TWO TYPES OF PHRASAL COMPOUNDS IN CINYANJA

By

NAOMI NJOBVU

Student No: 215081676

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Linguistic Science
School of Arts, University of KwaZulu-Natal
Howard College Campus,
Durban.

Supervisor: Professor Jochen Zeller

December, 2020

Declaration

I, Naomi Njobvu, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a. Their words have been re-written but the general information attributed to them has been referenced.

b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed

.....

Dedication

To my late parents Mr Cornelio David Njobvu and Mrs Henriette Jellifa Njobvu.

Acknowledgements

I attest to the fact that my PhD journey was tough, rough, and not without tears. I, therefore, owe every gratitude to God for giving me the grace to complete my thesis, and to face life's challenges with love and patience for with Him nothing is impossible.

I would like to express my sincere gratitude to my supervisor Professor Jochen Zeller for his academic guidance, constructive criticism and sound advice from the conceptualisation of my topic to the final report. I appreciate his in-depth knowledge of linguistics immensely. The questions and the discussions we had, helped me to think and rethink critically and analytically. I also thank him for his concerns about my welfare.

To the University of KwaZulu-Natal, I am grateful for the financial support towards my school fees, for without the fee remission I was accorded, I would not have managed to finish my academic work.

I am grateful too to the University of Zambia for the financial support that enabled me to carry out my research and maintain my upkeep. In a special way, I thank Professor Felix Masiye, the Dean of the School of Humanities and Social Sciences, Dr Kelvin Mambwe and Dr Hambaba Jimaima who were heads of the Literature and Languages Department at the University of Zambia for the administrative facilitation extended to me during my studies at the University of KwaZulu-Natal.

I feel indebted to many people who contributed in different ways during my PhD studies. Special thanks to Dr John Simwinga, for encouraging me to embark on my studies. I appreciate the brilliant academic discussions and the great ideas he shared with me. Dr Sande Ngalande and Mr Peter M. Sakala, I appreciate most sincerely the academic discussions we had they helped to shape my work. Mr Rene M. Mwansa, Mrs Eunice Mukonde Mulenga, Mrs Bertha Mwale Moyo and Mrs Bessy Mwamba, thanks for always being there whenever I needed help. I also thank my friends at the University of KwaZulu-Natal, Dr Lydia Hangulu Mudenda and Dr Joy T. Adu for academic, moral and emotional support whenever life seemed unbearable to me.

Finally, I thank my family for their support: Alex, Catherine, Cecilia, Kennedy, Kenneth and Cornelius. To my son, Christian, thank you for being understanding and for always making me proud.

Abstract

Phrasal compounds such as *two-for-the-price-of-one sales* pose problems for linguistic theories because they behave like words, but clearly include syntactic phrases. The main aim of this thesis is to analyse two types of phrasal compounds in Cinyanja, the verb-noun (VN)-compounds such as *cigona-mubawa*, 'drunk' (lit. 'it sleeps-in-bar') and associative (Assc)-compounds such as *njinga ya moto*, 'motorbike' (lit. 'bicycle of fire'). I argue that these two types of compounds differ in whether or not the phrasal parts inside the compounds are accessible to syntactic rules, i.e. whether or not they are subject to the Lexical Integrity Hypothesis. I show that the phrasal part of VN-compounds obeys Lexical Integrity and is not accessible to syntactic rules such as modification, extraction, or gapping. From this, I conclude that VN-compounds are formed by a process called reification (Harley, 2009), where a phrasal structure is reanalysed as a simple root. In contrast, the phrasal parts inside Assc-compounds are accessible to syntactic rules such as pronominalisation, modification and gapping, in violation of Lexical Integrity. I, therefore, analyse the Assc-compounds as (partly transparent) phrasal idioms. My analysis is based on previous work by Carstens (1991, 2008, 2018) on the syntax of Bantu noun phrases, Harley's (2009) analysis of phrasal compounds in the *Distributed Morphology* framework, and Jackendoff's (1997) theory of *lexical licensing*, as applied to phrasal idioms. My results inform our understanding of the processes of compound formation as well as the analysis and lexical representation of phrasal idioms.

Table of contents

List of tables

List of abbreviations

A	Adjective
a	Adjectiviser
A-A	Adjective-Adjective
ADV	Adverb
af	Affix
Agr	Agree
AM	Associative Marker
A-N	Adjective-Noun
ANP	Associative Nominal Phrases
AP	Adjective Phrase
APPL	Applicative
ART	Attributive
ASS	Associative
AV	Adjective-Verb
C	complementiser
CAUS	Causative Marker
CM	Class Marker
CP	Complementiser Phrase
CRD	Coordinative
CS	Conceptual Structure
D	Determiner
DEM	Demonstrative
D_{LOC}	Locative Determiner
DM	Distributed Morphology
DP	Determiner Phrase
DP_{LOC}	Locative Determiner Phrase
EPP	Extended Projection Principle
FV	Final Vowel
GEN	Gender
K	Kase
KP	Kase Phrase
LCS	Lexical Conceptual Structure
LF	Logical Form

LIH	Lexical Integrity Hypothesis
LOC	Locative
LPS	Lexical Phonological Structure
LSS	Lexical Syntactic Structure
Mod	Modifier
MP	Minimalist Program
n	Nominaliser
N	Noun
N-A	Noun Adjective
NC	Noun Class
NEG	Negative
N_{LOC}	Locative Noun
N-N	Noun-Noun
NP/nP	Noun Phrase
NP_{LOC}	Locative Noun Phrase
Num	Number
N-V	Noun-Verb
OM	Object Marker
P	Preposition
PA	Preposition-Adjective
PF	Phonological Form
PIC	Phase Impenetrability Condition
PN	Preposition-Noun
PossDP	Possessor Determiner Phrase
PP	Prepositional Phrase
PRF	Perfect Tense
PRN	Pronoun
PRS	Present Tense
PS	Phonological Structure
PST	Past Tense
PV	Preposition-Verb
REFL	Reflexive
RHR	Right Hand Head Rule
RM	Representational Modularity

REC	Recent past
SBJ	Subject
SBJV	Subjunctive
SG	Singular
SM	Subject marker
Spec DP	Specifier of Determiner Phrase
Spec	Specifier
$SpecDP_{LOC}$	Specifier of Locative Determiner Phrase
SS	Syntactic Structure
SUB	Subordinative
SVO	Subject-Verb-Object
SVOX	Subject-Verb-Object-Oblique
T	Tense
TP	Tense Phrase
UG	Universal Grammar
V	Verb
VI	Vocabulary Insertion
V-N	Verb-Noun
VO	Verb-Object
VP/vP	Verb Phrase
V-V	Verb-Verb

List of symbols

Ø	zero prefix
+	plus
>	to; becomes
&	and
√	root
=	equals to

CHAPTER 1: INTRODUCTION

1. General introduction

Compound words are found in many languages and are an integral part of human communication. Many studies have been conducted on compound words in different languages (see e.g. Bauer, 1983; Fabb, 1984; Lieber, 1992; Plag, 2003; Lieber, 2009b; Olsen, 2017). Compound words generally consist of more than one word, for example, *schoolbook*, *teacup* and *blackboard* which are free bases (i.e. a basic meaningful part of a word to which affixes can be added). However, a typical compound is a word that consists of other words which can themselves be complex, for example, *truck-driver*, where the word *driver* is derived from the verb *drive* by adding the suffix *–er*. Normally, compounds do not include inflectional morphology, like agreement, because compounds are formed in the lexicon where inflectional material is not allowed inside the words but only added in syntax (see Di Sciullo and William, 1987; Lipka, 1992).

However, the so-called phrasal compounds are a problem for this standard view because they allow phrases to combine with words in the formation of complex words. The phrases inside phrasal compounds may also include inflectional morphology (e.g. *-ing* in (1a); the plural *–s* in (1b)) and functional categories (e.g. *the* in (1c)), as shown in the examples.

1. a. stuff-blowing-up effects
   b. Prince-of-Thieves film
   c. state-of-the-art computer

The compounds in (1) are lexical items that are derived by combining phrases, such as the NP *state of the art*, with bare nouns such as *computer*. The outcome is nominal compounds that have the same distribution as other compound nouns in English.

Another category of lexical items that includes phrasal structure, possibly with functional material, is idioms. These include expressions such as;

2. a. kick the bucket
   b. spill the beans

While the examples in (1) are compound words which include phrasal material, the idioms in (2) are themselves phrasal constructions, namely VPs. These VPs have the regular syntax of transparent VPs in English, but have been associated with special meanings, which makes these expressions idiomatic.

In this thesis, I analyse two types of compound constructions in Cinyanja**,** a language spoken in Zambia. The first is the verb-noun (VN)-compound, for example, *cigona-mubawa*, 'drunk' which literally means 'it sleeps-in-bar' where 'sleeps in the bar' is a verb phrase (VP). The second is the associative (Assc)-compound, for example, *njinga ya moto*, 'motorbike', which is literally translated as 'bicycle of fire'. Both types of compound capture the fact that idiosyncratic lexical information can be associated with complex phrasal structures. However, I argue in this thesis that these two types of compounds differ with respect to how phrasal structure is incorporated into the expression. What I will show is that Cinyanja verb-noun compounds are genuine phrasal compounds, where a phrasal constituent is part of a compound word. I show that in these compounds, the phrasal structure behaves like a syntactic atom; arguably, syntactic rules have been used to produce the phrasal structure, but the outcome of the rules has been reanalysed ("reified") as a lexically and syntactically basic unit, a word or root (see Harley, 2009). Consequently, the phrasal structure is completely invisible for any syntactic operation, and the phrasal compound behaves like any other word. In contrast, I argue that associative compounds are phrasal idioms: they are phrasal constructions with the syntax of regular associative constructions in Cinyanja, but they have become lexicalised and are represented as phrasal lexical items with special meaning. In these phrasal "compounds", the phrasal structure remains phrasal, and only the semantics is no longer compositional. I demonstrate that, while the non-compositional semantics may prevent the phrasal constituents of the compound from being completely transparent, the compound still behaves like a regular phrasal construct in the syntax in many important respects. In particular, I show that phrasal parts of associative compounds are accessible to syntactic rules, under specific conditions that will be discussed in detail in this thesis.

This chapter introduces a brief sketch of the sociolinguistics and grammatical properties of the Cinyanja language in section 1.1. In section 1.2, I discuss the research problem and present the research questions and objectives which guide this study. Section 1.3, provides information on the sources of data and how the data was classified, and section 1.4 discusses an overview of some of the key assumptions that are made in the theories that have been used in the thesis. The last section of the chapter outlines the different chapters and what each covers.

1.1 Brief sketch of Cinyanja

As stated in the title, this study is designed to provide a detailed analysis of phrasal compounds in Cinyanja. The language is called "Cinyanja" by the Nyanja people and the name *Nyanja* means "lake". Cinyanja is thus a language of the lake dwellers, the Chewa people and others who have adopted it (Englund, 2001). As with other languages, *ci-* attached to *Nyanja*, is a class 7 prefix attached to language nouns. Sometimes the Cinyanja language is simply referred to as "Nyanja" without the prefix.

Cinyanja is a Bantu language spoken in the southern and south-central regions of Africa. Like other Bantu languages, it is grouped according to language zones as devised by Malcolm Guthrie in the 1940s. According to Guthrie (1967), the Nyanja/ Chewa language group belongs to Zone N31 with Cinyanja and Chichewa as sub-groups. Nyanja is classified as N31a and Chewa as N31b. In terms of geographical distribution, Cinyanja is a Bantu language spoken in Zambia, Mozambique and Zimbabwe while in Malawi where it is widely spoken, it is called Chichewa. The Cinyanja/ Chichewa language group, as it is mostly referred to, is spoken by approximately 15 million people (Paas, 2005). In Zambia, based on the 2000 census statistics, 20.6 per cent of the population speak Cinyanja as a first language (Marten and Kula, 2007: 295).

In Zambia, Cinyanja is one of the 7 local languages that were designated official status by region alongside Bemba (spoken in the Northern Province, Muchinga Province, Luapula Province, Copperbelt Province and some parts of Central Province), Kaonde, Lunda and Luvale (spoken in the North-western Province), Lozi (used in the Western Province) and Tonga (spoken in the Southern Province). According to this linguistic zoning in Zambia, Cinyanja is mainly spoken in the Eastern Province of the country (Kashoki, 1978, Miti, 2001 and Mchombo, 2004). It is also the designated regional official language for the Eastern Province and Lusaka Province (i.e. the capital city of Zambia) alongside English. Although the government of the Republic of Zambia recognises Cinyanja and other local languages as official, the recognition is only by proclamation and not enshrined in the constitution of Zambia which recognises English as the only official language in the country. The official status which English enjoys, gives it elevated status in different domains, for example, commerce, education and media especially in urban settings (Marten and Kula, 2007). Although Cinyanja does not have the status that English has in Zambia, it is also used in some domains such as local court proceedings, business places, and religious gatherings. Other key areas where Cinyanja is used include the media and education.

Cinyanja is used in broadcast and the print media. In the broadcast media, programmes like local news, some documentaries, entertainment (e.g. drama, music), and cultural programmes are aired in Cinyanja (Hamungole, 2015). On radio, a slot is given for local news in Cinyanja. Additionally, some radio programmes ranging from political and economic topics to some phone-in programmes where callers comment on different issues that concern the general public (see also Marten and Kula, 2007) are also presented in Cinyanja. In the print media, a magazine entitled "Tsopano" was printed periodically in Cinyanja. Some books are also printed in Cinyanja although the majority of these books are used for educational purposes in schools. In education, Cinyanja is mainly used in the Eastern Province and some public schools in Lusaka as the medium of instruction in the first four grades of primary education, and thereafter as a subject from upper primary to secondary school. At tertiary level, Cinyanja is taught through the medium of English only to those specialising in African languages.

Having presented a brief sociolinguistic overview of Cinyanja, I now turn to Cinyanja grammar. Like most Bantu languages, Cinyanja is a tonal language where the pitch of the syllables is significant. There are two types of tones in Cinyanja; High and Low. Two tone patterns are distinct in Cinyanja; the first is the lexical tone which distinguishes the lexical meaning of a word, for example, *mténgo* 'tree' and *mtengo* 'price' (see also Kiso, 2012). The second is grammatical tone, which is a tone that shows grammatical relationships notably on verbs and other functions. For instance, the verb *–fotokoza* 'explain' has no tone but can obtain tone when used to distinguish the simple past as in (3a) and the recent past as in (3b).

3. a. ndi-ná-fótokoz-a[1]
   1-PST-explain-FV
   'I explained'

   b. ndi-na-fótókoz-a
   1-REC.PST-explain-FV
   'I explained recently'
   (Mtenje, 1987:172, Kiso, 2012:17)

The examples above show that tone is a salient feature in distinguishing lexical and grammatical meaning. However, I do not represent tone in my thesis because it is not relevant to my topic. Moreover, tone is not marked in the orthographies in Cinyanja.

[1]I will adapt the glosses from the examples taken from the literature to the system used in this thesis.

