004). *The syntax of Chichewa*. Cambridge, UK: Cambridge University Press.
Mchombo, Sam (2005). Nyanja. *Encyclopedia of Languages and Linguistics*. 2.Volume 8. Oxford, Elsevier. (ed.) Keith Brown pp.744
Mchombo, S. A. and R. M. Ngalande (1980). Reciprocal Verbs in Chichewa: A Case for Lexical Derivation. *Bulletin of the School of Oriental & African Studies, University of London 45*: 570-575.
Mchombo, S. A. and A. S. A. Ngunga (1994). The Syntax and Semantics of the Reciprocal Construction in Ciyao. *Linguistic Analysis 24*(1-2)
McNally, L. (1993). Comitative Coordination: A Case Study of Group Formation. *Natural Language & Linguistic Theory 11*: 347-379.
Meeussen, A.E. (1967) *Bantu Grammatical Reconsruction.* Annales du muse, Royale de l'Afrique centrale. Tervuren 8 :80-121.
Mugane, John. 1997. *A Paradigmatic Grammar of Gikuyu.* Stanford, CA: CSLI Publications.
Muller, G. (2005) Remarks on cyclic linearization and order preservation. *Theoretical Linguistics.*
Mutonyi, N. (1992) Cs and Vs or moras: The case of Bukusu prosodic structure. *OSU Working Papers in Liguistics* 41: 46-84.
Mutonyi, N. (1996) Tense, aspect and Bukusu verb tones. *OSU Working Papers in Linguistics* 48: 77- 110.
Mutonyi, N. (2000) Aspects of Bukusu morphology and phonology. Ph.D dissertation . Ohio State University.
Nissenbaum, J. (2000) Covert movement and parasitic gaps. *Proceedings of NELS* 30, 2000.
Nurse, D. and Philippson, G. (2003) *The Bantu Languages.* Routledge.
Payne, T.E. (1997) *Describing Morphosyntax.* Cambridge: Cambridge University Press.
Pollock, J-Y. (1989) Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20: 365-424.
Pullum, G. (1991) English nominal gerund phrases as noun phrases with verb phrase heads. *Linguistics* 29: 763-799.
Pylkkanen, L. (2002) Introducing arguments. Ph.D dissertation, MIT.
Radford, A. (1981) *Transformational Syntax:A students' guide to Chomsky's extended standard theory.* Cambridge: Cambridge University Press.
Reidel, K. (2009) *The Syntax of Object Marking in Sambaa: A Comparative Bantu Perspective.* Utrecht: LOT.
Reinhart T. (1997) Strategies of anaphora resolution. *OIS Working Papers.*
Reinhart, T. and Reuland, E. (1993) Reflexivity. *Linguistic Inquiry* 24: 657-720.



UNIVERSITY OF NAIROBI
LIBRARY

Reuland, E. (2001) Primitives of Binding. *Linguistic Inquiry* 32: 439-492.
Reuland, E. (2005) Binding Conditions: How are they derived? In Muller, S. ed. *The Proceedings of the 12th International Conference on HPSG*: 598-593. Stanford: CLSI.
Rugemalira, J (1993) Runyambo Verbal Extensions and Constraints on Predicate Structure. Doctoral dissertation, University of California, Berkeley.
Rugemalira, J (1993) Bantu multiple object constructions. *Linguistic Analysis* 23: 226-252.
Sells, P., Zaenen, A., and Zec, D. (1986) Reflexivization variation: Relation between syntax, semantics and lexical structure. In: Masayo, I, Wechsler, S and Zec, D. eds. (1986) *Working Papers in Grammatical Theory and Discourse Structure: Interraction of Morphology, Syntax and Discourse.* Stanford: CSLI Publications. 169-238.
Sibanda, G. (2004) Verbal phonology and morphology of Ndebele. Doctoral dissertation, University of California, Berkeley.
Sikuku, J. (1998) The morphosyntactic structure of LuBukusu anaphoric relations: a government and binding approach. Mphil. Thesis. Moi University.
Simango, Ron. 1999. Lexical and Syntactic Causatives in Bantu. *Linguistic Analysis* 19:69-86.
Safir, K. (1996c) Semantic atoms of anaphora. *Natural Language and Linguistic Theory* `14: 545-589.
Safir, K. (2004) *The Syntax of Anaphora.* Newyork: Oxford University Press.
Schroeder, H. (2002) Word order in Toposa. An aspect of multiple feature checking. Ph.D dissertation. University of Nairobi.
Tang, S-W. (1998) Parametrization of features in syntax. Doctoral disseratation, University of California, Irvine.
Thrainson, H. (1991) Long distance reflexives and the typology of NPs. In: Koster, J. and Reuland,E. eds. *Long Distance Anaphora.* Cambridge: Cambridge University Press. 49-76.
Wasike, A. (1992) The simple sentence in Lubukusu: A GB approach. MA Thesis, University of Nairobi.
Wasike, A. (2007) The left periphery, wh-in-situ and A-bar movement in Lubukusu and other Bantu languages. Ph.D dissertation, Crnell University.
Welmers, W.E. (1973) *African Language Structures.* Berkeley: University of California Press.
Yang, D-W (1984) The extended binding theory of anaphors. *Language Research* 19: 169-192.
Zwart, J-W. (2001) Syntactic and phonological verb movement. *Syntax 4.* 34-62.