As a Bantu language, Cinyanja has an agglutinative morphology of words (Guthrie, 1971), especially in the area of verbal inflection. Generally, words contain different morphemes which play a role in determining the meaning as shown below.

4. a. A-na-dzi-thyol-a
1.SM-PST-REFL-break-FV
'She fractured herself'.

b. M-nyamata a-na-thyol-a m-pando.
1-boy 1.SM-PST-break-FV 3-chair
'The boy broke the chair'.

c. M-nyamata a-na-u-thyol-a m-pando.
1-boy 1.SM-PST-3.OM-break-FV 3-chair
'The boy broke (it) the chair'.

d. M-nyamata sa-na-u-thyol-e m-pando.
1-boy 1.SM.NEG-PST-3.OM-break-SBJV 3-chair
'The boy did not break (it) the chair'.

In the above examples, the affixes (e.g. subject markers, tense markers, reflexive markers, object markers, negation affixes) attached to the verbal structure contribute to the meaning of the word or clause.

The basic clause structure in Bantu languages is subject > verb > object (SVO). Cinyanja fits in this typology of word order where a noun phrase functioning as a grammatical subject precedes a verb and the one functioning as a grammatical object follows the verb as in (5).

5. a. M-nyamata (S) w-a-gul-a (V) galimoto. (O)
1-boy 1.SM-PST-buy-FV 5.car
'The boyhas bought a car'.

b. Ci-gawenga (S) c-a-ph-a (V) m-phunzitsi. (O)
7-thug 7.SM-PRF-kill-FV 1.teacher
'The thug has killed a teacher'.

In (5a) the nominal subject is *mnyamata* 'boy' and the nominal object is *galimoto* 'car' while in (5b) the nominal subject is *cigawenga* 'thug' and the object is *mphunzitsi* 'teacher'.

In Cinyanja, nominal morphology displays a bimorphemic structure comprising a nominal stem and a noun class prefix which is always a bound morpheme. The noun class (NC) prefixes are morphological markers of number and gender and play a distinct role in the agreement system. Traditionally, noun classes are identified using a numbering schema where Arabic numerals are used; for example, NC1, NC2, NC3, NC4 etc. (see also, Corbett & Mtenje, 1987; Carstens, 1991; Msaka, 2019). Since the current study is based on Cinyanja phrasal compounds where noun class prefixes are a prominent feature and constantly referred to in the thesis, the understanding of the noun class system is important.

As with all other Bantu languages, nouns in Cinyanja belong to a noun class system and this system uses prefixes to express the class of different nouns (Carstens, 2008; Msaka, 2019). The classes are associated with paired numbers showing the singular and plural form of the nouns. The pairs include classes 1-10 and 12-13. Other classes are not paired but are significant in the classification of Bantu nouns (e.g. class 14 and 15) and the locative classes 16, 17 and 18. Cinyanja has 17 noun classes in total. Table 1 gives a summary of Cinyanja NC prefixes.

Table 1.1: Noun class prefixes in Cinyanja

Prefix number	**Class prefix**	**Example**
1	mu-	mu-nthu ‘person’
1a	-	Ø-tate ‘father’
2	a-	a-nthu ‘people’
3	mu-	m-tengo ‘tree’
4	mi-	mi-tengo ‘trees’
5	li/dzi-	li-wu ‘word’ dzi-la ‘egg
6	ma-	ma-wu ‘words’ ma-zila ‘eggs’
7	ci-	ci-tsime ‘well’
8	zi-	zi-tsime ‘wells’
9	n-	ng’ombe ‘cow’
10	n-	ng’ombe ‘cattle
11	--	-
12	ka-	ka-mtengo ‘a small tree’
13	ti	ti-mitengo ‘small trees’
14	u-	u-lesi ‘laziness’
15	ku-	ku-lira ‘to cry’
16	pa-	pa-mutu ‘on the head’
17	ku-	ku-mudzi ‘to the village’
18	mu-	mu-nyumba ‘in the house’

Generally, in Cinyanja, noun class 12/13 prefixes derive diminutives while noun class 7/8 prefixes derive augmentative/ pejorative meaning when attached to primary prefixes as below.

6 a m-tengo
3-tree
‘a tree’

b. mi-tengo
4-tree
‘trees’

c. ka-m-tengo
12-3-tree
'small tree'

d. ti-mi-tengo
13-4-tree
'small trees'

e. ci-m-tengo
7-3-tree
'big tree'

f. *zi-mi-tengo
8-4-tree
Intended: big trees

In (6c & d) above, NC 12/13 prefixes (i.e. singular and plural) derive a diminutive meaning when attached to nouns with primary prefixes and in (6e), NC 7 singular prefix *ci-* attached to *mtengo*, 'tree' has an augmentative meaning. Although NC 8 prefix *zi-* is the plural for NC 7 prefix in (6f), it does not attach to all nouns with ordinary primary prefixes. However, NC 12 and 7 prefixes can replace ordinary primary prefixes to derive diminutive and augmentative meanings (see also Carstens, 2008 discussion for Swahili) as shown in the example below.

7. a. Ma-thumba anga a-a-gw-a.
6-bag 6.my 6-PRF-fall-FV
'My bags have fallen'.

b. ka-thumba k-anga k-a-gw-a.
12-bag 12-my 12-PRF-fall-FV
'My small bag has fallen'.

c. ci-thumba c-anga c-a-gw-a
7-bag 13-my 13-PRF-fall-FV
'My big bag has fallen'.

In (7) above, when NC 12 prefix *ka-* in (7b) replaces NC 6 prefix *ma-* in (7a), the noun *thumba* 'bag' has a diminutive meaning, and in (7c) where NC 7 prefix *ci-,* replaces *ma-* the noun *thumba* has a pejorative/ augmentative meaning in Cinyanja.

In some cases, nouns have a zero prefix and the only way to know the class is by establishing the concordial/ agreement it controls in a phrase or sentence (see Miti, 2001; Msaka 2019) as shown in the example below.

8. Tebulo la-li-kulu l-a-thyok-a.
   5.table 5.ASS-5-big 5.SM-BE.PRS-break-FV
   'The big table is broken'

In (8) above, *tebulo* 'table' is classified as a class 5 noun based on the concordial agreement it governs in the sentence.

As (8) already illustrates, a prominent feature of nouns in Bantu languages is that they control agreement on verbs and other sentence modifiers. (9) illustrates this further.

9. a. M-nyamata u-ja wa-m-tali w-olemera w-a-gul-a
   1-boy 1-that 1.ASS-1-tall 1.ASS-rich 1.SM-PRF-buy-FV
   galimoto.
   5.vehicle
   'That tall rich boy has bought a vehicle'.

   b. Ana a-kazi onse a-wa okongola a mfumu
   2.girl 2-female 2.all 2-these 2.beautiful 2.ASS 1.king
   a-ku-phik-a ca-kudya.
   2.SM-PRS-cook-FV 7-food
   'All these beautiful daughters of the king are cooking food'.

The examples in (9) above show that the verb and other sentence modifiers agree in noun class with the subject noun. The verbal prefixes, for example, the subject markers (SM) and object markers (OM) are agreement markers.

1.2 Research problem

As stated at the beginning of this introduction, this study analyses two types of phrasal compounds in Cinyanja. The first is made up of a verb and a noun- VN-compounds:

10. a. m-pala-ma-tabwa
1-scrape-6-plank
'plank scraper'
'carpenter'

b. ci-gona-mu-bawa
7-sleep-18.LOC.5bar
'it sleeps in the bar'
'drunk'

The second type of phrasal compound comprises two nouns linked by an associative marker- Assc-compounds:

11. a. njinga ya moto
9.bicycle 9.ASS 3.fire
'bicycle of fire'
'motorbike'

b. mw-ana wa leza
1-child 1.ASS 1.god
'child of god'
'rainbow'

The two types of phrasal compounds presented above raise questions as regards their status because they are clearly lexicalised expressions. Further, they also include agreement markers, for example, in (11b) *mwana wa leza* 'child of god', the subject noun is marked with the noun class 1 prefix and controls agreement (*w-*) on the associative marker *-a* 'of', within this expression. Also, in example (10), the phrasal compounds include inflectional material, for example, noun class 6 *ma-* in *mpala-**ma**tabwa*, and the class 18 locative prefix *mu-* in c*igona-**mu**bawa*. All these issues raise questions as regards the analysis of these compound expressions.

As shown above, there are two ways in which phrasal material can appear inside a lexical entry. Either the lexical item is a phrasal compound (for example *stuff-blowing-up effects, over-the-fence gossip*) or it is a phrasal idiom (e.g. *spill the beans*, *bury the hatchet*). Although both types of lexical items include a well-formed phrasal part, syntactic rules do not apply in the same way to these phrases. In phrasal compounds, a part of the compound expression cannot be manipulated by syntactic operations such as modification:

12. a. *bikini-**best**-girls-in-trouble genre
    b. best bikini-girls-in trouble genre

Example (12) shows that a modifier such as *best* cannot be inserted between the constituent parts of the phrasal compound as in (12a). Instead, modification of the phrasal compounds affects the entire compound as in (12b). This can be contrasted with the phrasal idioms in (13) below, where modification of a part of the phrase is possible.

13. a. bury the hatchet 'reconcile a disagreement
    b. bury the *political* hatchet 'reconcile a political disagreement '

In (13b) above, a part of the phrasal idioms is modified by an adjective, whose meaning can then be incorporated into the special meaning of the idiom. While not all idioms allow this type of modification, this example shows that syntactic operations sometimes have access to the internal structure of these idioms.

The constraint that rules out the modification of the phrasal part of the compounds in (12) is referred to as the *Lexical Integrity Hypothesis*. This hypothesis states that syntax cannot have access to the internal structure of words. It is built from Chomsky's (1970) lexicalist hypothesis and has been advanced by other linguists such as Lapointe (1980) in (14a) and Di Sciullo and William (1987); others are Bresnan and Mchombo, (1995). The definition in (14) provides two prominent formulations of the Lexical Integrity Hypothesis:

14. *The Lexical Integrity Hypothesis*
    a. *Generalised lexical hypothesis* (Lapointe, 1980)
       No syntactic rule can refer to elements of morphological structure

b. *The thesis of atomicity of words* (Di Sciullo and William, 1987)
Words are “atomic” at the level of phrasal syntax and phrasal semantics. The words have “features,” or properties, but these features have no structure, and the relation of these features to the internal composition of the word cannot be relevant in syntax.

Di Sciullo and William (1987) claim that syntax cannot have access to the internal structure of complex words because it lacks the "vocabulary" to analyse morphological objects. The examples in (12) above show that this principle can never be violated in phrasal compounds. The fact that Lexical Integrity is not an issue for transparent phrasal idioms such as those shown in (13b) shows that these are not words, but complex phrasal constructions. As I discuss in my thesis, this disparity in the way syntax operates on these complex lexical expressions can also be observed with the two types of Cinyanja compounds I investigate here. I show by analysing the structure of these compounds, that VN-compounds are subject to the Lexical Integrity Hypothesis. They, therefore, meet the conditions for the analysis as genuine phrasal compounds, thus, they are words. In contrast, I show that Assc-compounds are not subject to the Lexical Integrity Hypothesis, and I conclude from this that these compounds are best analysed as phrasal idioms.

The phrasal compounds being analysed in this thesis include phrasal material such as example (12) above. Cinyanja VN-compounds are also phrasal compounds. In contrast, Assc-compounds are phrasal idioms. In this thesis, I refer to both VN-compounds and Assc-compounds as “compounds” or compound expressions but this use of the terminology implies that not every compound is necessarily a word as demonstrated by Assc-compounds which are not words. However, I use the term “word” in the sense of a syntactic atom which incorporates the idea of lexical integrity.

Having clarified the use of these terms, I want to state that my analysis of the phrasal compounds was motivated by a number of questions. The main question and the first to be addressed concerns the structure of the compound constructions I examine, particularly their internal structure. In this study, I show by analysing the two types of compounds that both are based on phrasal structure: the VN-compounds include a VP phrasal structure and the Assc-compounds include the nominal phrasal structure typical of regular associative constructions.

A second question concerns the nature of the phrasal material in these expressions. Notably, Cinyanja phrasal compounds include noun class prefixes and locative prefixes (i.e. the VN-compounds) and

associative markers (i.e. Assc-compounds). Some linguists have analysed noun class prefixes as derivational (Mletshe, 2010; Schadeberg, 2001:13; Msaka, 2019) or as having a mixed nature namely derivational and inflectional (Mufwene, 1980; Sproat, 1985, Bresnan and Mchombo, 1995), while others have analysed them as having an inflectional nature (Carstens, 2008). What these studies have not addressed is the nature of the noun class prefixes and locative prefixes inside the VN-compounds in Cinyanja. Similarly, studies on associative constructions in Bantu languages have discussed the interpretation and the syntax of these constructions (Mugane, 1997; Delfitto et al., 2011) but none of these studies has discussed the nature of the associative marker in the associative compounds in Cinyanja. In this thesis, I address this gap. I show, following Carstens (1997, 2008), that the noun class prefixes and the locative prefixes that appear inside the VN-compounds and the associative markers that appear inside the Assc-compounds are not derivational but inflectional, because they represent functional categories. This leads to the conclusion that the compound constructions I examine include not only lexical phrasal categories but fully-fledged syntactic structures that include functional material.

The third aspect addressed in this study is the morpho-syntactic status of these phrasal compounds. Superficially, the two types of phrasal compounds comprise general properties of compound words: they are complex expressions composed of more than two elements, they are generic (i.e. the non-head is non-referential) and their constituent parts cannot be separated (Plag, 2003; Altakhaine, 2016). However, given that these expressions are based on or incorporate phrasal structure, the question arises regarding the status of these complex expressions either as words or phrases. To answer this question, I examine in detail the impact of the Lexical Integrity Hypothesis on phrasal compounds in Cinyanja. Different tests (e.g. pronominalisation, modification and inseparability) will be used to ascertain if, or to what extent, the phrasal material incorporated into the compounds is accessible to rules of syntax. The study will show that lexical integrity impacts the two phrasal compounds differently. Despite the two phrasal compounds having a resemblance to compound words, I will show that the two complex structures are different. In the VN-compounds, lexical integrity is not violated, and I therefore argue that these are genuine (phrasal) compounds. In contrast, I demonstrate that Assc-compounds tolerate violations of lexical integrity, and I therefore argue that these compounds are phrasal idioms.

To summarise, this study contributes to the existing body of knowledge in several ways. The first contribution is that it provides information on the internal structure of the two types of phrasal compounds which to the best of my knowledge have not been investigated in Cinyanja. I show in this thesis that the VN-compounds are formed when a VP is nominalised. For the Assc-compounds, I show

that they have the regular syntax of associative phrases, which are formed by linking a head noun with a possessor DP using the associative marker.

The second contribution is that my study provides information on the noun class prefixes, locative prefixes and associative markers found inside these phrasal compounds. In this study, I show that these elements inside the compounds belong to functional categories and that they are therefore, inflectional.

The third contribution is that this study also offers a syntactic analysis of the phrasal compounds and their internal phrasal structure. The study reveals the two different syntactic representations of the two phrasal compounds. It shows that the phrasal part in VN-compounds cannot be accessible to syntactic rules and behaves like a syntactic atom. I argue that VN-compounds involve the reanalysis ("reification") of the phrasal structure as a lexical root and that this reified root forms the basis on which the phrasal compounds are formed. In contrast, the phrasal parts of Assc-compounds are to some extent accessible to syntactic rules. They are analysed as phrasal idioms. I propose a lexical representation of these compound constructions which links their special meaning to their phrasal syntactic representation, and I show how some of the intriguing properties of Assc-compounds that I discuss follow from their lexical representation. This analysis of Assc-compounds as phrasal idioms is a novel contribution to Bantu linguistics.

Finally, the fourth contribution of my thesis is empirical. I have collected numerous examples of Cinyanja compound constructions (see section 1.3 below on methods of data collection), and their properties are examined in detail in this thesis. This examination has revealed some curious aspects of the interpretation as well as the grammatical status of these expressions which to the best of my knowledge have not been noted in Bantu research before.

1.3 Methods of data collection

As stated above, this study is centered on two types of Cinyanja phrasal compounds, namely; VN-compounds and Assc-compounds. Data on these complex expressions was collected using secondary sources and this was achieved by reading through these sources to identify the two types of phrasal compounds. The following were the main data sources for the phrasal compounds in Cinyanja; *Chinyanja (Chichewa)-English Dictionary* (Vermeullen, 1979); *English-Chinyanja Dictionary* (Paas, 2005); *Chichewa/Chinyanja English Dictionary* (Paas, 2013) and *Oxford Chichewa-English/ English-*

Chichewa Dictionary (Paas, 2016). Other sources included; *The Syntax of Chichewa (*Mchombo, 2004:117-118)[2] and reviewed articles in Bantu languages. Further, I used my personal knowledge of Cinyanja to supplement secondary data**.** I should indicate, in this regard, that as a native speaker of Cinyanja with mother tongue competence of the language, I was able to add some words to the data.

The main purpose of using secondary sources was to extract compounds with a noun-noun (NN) and also a verb-noun (VN) combination from previous research. In the method adopted for this research, a mini-corpus of NN-compounds and VN-compounds that include inflectional material and are therefore phrasal was compiled. The mini-corpus was further divided into three different categories based on the position and type of inflectional material.

The first category of compounds consisted of VN-compounds with an initial noun class prefix that consists of a verb and a noun without inflectional material internal to the compound words:

15. cipha-dzuwa
ci-ph-a-dzuwa
7-kill-FV-5.sun
'killer of the sun' (literal meaning)
'a beautiful girl' (lexicalised meaning)

The second category of phrasal compounds comprised VN-compounds with visible class markers internal to the compound words:

16. mukonda-maungu
m-kond-a-ma-ungu
1-love-FV-6-pumpkin
'lover of pumpkins' (literal meaning)
'pumpkin lovers' (lexicalised meaning)

The third category of phrasal compounds comprised two nouns linked with an associative marker which are lexicalised as shown below.

[2]Mchombo (1978) PhD thesis would possibly have been a good source of data, but I did not manage to find it.

17.	Njinga	ya	moto	
	9.bicycle	9.ASS	3.fire	
	'bicycle of fire'			(literal meaning)
	'motorbike'			(lexicalised meaning)

The selection and categorisation were based on the general definition of a compound word as a combination of two or more words (Fabb, 1998:66; Plag, 2003:5); where the term "word" is used to also refer to meaningful linguistic units that cannot be further segmented (Di Sciullo and Williams, 1987; Lieber, 2009a). In this study, names of people (kinship terms) and names of buildings (e.g. shops) that met this definition of a compound word were also considered compounds.

1.4 Theoretical framework

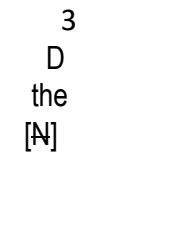
1.4.1 Principles and Parameters

In this study, I apply some of the ideas and assumptions of the Principles and Parameters theory (Chomsky, 1981, 1986), including ideas expressed in the Minimalist Program (MP; Chomsky 1995 and subsequent work). The Principles and Parameters approach to language is based on the proposition that there are universal principles that apply to all natural languages. While there are certain principles of grammar that are universal, parameters explain language variation. For one to have mastery of a language, one needs to have knowledge of the universal principles, and also to acquire the parameter settings for that language (Chomsky, 1986, 1995; Haegeman, 1994). Chomsky's (1995) MP is a natural extension of the Principle and Parameters theory which deals with principles of economy, simplicity and naturalness.

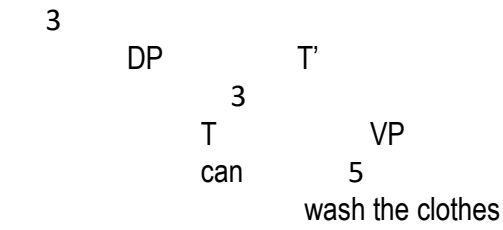
One of the key syntactic principles of the Principles and Parameters theory in the MP is the Binarity Principle, which states that all phrases are binary branching. Syntactic constructions such as phrases and clauses/sentences are formed through the operations of Merge and Move. Merge is a process where a constituent combines with another constituent to form a new constituent. Chomsky (1995) notes that the simplest type of Merge is where two elements X and Y are combined to form a new element Z. Chomsky (2001:18, 2008:139) identifies two types of Merge. The first is *external* Merge where the two elements X and Y are separate objects from the lexicon. The second type is *internal* Merge, where Y is already part of X and then re-merged with X as a copy. Internal Merge is equivalent to Move.

Syntactic structures such as phrases or sentences are built up through a series of merger operations. Since words have selectional features, they specify which categories they can merge with. For example, the determiner *the* is a D-head and has an uninterpretable selection N feature. This indicates that it combines with a noun or NP to form a DP. When D merges with the NP complement, the selection feature of [N] of D is deleted as shown in (18a). In (18b), the modal *can* has a sectional feature of V which determines that the modal (which belongs to the functional category T) combines with a verb or VP (Radford, 2004:58-59):

18 a.



b.



A fundamental operation of the Minimalist Program is the operation Agree (Chomsky 2000), which establishes a relation between two elements that share certain grammatical features (Kremers, 2003:6). Agree involves two core elements, a *probe* (i.e. an uninterpretable feature of a functional category) and a *goal* (i.e. the corresponding interpretable feature of another functional category). To enter into an Agree relation, the category which includes the probe and the goal has to be active. A goal category is active when it has an uninterpretable feature. A probe probes for an active goal (the same matching feature as the probe, but valued). When the probe has found a goal, the goal values the probe, and agreement is established. Following Carstens (2005) and Baker (2003, 2008), I assume that noun class agreement in Bantu languages such as Cinyanja is also established from underlying probe-goal relations. I illustrate this using Cinyanja examples:

19. a. M-tsikana w-a-pwany-a ma-zila.
1-girl 1.SM-PRF-break-FV 6-egg
'The girl has broken the eggs'.

b. *M-tsikana a-a-pwany-a ma-zila.
1-girl 6.SM-PRF-break-FV 6-egg
Intended: 'The girl has broken the eggs'.

In (19a), the verb agrees with the subject noun rather than the object DP. The subject has phi-features (phi-features are person, gender and number) expressed in the form of noun class features. For agreement to be established, the probe has to look for the goal with related phi-features within its c-command domain to agree with. In (19a), the subject DP has valued the phi-features of T, which are expressed on the verb. Therefore, an agreement relation is established between the probe *wapwanya* and the goal *mtsikana*. Notice that the goal (the subject-DP) c-commands the probe (T) in (19a). This is in consonant with Carstens (2005) and Baker (2008) that heads in Bantu agree upwards. In (19b), where the object DP shares the phi-features with the verb, there is no upward agree relation established and the sentence is ungrammatical.

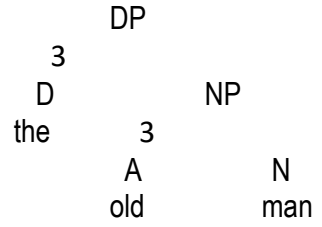
1.4.2 The DP-hypothesis

Since this thesis is mainly concerned with nominal expressions, it is important to carefully consider the syntax of noun phrases. Traditionally, the noun is viewed as the head of the noun phrase (NP), for example, *man* in *old man* is the head of this phrase. However, according to Abney's (1987) DP-hypothesis, a nominal expression is analysed as a determiner phrase (DP) which is headed by a determiner (D). The D-head takes an NP as its complement. Determiners include elements such as articles, demonstratives and genitive pronouns. In English, determiners precede nouns as shown in the examples below.

20. a. the car
    b. this car
    c. my car

Example (20) above shows that the combination of the determiner, for example, *the* in (20a) and the content word *car,* which projects an NP, forms the DP *the car*. Abney (1987) notes that determiners are functional heads of a nominal type, and the DP is an *extended projection* of the noun:

21. a. the old man

    b.


In Abney's (1987) DP-analysis, the D-node is considered the site of the grammatical features of the noun, namely person and number (the so-called phi-features, which also include gender; Chomsky 1995). The idea that each of these phi-features is represented by a separate functional category is introduced in Chapter 3.

It is important to note that even in constructions or languages without overt determiners, the noun phrase is a DP whose head bears the grammatical phi-features. This assumption is relevant for a Bantu language such as Cinyanja, which does not have overt determiners.

In English, demonstratives are determiners and therefore cannot co-occur with other determiner elements at the same time in the same environment in a DP as shown in the examples:

22. a *this the car
    b. *my the car
    c. *this my car

In contrast, demonstratives in Bantu languages are typically not analysed as determiners, but as modifiers and therefore do not occupy the D position. In Cinyanja, demonstratives typically follow the head noun, as shown in the examples below:

23. mw-ana **u-yu**
    1-child 1-this
    'this child'

Pronouns like nouns are also analysed as Ds/DPs. The functional structure of the DP in Bantu is introduced in Chapter 3 and the syntax of nominal modification in Bantu is discussed in detail in Chapter 4.

In the next section, I discuss the clausal architecture.

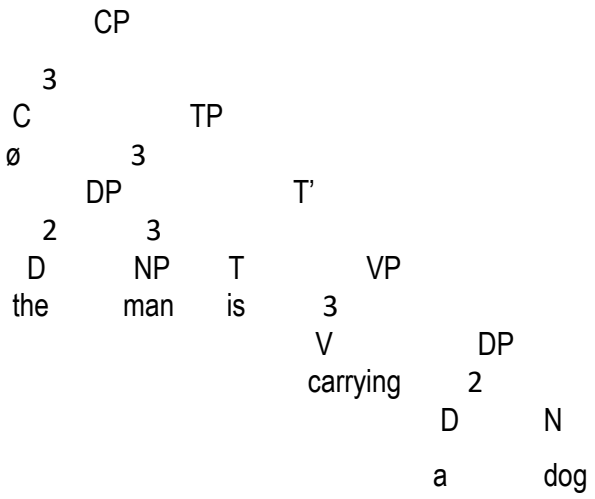
1.4.3 Clausal architecture and agreement

In this section, I briefly present the clause architecture according to Chomsky (1995). One of the principles of universal grammar (UG) is that each syntactic structure must be binary. Clauses, like

other phrases, are formed by the same binary merger operation (Radford, 2004) as shown in (24b) which represents the example in (24a).

24. a. The man is carrying a dog.

b.



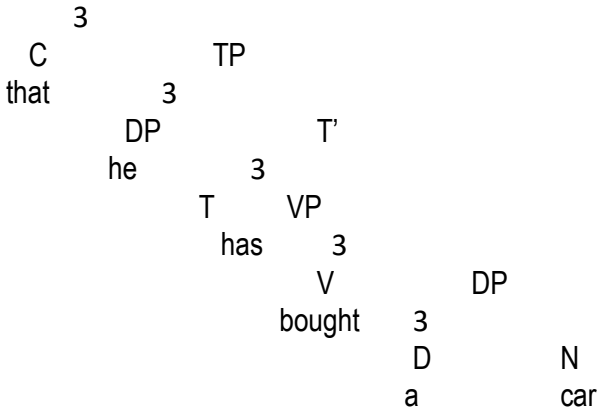
(24b) is a simplified representation of the sentence in (24a), which will shortly be revised. It shows that the V is the head of the VP, and this VP merges with an auxiliary. Auxiliaries are represented as belonging to category tense (T). T projects to form a T-bar (T'). The T' merges with the subject-DP to form the extended projection TP. In the TP clause, the DP *the man* is the specifier of T. So in (24b) above, the CP is the projection of C and C is the head of the entire clause.

The structure in (24) is representative of the main clause, but recall that a sentence can also have an embedded clause with a subject and finite verb. An embedded clause is a clause within a clause and it is introduced by a complementiser (*that, for* or *if*) in English. It is also possible to have an embedded clause with a null complementiser as in (25c).

25. a. I believe [that he has bought a car].
    b. I wonder [if Thembi would cook].
    c. Thembi said [ø she would bake the scones].

In (25a) above, the embedded clause is a complementiser (CP)-clause *that he has bought a car*. The complementiser merges with the TP and projects its own phrase, a CP.

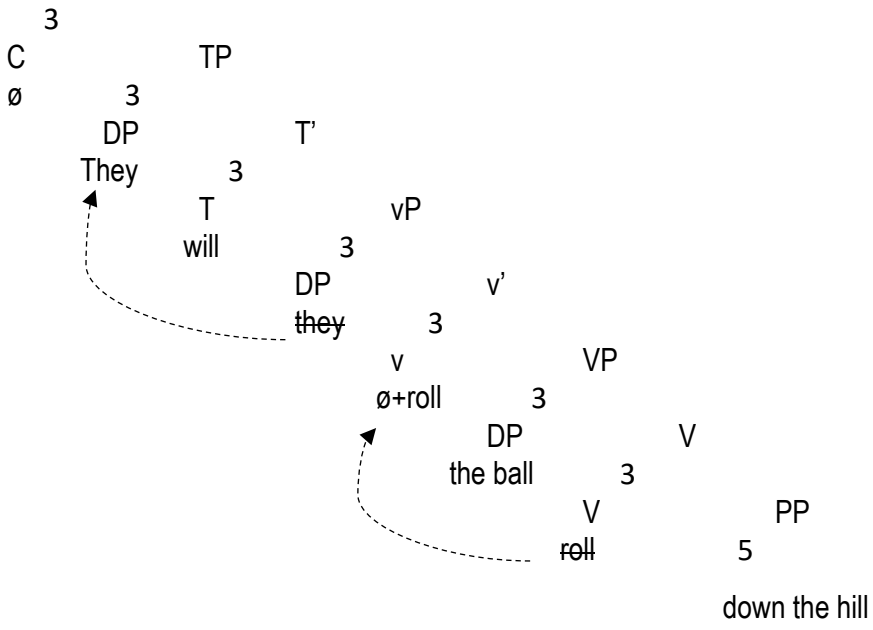
26. CP



The example in (26) shows that the complementiser (C) is the head of the complementiser phrase (CP). In the MP, all sentences are represented as CPs, including main clauses, where C is null. This is illustrated in (27b) to represent (27a).

27. a. They will roll the ball down the hill.

b. CP



(27b) represents the sentence in (27a) as a CP, with C taking the TP as its complement. (27b) represents two more innovations from (24b) above. First, it includes another projection between T and VP, that of the so-called light verb (little v). The role of the little v is to introduce the external argument in its specifier and to assign accusative case to the object-DP in the VP. According to Chomsky (2007), the little v has strong affixal features; as a result, it attracts the lexical verb, which moves to v and adjoins to it as shown in (27b) above.

The second new aspect of (27b) is subject movement. The subject is not base-generated in SpecT, as in (24b), but in the specifier of little v, and moves to SpecT. Since T has an extended projection principle (EPP) feature, it triggers movement of the DP *they* to Spec, TP to satisfy the EPP requirement.

Chomsky (2000, 2001, 2008) suggests that syntactic derivations proceed in cycles which are determined by so-called *phases*. According to Chomsky, a phase is a syntactic category corresponding to a complete proposition, and he, therefore, identifies the CP and a transitive vP (little verb phrase) as phases. Once one of these phases has been completed, the domain of the phase head (VP in the case of v; TP in the case of C) is transferred to the phonological component (PF) and the semantic component (LF). The syntactic derivation proceeds, but as a result of Transfer, the material inside the domain of the relevant phase head has become inaccessible to further derivations. In this study, I assume, following Abney (1987) and Pesetsky and Torrego (2001) that DP, like vP and CP, is also a phase. More details of derivations by phase as they relate to phrasal compounds are presented in Chapter 3.

1.4.4 Distributed Morphology and Lexical Licensing

In this section, I focus on Halle and Marantz's (1993) and Marantz's (1997) proposition of late insertion in Distributed Morphology and Jackendoff's (1997) theory of Lexical Licensing where the three components of grammar (syntax, phonology and semantics) are linked through correspondence rules.

Distributed Morphology (DM) is an alternative to the lexicalist approach which holds that morphology/ the lexicon feed into syntax. Halle and Marantz (1993) are the precursors of the DM-framework. DM is concerned with the relationship that exists between words and their internal structures that are generated in the syntax. The basic principle of DM is that there is a unique generative system (the syntax) that is involved in the formation of both complex words and complex phrases (Halle and Marantz, 1993; Harley and Noyer, 1999; Embick and Noyer, 2007).

The main idea in DM is that morphology is not associated with a single place in the architecture of grammar, rather, it is 'distributed' among multiple components of grammar (Halle and Marantz, 1993). Therefore, in processing a particular derivation, three lists are accessed which are a substitute for the lexicon. The first list is the "Narrow lexicon" (Marantz, 1997) which provides primitives to the syntax to

operate with. These primitives are functional morphemes and lexical roots, the terminal nodes in syntax (Harley and Noyer, 1999). Roots are open class categories and traditionally include verbs, nouns and adjectives. Harley (2009) notes that roots can be perceived as having a purely lexical content the interpretation whereof may vary depending on the context they are used in. In current versions of DM, roots are assumed to be acategorial in nature and only realise their category when they merge with a category-defining feature bundle, i.e. a functional head which determines the category of the root it merges with. These heads are "little" n, a, or v (see also Marantz, 1997; Zeller, 2017b) which can also be realised by affixes. For example, a root √HOUSE can be interpreted as a noun when it merges with a categorising head n as in (28a), or as a verb when it merges with a category-defining head v as in (28b) (Okubo, 2013; Harley, 2014). These heads determine the higher functional categories (e.g. T, D) which in turn determine the number and tense morphology.

28. a. The *houses* on the hill are massive.
    b. The old man *housed* strangers.

In this list, the syntactic terminals, the roots and functional morphemes are not supplied with phonological features (Embick & Noyer, 2007). These features are given in the second list.

The second list in DM is the Vocabulary, which provides the syntactic terminal nodes (i.e. both roots and functional morphemes) from list one with language-specific strings of phonological content through a mechanism called vocabulary insertion. The Vocabulary is considered a list of phonological realisations (i.e. exponents) inserted late, i.e. after syntactic operations, in the mapping to PF. Different vocabulary items that match the node targeted by vocabulary insertion compete for insertion, giving rise to contextual allomorphy. For instance, if a nominalised root √CAT merges with a functional head (say, Num) that is specified for plural number, Num needs to be realised by a plural inflection. In English, different morphemes will be available for consideration, among them the *-s* morpheme, the null *Ø* morpheme and the *–en* morpheme. However, the *Ø* and the *–en* morphemes can be chosen only in specific environments, as they are specified for the specific roots that license these exponents (e.g. √OX for *–en*, or √SHEEP for *Ø*). Since the root √CAT is not among the roots that license these vocabulary items, the –s morpheme as the least specified vocabulary item must be chosen for insertion and wins the competition. Vocabulary insertion then connects the phonological feature with the grammatical features that match the syntactic terminal node Num (Marantz, 1997).

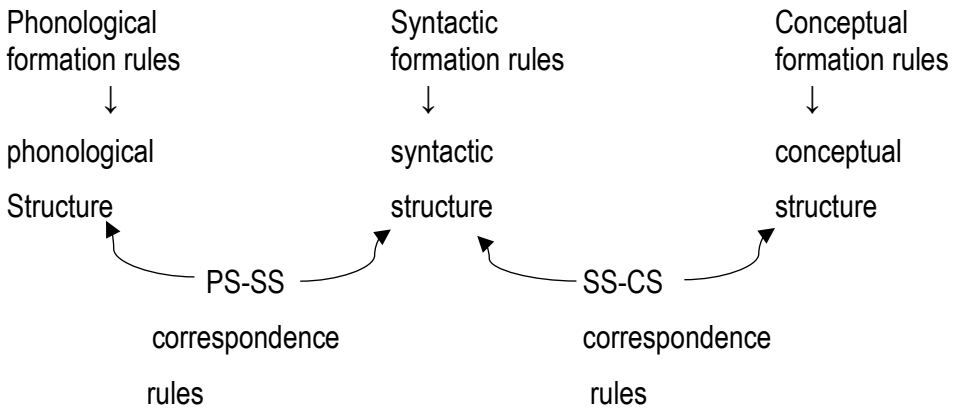
The third list, namely the encyclopaedia, is a list of semantic information associated with terminal nodes or larger syntactic objects such as phrasal idioms (Embick and Noyer, 2007). The morphemes

particularly in list 1, are considered the building blocks of words, phrases and eventually sentences in DM.

As regards the analysis of compounds in this study, and following the DM framework, I argue in Chapter 3, based on an idea articulated in Harley (2009), that VN-compounds in Cinyanja involve the reanalysis of phrasal material as roots. These roots then require a categorising functional head to determine their grammatical status. The difference between VN-compounds and Assc-compounds (discussed in Chapter 4) is that the latter are not root words but phrasal idioms, which are represented as complex phrases in the lexicon. The lexicon is considered a storehouse of all lexical information that syntax operates on and includes words and complex structures such as phrases.

My analysis of phrasal idioms in Chapter 4 is based on the theory of idioms and Lexical Licensing developed in Jackendoff (1997). According to Jackendoff (1997), a lexical item is interpreted as establishing a correspondence between three different modules of grammar; phonological structure (PS), syntactic structure (SS) and conceptual structure (CS). According to Jackendoff (1997:38), the three components form linguistic knowledge (i.e. grammar).

29. Architecture of linguistic knowledge



(Jackendoff, 1997:39)

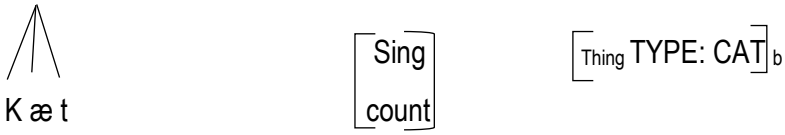
The structure above shows how information structure is organised in the mind and Jackendoff calls this Representational Modularity (RM). The structure represents the fact that each of the modules is distinct; no phonological information can be in the syntax and no conceptual information in the phonological structure.

As stated above, Jackendoff (1997) considers the three generative systems of grammar (i.e, phonology, syntax and semantics) to be independent and to operate on distinct rules but connected

through interfaces (see also Ackema and Neeleman, 2004). According to Jackendoff, the link between the interfaces of the three systems (PS, SS and CS) is provided by *correspondence rules*. These correspondence rules can be general, and provide information on how the components interact, for instance, how phonology interacts with syntax and syntax with the conceptual structure.

For Jackendoff, the lexicon is part of the interface component of grammar. Lexical items license the correspondence between parts of phonology, parts of syntax and parts of conceptual structure.[3] For example, a lexical item such as "cat" can be represented as connecting three different types of structure, the lexical phonological structure (LPS), the lexical syntactic structure (LSS) and the lexical conceptual structure, (LCS), as shown in (30):

30. a. LPS $Word_b$ b. LSS N_b c. LCS



In (30) above, the lexical item *cat* is represented by three parts of information linked to each other by a subscript index. The lexical index b represents the fact that the phonological structure in (30a) corresponds to a singular count noun in syntactic structure in (30b), and to the concept of a CAT in (30c). Importantly, the format of lexical representation for simple words also applies to complex lexical structures such as idioms whose meaning is learnt and perceived to be listed in a person's mental lexicon. Since associative compounds are considered to be phrasal lexical elements that have the status of idioms, I use Jackendoff's theory of lexical licensing to analyse and represent these complex structures in Chapter 4.

1.5 Structure of the thesis

This thesis comprises five chapters and is organised as follows:

Chapter 1 is the general introduction of the thesis. It has provided a sociolinguistic and grammatical sketch of Cinyanja, discussed the research problem, questions and objectives, described the methods of data collection and presented the theoretical framework that is adopted in my thesis.

[3]With respect to the view demonstrated above, the theory of Lexical Licensing is compatible with DM, which also assumes that semantic and phonological information is linked to terminal nodes of syntactic structures. The LPS is associated with vocabulary items, the LSS relates to the syntactic terminal nodes while the LCS corresponds to the encyclopaedia.

Chapter 2 provides an introduction to compounding and compound words. It discusses the theory of compounding where different definitions of compound words are analysed. This is followed by a discussion of the key features of compound words. Phrasal compounds and other types of compounds as well as compounding patterns in English are also reviewed.

In Chapter 3, I discuss VN-compounds in Cinyanja. These VN-compounds include noun class prefixes and locative prefixes. The chapter argues that the noun class prefixes and the locative prefixes inside the VN-compounds are inflectional. The chapter also investigates the internal structure of these phrasal compounds and argues that they have a VP phrasal structure inside them which is nominalised by a noun class prefix to form the VN-compound. I also argue in this chapter that the phrasal part is a phase which is reanalysed as a root and re-enters the derivation as an unanalysable unit.

Chapter 4 discusses Assc-compounds, which superficially look like two nouns linked by an associative marker. The chapter investigates the internal structure of the Assc-compounds and concludes that they have a phrasal structure. I show that Assc-compounds have mixed properties with respect to Lexical Integrity. While the head noun and the phrasal structure in Assc-compounds cannot be interrupted, the possessor DP can be accessed by syntactic operations. Based on these properties I argue that Assc-compounds are phrasal idioms.

Finally, Chapter 5 presents the conclusion and recommendations for further research.

CHAPTER 2: COMPOUND WORDS AND COMPOUNDING PATTERNS

2. Introduction

Many studies have been conducted in various languages on compound words. Some studies have looked at the structure of compound words while others have tried to establish the differences between compound words and other linguistic objects that have a similar structure like syntactic phrases. Other studies have also looked at different classificatory criteria to establish a universal definition of compound words (Fabb, 1998, Haspelmath, 2002; Bisetto and Scalise, 2005; Scalise and Bisetto, 2009; Lieber and Stekauer, 2009 and others). Despite all these studies, no universal definition, structure and classificatory criteria have been established about compound words. This chapter, therefore, addresses some of these issues by reviewing different sources of literature on the theory of compounding, and also introduce relevant properties that serve as background information on phrasal compounds analysed in Chapter 3 and 4.

The discussion progresses by first defining a compound word which leads to the analysis of different views by various scholars on what a compound word is. This is meant to establish a definition that best describes compound expressions in Cinyanja. The views of Lieber and Stekauer (2009), Bauer (2009), Haspelmath and Sims (2010), Carstairs-McCarthy (2002), Nakov (2013), Fiorentino (2006), Bauer (1983), Plag (2003), and Altakhaineh (2016a) are examined. Features from Plag's (2003) definition and Altakhaineh (2016a) which seem to cut across languages are discussed in detail. This leads to a discussion of different types of compound words. The part of the study dealing with compounding patterns common in English considers views from Bauer (1983), Fabb, (1998), Plag (2003), Bisetto and Scalise (2005), Scalise and Bisetto (2009), Lieber (2009c), Padrosa-Trias (2010), and others. Finally, a summary of the chapter is presented.

2.1 What is compounding?

Compounding is one of the branches of morphology which deals with word formation. Some scholars consider compounding the most productive process of word formation in most languages (Plag, 2003: 169; Booij, 2009:203; Kiefer, 2009:527; Altakhaineh, 2016b). According to Ralli (2010:57), the process of compounding involves combining words or stems to form new complex items. Fiorentino (2006:15) defines compounding as a process of word formation that involves the joining of two or more open class morphemes to form larger complex words. For example, the combination of *tea* + *cup* forms the word *teacup,* and *black* + *board* + *eraser* forms *blackboard eraser*. Therefore, the combined words, for example, *teacup* and *blackboard eraser* are called compound words. These words are so diverse that scholars define them differently. Some of the definitions are restrictive while others give a broader

presentation. In the next section, I discuss various definitions of compound words to establish which one is compatible with the type of “compounds” discussed in this thesis.

2.2 On the definition of a compound word

Several studies have been conducted on compound words in different languages (Marchand, 1960; Roeper and Siegel, 1978; Bauer, 1983, 2001, 2009; Fabb, 1984, 1998; Lieber, 1992, 2009a; Plag 2003; Harley, 2009; Padrosa-Trias, 2010, Altakhaineh, 2016a and others). Linguists have defined compound words from different perspectives to try and provide a definition which can be considered valid across languages. For example, Fabb, (1998:66) defines a compound word as a combination of two or more words. Similarly, Nakov (2013:3) states that a compound is a combination of two or more existing words (e.g. *newspaper*, *babysit*). Marchand's (1960:11) definition seems to refine both former and latter definitions by adding that a compound comprises two or more words which combine to form a morphological entity. However, these three definitions are not very reliable because they are based on the concept ‘word’, which itself needs to be defined. It is not clear what the term "word" involves in these definitions. Additionally, the definitions do not include compounds that have phrasal elements inside them (called "phrasal compounds"). For example, *over-the-fence gossip* and *God-is-dead theology* are also considered compound words, but include full phrases (a prepositional phrase and a sentence respectively), and are therefore phrasal compounds.

For Lieber (2009a:43), a compound word is “composed of two (or more) bases, roots, or stems”. This definition is restrictive because compound words that have inflected words are not considered. Carstairs-McCarthy (2002:68) on the other hand, explains that compound words are formed when roots combine with the much smaller category of phrasal words. In this definition it is not quite clear what is meant by “the much smaller category of phrasal words”. Although Carstairs-McCarthy has included an important aspect of phrases in the definition, the definition does not cover compounds that have inflections on the first elements (for example, *suggestions box* and *students commissioner*) and synthetic compounds, (i.e. compounds with a deverbal head, for example*, bus-driver*). These derived words are not roots, yet they are compound constituents. Additionally, it is not clear whether the notion of “root” also includes bound roots. The notion “root” is critical and needs a clear explanation because, in some languages like Cinyanja, although some roots may carry the meaning of the verb, they may not be able to stand alone unless affixes are attached. For example, even if the root *-dy-* carries the meaning of the verb, it is a verb when the affixes *ku-* and *-a* are attached to realise *ku-dy-a,* ‘to eat’. Another problem with Carstairs-McCarthy’s definition is that some words that have complex structures

such as recursive compounds are not included in the definition. These are compound words formed by a repeated stacking of new words to form a new compound as shown in the example below.

1	a.	[[blackboard] eraser]	→blackboard eraser
	b.	[[[water resource] research] centre]	→water resource research centre

Other definitions that are worth considering are from Haspelmath and Sims (2010:137) and Bauer (1983:29). Haspelmath and Sims define a compound as a complex lexeme that is made up of two or more base lexemes. Bauer's (1983) definition reinforces the idea by stating that a compound may be a lexeme containing two or more potential stems that have not subsequently been subjected to derivational processes. In these definitions, the key term that is used is lexeme, which Bauer (2009:343) explains as referring to all forms of the same word, (e.g. the lexeme *come* has the forms *come*, *comes*, *came* and, *coming*). Assuming that Haspelmath and Sims use the word lexeme in the sense of Bauer (1983), then their definition, unlike Bauer's (2009), is quite limiting because it restricts the use of lexeme to only a basic unit of a word form. This implies that compounds that have inflections, for example*, parks commissioner and programs coordinator*, are not considered compounds. Much as these definitions seem to apply to a wide range of compounds, they do not cover compounds that include phrasal elements (i.e. phrasal compounds).

Another definition which seems to cut across languages is from Plag (2003). Plag (2003: 135) defines a compound word as "a word that consists of two elements, the first of which is either a root, a word or a phrase, the second of which is either a root or a word." Plag's definition is comprehensive and it accommodates compounds from various languages. Plag explains each feature in the definition explicitly. For instance, the term 'element' has been used as a generic term to mean either a root, word or phrase.

According to Plag, the use of the term 'two elements' in the definition should not be interpreted literally because it covers even complex words such as *award committee member* and *bathroom towel designer.* For example, *bathroom towel designer* involves the compound *bathroom* which consists of two elements, that is, *bath* and *room* and *bathroom towel* also has two elements namely *bathroom* and *towel* and *bathroom towel designer* comprises two elements too, and these are *bathroom towel* and *designer*. Plag argues that though these words are complex, they can still be analysed in a binary manner. Another critical aspect of the definition is the inclusion of the term 'phrase' which most definitions fail to capture.

However, although Plag's definition looks elegant, it is still too restrictive, because the definition allows phrases inside compounds only when the phrases are the first elements. This suggests that languages with left headed compounds that have phrasal elements as second members of the compound such as the ones discussed in Chapter 3 of this thesis are not covered by the definition. Due to this observation, I now turn to Altakhaineh's (2016a) definition of compound words. Altakhaineh (2016a:40) defines a compound word as a 'complex word that consists of at least two adjacent elements, where the non-head is normally non-referential. Each of these elements is either a word, combining form or a phrase, so that the whole compound is a combination of these elements.' Altakhaineh's definition seems to cover most of the important features of compound words in various languages. Importantly, the non-restrictive position of phrasal elements which seems to characterise Cinyanja phrasal compounds analysed in this study, has been included in the definition. Other interesting features that emerge from Altakhaineh's definition of compound words which are critical to the understanding and eventual analysis of compound words in general are; adjacency, non-referentiality, word and combining forms. These features are reviewed in the subsequent sections together with other critical features of compound words such as root, headedness and affixation.

2.3 Features of compound words

2.3.1 Adjacency

The concept of adjacency as used by Altakhaineh (2016a, 2016b) relates to what Lieber and Stekauer (2009) call insertion/separability. Altakhaineh (2016b) defines adjacency as the impossibility of inserting a word between two words in a compound word. While it is possible to insert a word inside a phrase, for example, *a black ugly bird*, it is not possible with a compound word as shown below.

2. a. blackbird
   b. *black-ugly-bird
   c. ugly blackbird

As shown in (2c) above, in compound words, the adjective modifies the entire compound word. Altakhaineh (2016b) indicates that adjacency is a reliable criterion for distinguishing compounds from phrasal constructs in Arabic because an element such as a demonstrative *haaða* this' and quantifier *baʕdˤ* 'some' cannot be inserted between elements of the compounds. Additionally, Altakhaineh notes that another criterion that is similar to insertion is displacement and this relates to the impossibility of displacing a constituent part in a compound word:

3. a. Trucks are what he [drives_____]

b. *Truck is what he likes a [_____ driver]

(Altakhaineh, 2016a:21)

Altakhaine indicates that the gaps are the original positions of the displaced words *truck* and *trucks*. While (3a) is grammatical, (3b) is ungrammatical because a constituent part of the compound is displaced. According to Altakhaineh, the two criteria namely insertion and displacement are related because they both deal with the inseparability of the compound elements. Therefore, these can be categorised under one criterion called adjacency.

2.3.2 Referentiality

Referentiality or anaphoricity is one of the criteria for testing compoundhood. Altakhaineh (2016b) refers to referentiality as the relationship between language and how it connects to the world (see also Saeed, 2003). The referring expression generally includes pronominals *it, he, she, they, one* and others (see also Bresnan and Mchombo, 1995:201). In compound words, the non-head is considered non-referential but generic, that is, it does not refer to any specific entity or person. For example, in a compound word *truck-driver*, 'truck' is the non-head; it does not refer to a specific truck, and as a result, is considered non-referential. In English, the non-head is usually the first element in a compound word while the head is the second element. However, Bauer et al. (2013: 464) argue that in certain compound words the first element of the compound may refer to a specific entity. They cite the noun *army* in the compound word *army budget* as having a specific reference to the army of a certain country (see also Altakhaineh, 2016b: 12). Despite this observation by Bauer, Altakhaineh considers referentiality as one of the reliable tests for compoundhood in Arabic and Hebrew because in these languages, the non-head of phrases is always referential while that of compounds is always non-referential.

Other important features that form part of Altakhaineh's definition of compounding are the notions 'word' and 'phrase', which I discuss next.

2.3.3 Word and phrase

The term "word" has for decades been a difficult concept to define. Di Sciullo and Williams (1987) assert that in morphology the term "word" is used to refer to meaningful linguistic units that are

unanalysable called morphemes; and in syntax, a “word” is defined as a syntactic atom or a unit which cannot be analysed. Lieber (2009a:12) asserts that in morphology a word is a sequence of one or more morphemes that can stand alone, whereas a morpheme is defined as the smallest meaningful unit of a language. For Plag (2003:9), a word is defined as an entity that has a part-of-speech specification (e.g. an article, verb, noun, adverb, adjective, preposition etc.), and that is a syntactic atom, which usually has one main stress and is an indivisible unit; that is, it cannot allow any insertion of material inside it. This is a detailed attempt to define a word by using different properties and features that describe a word in various languages. In the definition of compound words, Plag explains that the term "word" also represents grammatical words such as functional words (e.g. *the, some*, *a*), bare roots (e.g. *cat, house, cut*), and inflected words (e.g. *parks commissioner*, *teeth marks*). It seems Plag’s use of the term "word" is valid because it incorporates inflected words which are also a major feature of most compound words in Bantu languages as will be demonstrated in Chapter 3 with reference to Cinyanja VN-compounds.

Another important attribute that Altakhaineh has added to the definition of a compound word is the term *phrase* which most linguists have failed to include in their definitions of possible compounds. In recent literature, phrasal compounds have been a topic of discussion on word formation (Salvin, 1984; Wiese, 1996; Lieber, 2009c; Harley, 2009; Pafel, 2017 and others). These compounds are significant because they have a unique structure where the first element is a phrase in English as shown in subsection 2.4.5. For example, in *sexually transmitted disease clinic*, the noun phrase *sexually transmitted disease* combines with the noun *clinic*, which is the head, to form a compound word. Plag (2003:7) defines a *phrase* as a group of words that expresses a unified concept, and Bauer (1983:207) adds that the unified constituents should have a head. The head, also called the nucleus, is the element that determines the syntactic category of the phrase. In a verb phrase, the head is a verb, in a noun phrase, it is the noun, etc. According to Harley (2009), the phrasal elements in compounds behave like phrasal units, and not like complex compounds, because if the latter was the case, the phrasal compound *over-the-fence gossip*, would be realised as **fence-the-over gossip* because of the Right-Hand-Head Rule (RHR) which is an unacceptable expression. The inclusion of the term "phrase" in the definitions is significant because it accommodates complex expressions such as phrasal compounds discussed in this thesis. Further, the understanding of the concept of phrasal compound will also assist in establishing if the complex expressions that are compound-like in some Bantu languages (e.g. in Cinyanja), are also compounds and phrases or just lexicalised phrases, such as *kick the bucket*. Other critical features in Altakhaineh’s definition of a compound word are root and

combined forms (see Plag, 2003; Altakhaineh, 2016a) which are discussed in the subsection that follows.

2.3.4 (Bound) Roots and combining forms

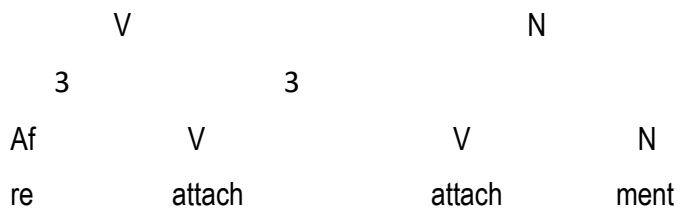
A root is the core part of a word which carries the meaning and cannot be segmented further. Plag (2003:74) describes a root as a "central meaningful part of a word to which affixes are attached." An affix is a bound morpheme which can be attached before a base (*prefix*) or after a base (*suffix*). However, some words are made up of elements that look like affixes on the one hand, and roots, on the other hand, called bound roots. Both affixes and bound roots cannot occur in isolation, unless they are attached to words. Plag (2003) argues that although bound roots resemble affixes, they behave differently because a bound root can combine with an affix (e.g. *bapt-ism, prob-able*) to create a new word while an affix cannot combine with another affix *(*e.g. *ism-able, re-ism)* to form a new word. Similarly, two bound roots cannot combine (e.g. **bapt-prob*) to form a new word.

Some bound roots are of Latin or Greek origin but have infiltrated the English language. These bound roots are referred to as combining forms (Altakhaineh, 20016a). These forms are also called neo-classical elements (Bauer, 1983:216; Plag, 2003; Lieber, 2009a:46; Mukai, 2006). For example, *-logy* in *bio-**logy*,** or geo- in ***geo**-graphy* are called neoclassical elements, and the resulting compound words are called neo-classical compounds. These combining forms resemble affixes because they can attach either to the initial part of the lexeme (for example ***photo**electric*) or to the final part of the lexeme, for example, *music**ology*.** It is also typical of the combining forms to combine with bound roots (e.g. *bio-logy*, *micro-scopy, sciento-logy*), and with words (e.g. *biochemistry*, *audiobook, hydro-electricity)* to form new words (cf Carstairs-McCarthy, 2002; Mukai, 2006:25; Conti, 2006:16). However, a unique feature of combining forms is that they can combine with other combining forms to form a new form, for example, *hydro* + *logy* forms *hydrology* while *morpho* + *logy* forms *morphology* in English. Although English has many combining forms, I will not discuss these any further because they are absent in most Bantu languages, Cinyanja in particular. What is crucial in discussing Cinyanja complex expressions is the term '(bound) root', because this property is present in Cinyanja compounds**.** Another critical feature of compound words is 'headedness' - discussed in the next subsection.

2.3.5 Headedness

Headedness is one of the features that distinguishes structures like phrases from compound words. In a syntactic structure, the head can be the first element that gives the entire structure (e.g. phrase) its grammatical category and semantic interpretation and may also be responsible for selecting the type of complement that should follow the head. For example, the adjective *jealous* may select a prepositional phrase such as *jealous of her friend* while the verb *buy* may select the noun *a book* as its complement (Centnarowska, 2016). However, this may be different from a morphologically complex word where the head is the right-most element of a word. This relates to what Williams (1981:248) proposes as the Right-Hand-Head Rule (RHR). As stated above, in a morphologically complex word, the head is the rightmost element of that word, and it can be a word or an affix:

4. a. reattach b. attachment



In considering example (4), headedness in morphology is not based on the relationship between constituent members but the position of the constituent member. For example, in (4a) the head is the rightmost member *attach* while in (4b) it is the rightmost suffix *–ment*. In suffixed words, the suffix determines the category of the word to which it has been attached. Prefixes, nevertheless, hardly determine the category of the word. For instance, in the word *re-attach*, the head is *attach* and it determines the category of the word as a verb while in the word *attachment*, the suffix *-ment* determines the category of the word as a noun. Prefixes, therefore, cannot be assigned a lexical category because they do not take a head position. Nonetheless, there is an exception to this rule with regards to the prefix *en-* in a word such as *enlarge* (see Williams, 1981: 249). The attachment of the prefix changes the adjective *large* to a verb. Affixes such as the prefix *en-* that create a new lexeme from a base are called derivational affixes while those that provide grammatical information of number, for example, the plural suffix *–s* and the past tense suffix *–ed* are called inflectional affixes. Affixes will be discussed more in subsection 2.3.6.

The RHR is not only applicable to derivatives but also compound words. In compound words, the head is the constituent that determines the semantic and lexical category of the compound word while the non-head can be a part of speech which modifies the head and is usually non-referential. For example, in the compound word *blue-sky, blue* is the adjective and the non-head of the compound whereas *sky*

is the noun and head of the compound. As the head of the compound, the noun *sky* shares its features with the entire compound, thus, determining the semantic and lexical category of the compound. The compound, therefore, is a hyponym of the grammatical head (Bauer, 1983:30, Guevara and Scalise, 2009). With regards to the non-head, the adjective *blue* modifies the noun as a kind of sky. In general, English compounds are right-headed with the rightmost element being the head. Morphologically complex words that have heads are called endocentric compounds.

As regards inflectional affixes, they are normally attached to the rightmost part of the entire compound word in English. This has lead scholars such as Di Sciullo and Williams (1987) to assert that inflectional affixes are added to the head of the compound word; for instance, *schoolbooks*. Nevertheless, there is an exception to this assumption because some compound words in English have inflectional affixes attached to the first member of the compound whose head is still the right-hand member. Examples of such exceptions as classified by Plag (2003), Lieber and Stekauer (2009), Ziering and van der Plas (2014: 1049) are shown below.

5. suggestions box
   parks commissioner
   students commissioner
   programs coordinator
   arms conscious
   sales-oriented
   sportsman
   systems analyst
   teeth marks
   overseas investor

These examples demonstrate that it is not always the head that bears inflectional elements as earlier indicated, but that inflection can also be realised on some non-heads in certain compounds. Anderson (1992:295) makes profound observations about the presence of inflections in some exocentric compound words. These compounds are generally not headed but as Anderson observes, some take regular inflections on the rightmost word; for example, *sabretooths* (a variety of tigers), *tenderfoots* (inexperienced beginners). This shows that inflection is not an outright mark of headeness**.** Anderson also indicates that in languages that have more overt inflectional features such as Icelandic, inflectional features may percolate from both heads and non-heads especially in adjective + noun (A-N)

compounds. This shows it is not absolute that the head takes inflections and that grammatical features of compounds percolate only from the head element of the compound.

Therefore, the RHR is not universal but is more applicable to English compounds. However, in other languages such as Italian, compounds are left headed (Booij, 2009:211) while in Catalan, compounds are both left and right headed (Padrosa-Trias, 2010).

2.3.6 Affixation

An affix is a bound morpheme that is attached to a root or stem. There are three types of affixes, and they are classified according to their position. Those that are added to the beginning of a root or stem are called prefixes while affixes added to the end of a root/stem are suffixes. Since these affixes are bound morphemes, they can maintain or change the syntactic category of a word as illustrated below.

6. Un-touch-able-s
   ↓ | ↓ |
   Prefix | suffix |
   ↓ (deriv.) ↓
   root suffix (inflect.)

In example (6) above the suffix *–able* is a derivational affix because it changes the category of the word *touch* while the prefix *un-* and the suffix *–s* are inflectional morphemes because they maintain the syntactic category of the word *untouchables*. Another type of affix, which is inserted right into the root of the word, is called an infix, for instance, the word *bloody* in the word *kangabloodyroo* (Bauer, 1983:90*).* It is important to note, however, that infixes are rare in English as well as Bantu languages. Therefore, they do not add much information to the current study.

The lexemes that form English compound words are mainly free morphemes or bare roots which have a distinct meaning of their own. For instance, in *schoolbook*, the elements are both free morphemes. In the case of a compound word having inflectional or derivational affixes, these often occur after the second lexeme or word as illustrated in example (7a & b).

7. a. flowerpot(-s)
   b. bus-drive(-er)

In case of a compound word having both derivational and inflectional affixes, two theories are distinguished. The first is that a derivational affix is merged with one of the constituents of the compound first, and then the inflectional affix is added to the entire compound word as shown in example (8b). The second view is that the derivational affix is merged with the compound combination, and then the inflectional affix is added after the derivational affix as shown in (8c). This is the case with most synthetic compounds which seem to be simultaneously a case of compounding and derivation.

8. a. bus drivers
   b. [bus [driv(e)-er]-s]]
   c. [[[bus drive]-er]-s]

However, it is not entirely the case that affixes are always attached to the rightmost member of the compound as suggested by examples (7) and (8) above. There are instances where suffixes are attached to the first word of the compound such as *mothers-in-law* and *parks commissioner.* Therefore, the claim that affixes are added to the second member of the compound which is the head (see Williams, 1981 RHR) is not entirely correct (see example 5 above). Additionally, some languages have inflections on both compound elements, for example, *sacs-poubelles* 'bin bags', which are also compounds in French (Rosenberg, 2007:350). Such examples show that the notion of affixation in compound words should not entirely be associated with headedness, but also with the nature of the language being studied.

In sum, I have shown in the above section some disparities in the definition of compound words and also discussed core features of compound words. Other important features of compounds that were reviewed in subsection 2.3.5 and subsection 2.3.6 are headedness and affixation respectively. The next section discusses the different types of compounds found in English. The presentation of these compounds will facilitate the analysis of the structure of Cinyanja compound words in the later chapters.

2.4 Types of compounds

In this section, I discuss different types of compound words, namely endocentric compounds, exocentric compounds, root compounds, synthetic compounds, and phrasal compounds. In the analysis of these types (especially of the endocentric and exocentric compounds), I refer to Bisetto and Scalise's (2005) subcategorisation of compounds as subordinative (SUB), attributive (ART) and coordinative (CRD)**.** Reference is also made to Padrosa-Trias (2010), Lieber, (2009b, 2009c) Scalise and Bisetto (2009), Bauer (1983), and others to discuss the types of compound words.

2.4.1 Endocentric Compounds

In endocentric compounds, the head and the entire compound have the same referent (Bauer, 1983; Lieber, 2009a; Buenafuetes, 2014)**.** In English, endocentric compounds are generally right headed. This is illustrated in example (1) above: a *blackboard eraser* is 'a kind of eraser'. This means that *eraser* is the head and *blackboard* is the non-head of the compound. The interpretation of the entire compound as *blackboard eraser* is dependent on the constituent head *eraser*. As indicated above, English compounds have their heads on the right-hand side (i.e. RHR), and the non-heads on the left (see Williams, 1981; Di Sciullo and William, 1987; Plag, 2003). In endocentric compounds, non-heads may be nouns, adjectives, prepositions, or phrases as in the following examples.

9.	*wind*mill	noun + noun
	*black*board	adjective + noun
	*under*dog	preposition + noun
	over-the-fence gossip	phrase + noun

According to Bisetto and Scalise (2005; henceforth B & S), endocentric compounds can be divided into three different types namely; subordinative (SUB), attributive (ART) and coordinative (CRD) compounds (see also Lieber, 2009a: 48). These types are based on grammatical relations that exist between the elements that form the compound words. In considering endocentric SUB compounds, B & S (2005) note that the two constituents in these compounds display a complement relation. In this type of endocentric compound, the first member is a complement/argument of the head. For example in *truck-driver*, *truck* is the complement of the deverbal head *driver* and can be interpreted as a *driver of a truck*. What is significant is that, typically, the two elements in the endocentric SUB compounds share a grammatical relation of complementation and most of them are nominal compounds, as shown in Table 2.1 below.

Table 2.1: Possessor-head relation of nominal compounds

Compound word		possessor relation	Type of compound
love story	NN	story of love	nominal compound
steamboat	NN	boat of steam	nominal compound
water pipe	NN	pipe of water	nominal compound
apron string	NN	string of/in/on/ an apron	nominal compound
babycare	NN	care for babies	nominal compound
chicken curry	NN	curry for chicken	nominal compound

Further, the grammatical relation can be that of an argument relation as shown in the examples below.

10. shoe repairer — repairer of shoes
bookseller — seller of books
coffee maker — maker of coffee
dishwasher — washer of dishes
truck-driver — driver of trucks
crime prevention — prevention of crime
wish fulfilment — fulfilment of a wish

[Padrosa-Trias, 2010:135]

From the examples in (10), what is clear about these compounds is that the second noun is the head of each compound and the non-head is an argument of the deverbal noun. For example; *shoe repairer* is the 'repairer of shoes' and the *bookseller* is the 'seller of books'. The only significant discrepancy between the compounds in Table 2.1 and those in example (10) is that the second noun in (10) is derived from a verb by adding a derivational suffix to the verb. The non-head is therefore the internal argument of the derived head which shows that the compounds are synthetic compounds.

From the examples in Table 2.1 and example (10), we notice that most of these endocentric compounds are generally right-headed and NN compounds are the majority. The non-head of the NN compounds is usually in an 'of relation'. However, some compounds, for example, *apron-string* may have other interpretations such as 'string of an apron', 'string on an apron' or even 'a string in an apron' (B & S, 2005:326).

Another endocentric type of compounds is the attributive compound (ART). ARTs are usually formed by combining either an adjective and a noun; for example, *blackboard, blackbird, blue-cheese* (i.e. the

adjective signifies a property of the head noun), or by combining two nouns where the non-head may be used metaphorically to express some kind of attribute to the head (B & S, 2005; Guevara and Scalise, 2009). These compounds are usually right-headed in English. It is worth mentioning that most endocentric ART compounds are nominal compounds in English, as shown in table 2.2 below.

Table 2.2: Endocentric ART Compounds

Compound word	**Semantic interpretation**	**Attribute**
snail mail (N + N)	a kind of mail that takes long to deliver	*slowness*, snail attribute
swordfish (N+ N)	a kind of fish that has a sword-like snout	*shape*, sword attribute
ghostwriter (N + N)	An author who is not the actual writer	*invisible*, ghost attribute

Although the compounds refer to a *mail*, *fish* or *writer*, they are endocentric because they express a modifier-head relationship.

Another type of compound, namely the endocentric coordinative (CRD) compounds, also referred to as dvandvas or copulatives (Bauer, 2009:351; Plag, 2003), consists of constituents whose meaning can be described as a conjunction of the meaning of its parts. The two elements that are linked in these compounds are both considered semantic heads and are rarely recursive in English compounds. In coordinative compounds, there is no modifier-head relationship or complement-head relationship between the compound's elements because, semantically, these compounds have two heads (B & S, 2005:32; Scalise and Bisetto, 2009:46) as illustrated in (11).

11. bitter-sweet — bitter and sweet
    fighter-bomber — fighter and bomber
    leaner-driver — learner and driver
    poet-doctor — poet and doctor
    woman-doctor — woman and doctor

These compounds have a unique composition because the non-head is absent. Contrastively, most compounds are described based on the presence or absence of the head, not the non-head**.** It is usual, syntactically and semantically, to interpret the non-head as a kind of attribute or complement to a given

head. However, in coordinative compounds, the constituents are on equal footing and they are both heads (Plag, 2003; Padrosa-Trias 2009). In the case of these compounds inflecting for plural, both elements of the compound word would take plural inflection, particularly if the compounds are NN compounds (Guevara and Scalise, 2009:124), for example, *poets-doctors* interpreted as 'poets and doctors'. Such examples also weaken the morphological analysis of headedness where only the right-hand member of the compound is affixed with plural inflection. Therefore, classifying coordinative compounds as endocentric is questionable because both elements can take the plural inflection and can head the compound word. I assume that this is one of the reasons why Padrosa-Trias (2010) does not consider coordinative compounds as genuine compounds, but that they are instances of asyndetic syntactic coordination (i.e. omission of conjunctions where several are required). Padrosa-Trias argues that an example like, *she was a producer and director and actor and author* can be presented as *producer-director* and *actor-author* which is a case of ellipsis, not compounding.

2.4.2 Exocentric compounds

Exocentric compounds normally lack a semantic head. In most cases, they have a referent outside the compound constituents. None of the compound constituents refers to the entire compound (Fabb, 1998:67; Scalise, et al., 2009; Basciano, et al., 2011). For example, *pickpocket* is an exocentric compound because it does not refer to a kind of pocket but to 'someone who picks things from peoples' pockets'. Nakov (2013) asserts that exocentric compounds that denote human beings have pejorative properties. This is illustrated by the examples below.

12.	redhead	a person who has a red head (i.e. due to red hair)
	flatfoot	'policeman'
	birdbrain	'someone with a head no bigger than that of a bird'
	faintheart	'someone with a faint heart'
	redcap	**'**a person who is associated with a red cap'

Bauer (2008) notes that compounds with a possessive reading like those in (12) are also referred to as 'possessives' or 'bahuvrihi' compounds.

Like endocentric compounds, exocentric compounds can also be interpreted as attributive, coordinative and subordinative (see B & S, 2005; 2009; Lieber, 2009c). Exocentric ART compounds display some kind of relation between constituents in the compound despite the head referring to

something outside the compound constituents. For example, in *greybeard,* there is an attributive relation between *grey* and *beard* because *grey* designates a particular property of beards, even though a *greybeard* is not a hyponym of any of the compound constituents. The actual interpretation of *greybeard* is 'a person with a grey beard'. According to Lieber (2009b:99), exocentric ART compounds operate in the same way as endocentric ART compounds except that the head in exocentric ART compounds represents a person or an entity in the world. This relates to what Booij (1992) refers to as a metonymic relation, where the semantic interpretation of the compound word is related to the behaviour or characteristics of the person, object or item being referred to. For example, *birdbrain* refers to a person who acts foolishly as though they had the brain of a bird. The semantic interpretation of the examples in (12) above fits into this description.

Another type of exocentric compound are coordinate compounds. In exocentric CRD compounds, like in endocentric CRD compounds, the compound constituents are semantically perceived to be conjoined (see Padrosa-Trias, 2009, 2010; B & S, 2005:327). Semantically, CRD compounds have two heads and none of the constituents is subordinate to the other. The constituents in the compound are both of equal weight. For example, *mind-brain* is interpreted as 'mind and brain'. Therefore, there is no modifier-head relation between these two constituents. Another example of an exocentric CRD compound is the NN compound *mother-child*. The interpretation that can be drawn from this example is that each head signifies a particular relation; that of 'a mother' in one instance, and 'a child' in another. Another observation that can be drawn from this type of exocentric CRD compound example relates to the relation of subordination between the elements that is, *child* is subordinate to *mother*

The third type, exocentric SUB compounds, refers to an entity in the world. Mostly, these compounds refer to an entity outside the compound constituents and may denote people, animals, plants or objects as shown in table 2.3.

Table 2.3: Compounds and denotatum

Compound	**Denotatum**
pickpocket	person

killjoy	person
spendthrift	person
wagtail	animal
tumble-dung	animal
heal-all	plant
catch-fly	plant
tear-thumb	plant
rotgut	object
scarecrow	object

According to Lieber (2009b:94), exocentric SUB compounds are poorly attested especially in English as opposed to Romance languages. In these compounds, the first element is a verb and the second a noun in English. The relation between the constituents is that of a verb-argument or verb-complement relationship. This is illustrated in the following examples;

13.	pickpocket	verb +noun	(VN)
	catchfly	verb + noun	(VN)
	cutpurse	verb +noun	(VN)
	killjoy	verb +noun	(VN)

The exocentric SUB compounds do not have a head among the compound constituents. The referent is external to the compound as a whole. Finally, it is noteworthy that subordinative exocentric compounds have a verb-argument relationship just like synthetic compounds. This also applies to some Bantu languages as will be attested in the discussion of some Cinyanja compounds in Chapter 3.

2.4.3 Synthetic compounds

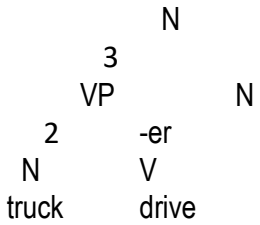
Synthetic compounds are always based on deverbal heads. The semantic relationship between the head and the non-head seems to be flexible; the non-head is typically an internal argument of the verb (Roeper and Siegel, 1978; Selkirk, 1982; Harley, 2009; Lieber, 2009a). Synthetic compounds are sometimes called verbal compounds (Fabb, 1998; Lieber 1983:265; Olsen, 2017) because of the verbal element such as *maker* in *shoemaker* and *driver* in *truck-driver*. The nouns *shoe* and *truck* are considered internal arguments of the verb. In most cases, the deverbal element may either be an *–er*

nominal*, an –ing* noun or adjective or the *–ed* passive participle. The nominalising affixes (i.e. category-changing affixes) can attach to verbs in compounds; for instance, *dog kill***er** (*-er*), *nut eat**ing** giraffe* (*-ing*) and *monkey bitt**en** tourist* (-en) (see Roeper and Siegel, 1978; Lieber, 2009c:358). Two main theories have dominated the literature on synthetic compound words; the syntactic and the morphological theory.

2.4.3.1 Syntactic theory of synthetic compounds.

The syntactic theory assumes the syntactic approach to synthetic compound formation which claims that synthetic compounds are formed when parts that are syntactic atoms combine (Fabb, 1984; Sproat, 1985; Booij, 2009; Harley, 2009; Padrosa, 2010). In this theory, two lexical items combine before the category changing affix is attached. Ackema and Neeleman (2001: 11) note that in the syntactic representation of synthetic compounds, the attachment of the category-changing affix happens after the lexical items have combined. For example, in *truck-driver*, the N *truck* and the V *drive* combine to form a complex V, which is later merged with the nominalising suffix *-er* to form the compound noun *truck-driver* as illustrated in (14).

14. truck-driver



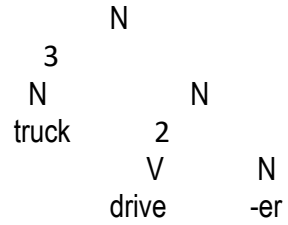
In the structure in (14), semantically, *truck* is interpreted as the internal argument of the verb *drive* and *driver* is considered the head. It has often been noted as a problem for this analysis that there are no verbal compounds such as “to truck-drive” in English. The existence of this verb is implied by the analysis of (14) above. Ackema and Neeleman (2004) note that the morphological analysis is an alternative theory of generating structures as discussed in the next subsection.

2.4.3.2 Morphological theory of synthetic compounds.

The morphological theory of compound words adopts the lexicalist approach where synthetic compound words are assumed to be formed by two simplex expressions from the lexicon (see Roeper

and Siegel, 1978; William, 1981; Sproat, 1985; Bresnan and Mchombo, 1995; Scalise et al, 2009; Lieber, 2009c). Depending on the semantic interpretation of the compound word *truck-driver*, Ackema and Neeleman (2004) note that the morphological representation of the internal structure of a synthetic compound would be analysed based on the structure in (15). In this structure the nominalising affix -*er* merges first with the V *drive* to realise an N *driver* which later merges with the N *truck* to form the compound *truck-driver* (Lieber, 2009c; Padrosa-Trias, 2010; Ackema and Neeleman, 2004)**.** For such a representation, *truck* is construed as a modifier of the noun *driver* as illustrated in (15).

15. truck-driver



In (15) above, the nominalising affix –*er,* is suffixed to the headword *drive*- not to the entire compound as in the structure in (14). Padrosa-Trias (2010: 26) notes that the structure in (15) is only possible when *truck* is a modifier, and not the internal argument, of *drive*.

Different scholars have different views about which of the structures in (14 & 15) presented above to adopt in analysing synthetic compounds. For instance, Fabb (1984:137) supports the structure in (14) as the correct representation of synthetic compounds based on the argument that synthetic compounds are a good example of words whose internal structure is syntactic, fully productive and semantically transparent. Fabb also indicates that these compounds have an internal composition which mirrors the internal argument structure of the verb phrase in English. However, Haspelmath and Sims (2010) favour (15) as the correct representation of the structure of synthetic compounds. This view is based on the argument that, generally, derivational affixes are attached to the semantic and/or formal head not to the entire compound as depicted in example (14).

Therefore, to determine which structure of synthetic compounds to adopt, it is important to consider the semantic interpretation of compounds, i.e. whether the non-head is an argument, in which case the structure in (14) is correct, or a modifier, in which case (15) is preferable. However, as it will be noticed from the analyses of Cinyanja VN-compounds discussed in Chapter 3, I consider the structure in (14) preferable. The explanation for this preference will be provided in Chapter 3. I propose that the

analysis of compound words where the verb merges with the noun before the affix is added suits the analysis of verb-noun (VN) compounds in Cinyanja, also discussed in Chapter 3.

2.4.4 Root compounds

Root compounds can be distinguished from synthetic compounds because, in synthetic compounds, there is always derivational morphology while in root compounds, the constituents are free morphemes. Mainly, the compound elements are either nouns, adjectives, or verbs. De Belder (2013) asserts that root primary compounds are formed by a direct merge of two roots. In these compounds, the head determines the category of the compound. The head which is usually the second element in English compounds is not derived from a verb (Lieber, 2009a:46; Haspelmath and Sims, 2010) as in synthetic compounds. The relationship between the head and the non-head is not restricted in root compounds. What is clear about these compounds is that the relationship between the elements in the compound is not that between the verb and its argument but that of modification (Harley, 2009; Di Sciullo, 2009). For De Belder (2013), root compounds consist of bare roots which have merged directly with one another as shown below.

16.	a.	doghouse	[N + N]
		flowerpot	[N + N]
	b.	blackbird	[A + N]
		blue-sky	[A + N]
	c	runway	[V + N]
		bathtowel	[V + N]
	d.	green-blue	[A + A]
		ice-cold	[A + A]

From the examples in (16) above, it is clear that this type of compound words is very productive and that it involves various types of word classes, even though in English**,** root compounding mostly derives nominal compounds. For root compounds, the semantic interpretation is quite liberal. For example, a *doghouse* can be a house for dogs, a house in which dogs are bred or a house with dogs painted on it.

2.4.5 Phrasal compounds

Another type of compound whose properties are important for this study is the phrasal compound. These compounds are formed by joining a phrase and a noun. Usually, the phrase is the first element and the noun is the second constituent of these compounds in English as shown in the examples below.

17. a. stuff-blowing-up effects
bikini-girls-in-trouble genre
comic-book-and-science-fiction fans
slept-all-day look
(Lieber, 2009a:153; see also Lieber, 1992:11)

b. floor-of-a-birdcage taste
over-the-fence gossip
God-is-dead theology
Sex-in-shiny-packets literature
(Pafel, 2016:12)

Lieber (2009a) notes that like with ordinary compounds, the constituents in phrasal compounds cannot be separated or modified despite being productively created (see also Scalise and Guevara, 2005; Lieber, 1992). As Lieber (2009a:153) and Harley (2009:142) indicate, modification in phrasal compounds affects the entire compound word, not a constituent part, as illustrated in (18).

18. a. *stuff-blowing-up exciting effects
b. exciting stuff-blowing-up effects
(Lieber, 2009a:153)

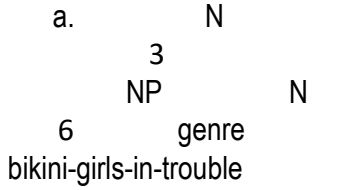
However, it is possible to insert a word in between elements of an ordinary phrase. For instance, to the phrase *small girl,* we can insert the word *pretty* to have *a small pretty girl.* Another criterion that would distinguish phrasal compounds from independent syntactic phrases is the modification of the first element with a degree adverb such as *very* (Lieber, 2009b). For example, the phrase *a very pretty girl* is permitted while **very pretty bikini-girls in-trouble genre* is not permitted because the phrase is part of the compound word.

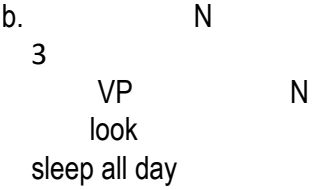
The examples in (18) above show that there is a relation between words and phrases in phrasal compounds. The phrases are built based on words and words based on phrases whose output is a word. What is interesting is that the phrasal part in these phrasal compounds acts like abound

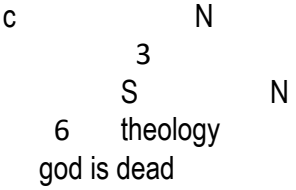
morpheme because it needs a host, that is, a noun that it combines with (see Pafel, 2016). Therefore, we can say that phrasal compounds straddle the borderline between morphology and syntax because even though they form part of a word, they have a phrasal structure which is the core feature in syntax as shown in the examples below.

19. a. god-is-dead theology.
    b. I-told-you-so attitude

In (19a) *god is dead* and in (19b) *I told you so* are complete sentences with a subject and predicate. Despite the sentence structure, Salvin (1984:56) notes that phrasal compounds with a phrase and those with a sentence are analysed in the same way because they both adjoin to the head noun as below.

20. a.


b.


c


The next section discusses some types of compounding patterns mainly in English although reference is made to other languages in the course of the discussion.

2.5 Types of compounding patterns in English

There are different types of compounding patterns in English. In this chapter, the classification has been done based on the resulting category. The nominal compounding pattern describes compounds consisting of two nouns (NN), an adjective and a noun (AN), or a verb and a noun (VN). The adjectival compounding pattern is found in compounds which may consist of a noun and an adjective (NA), two adjectives (AA), or preposition and an adjective (PA). Then the last but not least of the patterns is a verbal compounding pattern. In this pattern, the compound constituents may comprise a noun and a

verb (NV), two verbs (VV) or an adjective and a verb (AV). In English, the most common and most productive of all types is the nominal compounding pattern forming NN compounds**.**

In this section, I analyse each of the three types of compounding patterns starting with nominal compounding followed by adjectival compounding and finally the verbal compounding. The sources that I refer to are Bauer (1983), Plag (2003), B & S (2005), Conti (2006), Padrosa-Trias (2009) and Lieber (2009a, 2009b, 2009c). Other sources are referred to as well.

2.5.1 Nominal compounding patterns

The nominal compounding pattern is the largest class of compounds. Generally, these compounds are the most productive and can be described based on the presence or absence of the head. As noted above, in headed compounds, the first constituent is typically the non-head and the second is the head in English. The head is a noun and the non-head can be a noun (N), verb (V), adjective (A), adverb (ADV), preposition (P) or phrase (Plag, 2003, Carstairs-Mc Carthy, 2002). Nominal compounds can also comprise more than two members and are said to be recursive (Kiefer, 2009). Since there are many nominal compounding patterns, I discuss four different types which I consider critical to the discussion of the structure of Cinyanja compound words. These compound types are, NN, AN, VN and PN nominal compounding patterns.

2.5.1.1 NN Nominal Compound Pattern

As mentioned above, the NN nominal compound is considered the most common compounding pattern in English. These compounds may be formed by combining two nouns (NN), and the majority are endocentric. According to Bauer (1983:203), the best way to classify these compounds is to consider the semantic relationships and morpho-syntactic criteria. Bauer also notes that the NN category can also be exocentric although this type of pattern is not very productive. There are limited examples of NN exocentric compounds and among them are; *hatchback* and *skinhead.* A sub-category of exocentric NN compounds is appositional compounds. Bauer notes that in these compounds, the first element marks the sex of a person. For example; *boy-friend*, *manservant* and *woman-doctor*. For Bauer, this type of compounding is productive although it is usually used more with pronominal markers that identify the sex of animals such as; *she-goat* and *he-cheater*.

According to Bauer, the category of endocentric NN nominal compounds includes different patterns of compound words. One of these patterns is the combination of a gerund which ends in *–ing* and a noun as exemplified below.

21.	sleeping sickness	N + N	[is sickness which causes sleeping]
	shooting match	N + N	[is a match in which there is shooting]
	fishing rod	N + N	[rod used for fishing]

Lastly, Bauer explains that the most productive endocentric NN compound pattern is where two common nouns combine as shown in the examples.

22.	acid rock	brain surgery	cup cake
	bullet train	crew member	apple cake
	suicide seat	schoolbook	table mat
	clog dance	ginger bread	flowerpot

For B & S (2005), endocentric compounds and exocentric compounds are divided into subordinative (SUB), attributive (ART) and coordinative (CRD). These subtypes have already been discussed in detail in subsection 2.4.1 above, and most of the examples given belong to the category of NN compounds.

In a nutshell, the interpretation of NN nominal compounds is so diverse that it may depend on the argument structure of the head, and also the semantic relation between the two nouns. In some cases, the interpretation of the nouns may depend on the background knowledge of the interlocutors**.**

2.5.1.2 AN nominal compounding pattern

This compounding pattern is made up of an adjective in the non-head position and a noun in the head position. The pattern resembles noun phrases (NPs). Bauer (2004) notes that one way to distinguish the AN compounds from the AN phrase is through the marking of stress. Compound words mark stress on the left while phrases mark stress on the right in English. In (23), the stress-bearing elements are marked in bold.

23.	**Compound**	**Noun phrase (NP)**
	blackbird	black **bird**
	greenhouse	green **house**
	deep-structure	deep **structure**
	hard-hat	hard **hat**

AN compounds, like NN compounds, can either be endocentric or exocentric (Padrosa-Trias, 2010). Some of the examples that form the AN endocentric compounds are; *greenhouse*, *blackbird*, *white house*, and *blackboard*.

In some AN compounds, the meaning is not fully compositional. For example, the noun *bird* is the head of the compound *blackbird*, which is interpreted as a 'kind of bird', but the understanding of what kind of bird a *blackbird* is, cannot be construed from the compound constituent *black*. Plag (2003:151) notes that the high lexicalisation of AN compounds shows that these compounds are not as productive as NN compounds although the interpretation of these compounds, like that of NN compounds, is dependent on the modifier-head relation. Further, Padrosa-Trias (2010) indicates that in AN compounds, plural marking is done on the head just like in endocentric NN compounds. This shows that the head of the compound is what generally gives the compound as a whole its grammatical category.

For exocentric AN compounds, and as mentioned before, the head of the compound does not contribute to the meaning of the compound. The missing reference is outside the compound word. Some of the examples include:

24.	bluebell,	redhead	longlegs	flatfoot	greenback
	greybeard	redneck	longnose	bignose	redcap
	loudmouth	boldface	lazybones	hothead	red-eye
	paleface	fathead	shorthorn	redbreast	

Taking *greybeard* as one of the examples of AN exocentric compounds, one notices that the reference is not literally to the *beard* that is grey, but to a person with a grey beard. This kind of compound is interpreted metonymically and generally considered lexicalised. However, a closer look at these exocentric compounds shows that the second constituent, the noun, can function as the head while the adjective can function as a modifier. Take for instance the compound word *bluebell*, which refers

to a type of 'flower with a bell-like structure', the adjective blue seems to modify *bell* literally. Compounds like *bluebell*, *redhead*, *fathead* and *shorthorn* whose nouns can be modified resemble the attributive NN compounds. In attributive NN compounds, the non-head modifies the head noun. For example, a *bell jar* is a kind of *jar*, which means *bell* modifies *jar* and *jar* is the semantic head of the compound. When pluralised, *jar* takes the plural (e.g. bell jars). Drawing from the similarities of modifier-head relation between NN and AN, it seems some exocentric AN compounds have attributes of endocentric compounds (see also Padrosa-Trias, 2010).

2.5.1.3 VN nominal compounding

Nominal compounds that have the structure VN can also be categorised as exocentric or endocentric. For endocentric VN compounds, the noun is not the object of the verb but the head of the compound. In this compounding pattern, the verb takes the non-head position because it precedes the noun. As a consequence, it behaves like a modifier of the noun head. For example, a *payday* is 'a day when someone gets paid', *call boy*, is 'a boy who calls'. Canonically, verbs do not modify nouns, but adjectives do. This is the reason why it is not unproblematic to assume that the element in the non-head position is, in fact, a verb, and rather not a noun or an adjective. However, *pay* and *call*, for example, are also nouns in English. Bauer (1983) and Padrosa-Trias (2010) acknowledge this semantic ambiguity of the VN compounding pattern. However, Bauer (1983:205) argues that genuine VN compounds exist and may be expressed in compounds such as *crash pads*, *play pit* and *goggle-box*.

As for exocentric VN compounds, the semantic head of the compound is external to the compound. These compounds also resemble synthetic compounds because the second member, the noun, is understood as the object of the verb (Carstairs-Mc Carthy, 2002: 64):

25. pickpocket
    cut-throat
    killjoy
    cutpurse

In these examples, *pocket*, *throat*, *joy* and *purse* are objects of the verbs although they are not the semantic heads of the compounds. For instance, the compound word *pickpocket* does not refer to the pocket but to someone who picks items from people's pockets. Despite being headless, these

compounds are still nominal. All the VN compounds in example (25) above, denote people. However, some VN compounds denote animals, for example; *wagtail, tumble-dung*, while others denote plants such as *catchfly*, *heal-all*, or objects, for example, *breakwater* and *scarecrow* (Lieber, 2009c:287; Kornfeld, 2009; Padrosa-Trias, 2010:145). Most of the VN compounds that denote people have a *derogatory* interpretation compared to compounds that denote animals and objects. These compounds also tend to have lexicalised meaning (see also Bauer, 1983; Carstairs-Mc Carthy, 2002; Plag, 2003), and as a result, are not very productive. Padrosa-Trias (2010) argues that by being nominal, VN compounds have a zero N that refers to a person and gives the compound its nominal category. This implies that the grammatical relation between the compound constituents is that of subordination (i.e. the noun is an internal argument of the verb).

2.5.1.4 PN nominal compounding

These types of compound words are formed by combining a preposition as the first constituent with a noun (PN). The relation between the preposition and the noun is that of a modifier-head structure (Plag, 2003:152). Some examples are; *afterlife, outpost and underarm.* The preposition in these examples modifies the noun. For example, in the word *outpost*, the noun *post* refers to a type of station also called post in military terms and the preposition *out* describes the remoteness of the post. In this compound word, *post* is the head of the compound because the semantic interpretation of the compound as a noun depends on the features that the noun *post* possesses. The preposition *out* in this case modifies the head noun *post*, meaning the compound expresses an attributive relationship. Other similar examples are shown in example (26).

26.

outpost	aftertaste	ingroup
undercoat	overcoat	aftereffect
outbuilding	outroom	afterbirth
afterthought	overcoat,	afterlife

Some PN compounds also have a SUB nature and Padrosa-Trias (2010) identifies two types of subordinate PN compounds as endocentric and exocentric. Subordinate endocentric PN compounds relate to a P that has been subordinated to the head noun, for example, *in-joke*, 'a joke understood by a selected few people', and an *out-tray* as a 'tray for the content which is ready for dispatch'. For these compounds, the noun is both the semantic and formal head and takes the plural. As regards exocentric PN compounds, their reference concerns an entity outside the compound word. Examples are:

underbelly, which refers to the 'part of the body below the belly', *underground* may refer to 'a railway system which is under the ground'; *afterbirth*, which refers to the 'placenta and foetal discharge from uterus' and *underarm*, which refers to 'armpit' (i.e. a place under the arms). In these compounds, plural marking is on the second element despite the semantic head being external to the compound. This suggests that inflections affect right-hand elements whether the compound is endocentric or exocentric. It is noteworthy that some PN compounds are lexicalised (i.e. have a fixed meaning) and therefore not treated as a combination of a P and N semantically. Some of the examples include; *uprising* which now refers to 'rebellion', *underdog* which may refer to 'a weaker person' and *oversight* which refers to 'an unintentional omission of something' (Padrosa-Trias, 2010:149). Furthermore, some PN combinations may not be compounds but may be analysed as words derived via affixation of P-elements, for example, words such as *overdose, overtax* and *overkill* (see also Padrosa-Trias, 2010).

In sum, some PN compounds are identified as endocentric while others are exocentric. Further, the noun in PN compounds is described by a particular place/position, for example, *under* is describing a place or position of a noun, that is, a 'ground'.

2.5.2 Adjectival compounding patterns

Adjectival compounding patterns are diverse because they comprise different grammatical categories. The prototypical patterns may include either the verb or noun in the non-head position and the adjective in the head position; for example; *sugar-free*, *knee-deep* and *bush-green.* Other compounds in this category have an adjective in the non-head position. These are compounds like *white-collar (job)* (A+N) and *blue-sky* (A+N). Other adjectival compounding patterns do not have an overt adjective as one of the constituents of the compound. Some of the examples are *make-believe* (V+V) and *cutthroat* (razor) (V+N) (Bauer, 1983; Conti, 2006). Such compounds are considered exocentric because they function as adjectives despite the second compound element (i.e. the head) not being an adjective. In this section, I do not discuss all the adjectival compounding patterns that have been listed in the literature but I only discuss two types of adjectival compounding patterns namely; those that have an adjective as a head. These are NA compounds and AA compounds.

2.5.2.1 NA adjectival compounding pattern

The NA adjectival compounds comprise the noun in the non-head position and the adjective in the head position. The non-head can serve two roles; firstly, as a modifier and secondly, depending on

the relation with the adjectival head, as an argument of the head (Plag, 2003), as shown in the following examples;

27. a. dog-lean dog-tired
sky-blue knee-deep
ice-cold snow-blind
sea-sick olive-green
Crystal-clear

(Conti, 2006: 60, 170)

b. sugar-free
girl-crazy
class-conscious
structure-dependent
blood-red
lead-free

(Plag, 2003: 152)

Plag notes that the compounds in (27a) are characterised by a modifier-head relation expressed between the noun and the adjective. For Plag, the most common semantic interpretation for compounds in (27a), is comparison. For example, *blood-red* can be interpreted as, 'as red as blood', and *dog-lean* as, 'as lean as a dog.' Although this view is also shared by Conti (2006), she argues that it is difficult to determine the degree of comparison between the relationships of the constituents of the compound. She claims that an example such as *dog-tired* does not immediately invoke a picture of a dog that is 'as tired as….'

For the compounds in (27b), the non-head can be analysed as an argument of the head where the first element satisfies the argument position of the adjective. In the corresponding syntactic structure, the argument of the adjective would be preceded by a preposition. For example, *lead-free* can be interpreted as 'free of lead' and *class-conscious* as 'conscious of class' (see Plag, 2003).

2.5.2.2 AA adjectival compounding pattern

The AA adjectival compound involves combining two adjectives, (AA). As in the NA adjectival compound, the AA adjectival compound may express an attributive relationship where the first element

may be a modifier, for example*, icy-cold* and *bluish-green*. Bauer (1983) identifies two main semantic groups through which adjectival compounds can be classified and these are: appositional compounds (see also coordinative compounds) for example, *bitter-sweet* and *deaf-mute* and endocentric compounds, for example, *open-ended* and *ready-made*. In B & S's (2005) classification of compound words, appositional compounds and endocentric compounds do not belong to the same tier. Appositional compounds are classified either as endocentric (e.g. *snailmail*) or exocentric (e.g. 'crossword) to refer to puzzle in English. Therefore, the method adopted by Bauer (1983) of classifying AA adjectival compounds as either appositional or endocentric contradicts B & S's argument of classifying compounds as either headed or non-headed. In Bauer's classification, appositional and endocentric compounds belong to the same tier. This gives the impression that the two types are not related and that the appositional compounds cannot be endocentric.

B & S (2005), Scalise and Bisetto (2009), and Plag (2003) call compounds such as *bitter-sweet*, *blue-green* and *deaf-mute* coordinative compounds and classify them as either endocentric or exocentric. However, Padrosa-Trias (2010) suggests that the only way these coordinate adjectives can make sense is by inserting them in the non-head position of a compound headed by another noun, such as *harsh-rude tongue*, where *tongue* is an outside noun to the coordinate non-head, *harsh-rude*. In this case, the compounds have an endocentric interpretation because the coordinate non-head structure, *harsh-rude* modifies the head-noun *tongue*.

Concerning the analysis of the AA compounds, Szymanek (2009) identifies two possible structures of AA adjectival compounds, namely; coordinate combination (*e.g. fruit-vegetable shop*) and modifier plus head combination (i.e. with subordination) for colour terms such as *dark-red*, *light-yellow*, which show subordination of the first element. For instance, *dark-red* implies a kind of red which is dark thus *red* is the head and *dark* is a non-head.

Lastly, other AA adjectival compounds have derived adjectives as heads, and some of the examples include: *blue-eyed, clear-cited, university-controlled and hair-raising*. Plag (2003:153) identifies two possible ways to analyse the structure of such compounds. One way is to assume that the suffix *–ed* in *blue-eyed* is attached to the phrase as in *[[blue eye]-ed]* or to the head as in *[blue [eye-ed]].* Between the two structural analyses, Plag favours the structure where the affix is attached to the phrase because the structure provides a semantic interpretation of 'having *blue eyes*'. However, this is opposed to adjectivally used past participles which are heads of compounds such as *university-controlled*. For such a compound, Plag favours the structure where the *–ed* participle attaches to the

second word as in *[university [-controlled-ed]],* which gives the compound the interpretation of the non-head being in an agent argument relation with the head. The significance of this for the analysis of phrasal compounds is that the two alternatives are reminiscent of the two possible analyses of synthetic compounds and also seem to apply to VN compounds discussed in Chapter 3.

2.5.3 Verbal compounding patterns

Few verbal compounds in English are listed in the literature. Some linguist (see Booij, 2005; Lieber, 2009c) claim that verbal compounding in English is not a productive process. Bauer (1983:207-209) suggests the following combinations of form classes for verbal compounds: N+V (e.g. *head-hunt*), V+N (e.g. *shunpike*), V+V (e.g*. freeze-dry*), A+V (e.g. *soft-land*) and P+V (e.g. *overbook).* Bauer further indicates that most of the compound verbs are formed by backformation and conversion (see also Plag 2003; Booji, 2005). One of the compound patterns which is said to be a product of backformation and sometimes conversion is AV. However, Padrosa-Trias (2010) doubts if AV compounds are genuine compounds considering that they are derived from nominal or in certain instances adjectival compounds through backformation. Lieber (2009c) on the other hand argues that although backformation occurs in English compounds, it is not a very productive process of forming verbal compounds**.** Padrosa-Trias (2010) discusses three types of compound verbs: NV, VV and AV and asserts that all the three compound combinations are endocentric and that the verb acts as a formal and semantic head. With regards to the grammatical relations between the two compound constituents, only subordinative and attributive relations are assumed to be present while the coordinative relation is considered absent.

In this section, I discuss two types of verbal compounds. The first is the NV and the second is the VV pattern. These patterns, though rare, are traceable in some languages (e.g Catalan, English, Dutch, Mapudungun, Japanese, and Chinese) (see Padrosa-Trias, 2010; Booij, 2005; Baker and Fasola, 2009; Haspelmath and Sims, 2010, Altakhaineh, 2016).

2.5.3.1 NV verbal compounds

The NV verbal compound consists of a noun as the first constituent and a verb as the second one. The combination of a noun and a verb to form a verbal compound is generally referred to as noun incorporation (Baker, 1988; Booij, 2005:92). Some of the examples are *shop-lift, head-hunt* and *sky-dive*. The incorporated nouns in NV verbal compounds do not refer to specific objects in English; they

are non-referential. Additionally, the verbal compounds are also products of backformation from nominal or adjective compounds.

There are different views concerning the formation of NV compounds. Some linguists dispute the existence of these compounds (e.g. Selkirk, 1982), while others argue that they are formed by the process of backformation from nominals or adjectives (Bauer, 1983, Plag, 2003; Booij, 2005). This argument is based on the claim that a nominalising affix can be added to some verbal compounds. Some of the examples of words that are formed through the process of backformation include; *talent-spot* derived from *talent-spotter* and *trickle-irrigate* from *trickle-irrigation*. Padrosa-Trias (2010:156) argues that despite the possibility of some NV verbal compounds being derived through the process of backformation or conversion, they are not different from compounds that are base generated such as *sight-read*, *chain-drink* and others. She thus considers verbal NV expressions as compounds.

Further, Padrosa-Trias considers NV verbal compounds endocentric and categorises them into endocentric subordinative and endocentric attributive compounds. Examples of endocentric subordinative compounds may include *babysit* and *brainwash* while those of endocentric attributive compounds may include *chain-drink* and *ghost-write*. As regards the head, Padrosa-Trias notes that the verb in the verbal compound is both the formal head and the semantic head. Formal, because it determines the verbal status of the compound and also because it inflects for tense and agreement. Semantic, because the compound is a hyponym of the verb (i.e. the meaning of the entire compound is included in the meaning of the verb). Concerning the noun in the NV verbal compound, Padrosa-Trias explains that it can be interpreted as the instrument used to perform an action in certain instances. For example; *computer-generated (i.e.* something that is generated by a computer) and *steam clean* (i.e. cleaning using steaming). She also observes that the noun may function as an internal object of the verb when in fact it may not in a syntactic construction. For example, with the compound word *babysit*, one can construct a sentence such as *They babysat John all afternoon.* Here, the pronoun *they* is the external argument and *John* is the internal argument, not the noun *baby*. In this case, the entire compound functions as a verb.

2.5.3.2 VV verbal compounds

The VV type of verbal compounds is formed by combining two verbs (VV). Few examples exist in English so far.

28. freeze-dry stir-fry
drop-kick slam-dunk
roll-start

Lieber (1983:265)

Although VV verbal compounds are said to be rare and unproductive, Lieber (2009) notes the presence of VV endocentric compounds in English. In most of the VV verbal compounds, the relationship that exists between the two verbs is not systematic. This is in contrast with NN and AN compounds, where the relationship between the constituents in the compound is clear. By default, the constituents in VV compounds can be interpreted as having a modifier-head or argument-head relationship. However, it is quite difficult to identify the head in examples such as *stir-fry* because semantically, both constituents of the compound seem to be heads just like coordinative compounds. Moreover, having the first constituent as a verb may defeat the whole concept of modification because verbs rarely function as modifiers**.** This observation is supported by Baker and Fasola (2009:601), who claim that in Mapudungun, an indigenous language of South America, most VV compounds are partially idiomatic, and as a result, it is difficult to identify the element that is the head of the compound. However, Padrosa-Trias (2009) argues that VV verbal compounds are endocentric. Thus, the second verb is both a formal head (i.e. the past tense inflection generally attaches to it) and a semantic head (i.e. the compound as a whole denotes a subordinate action expressed by the second verb) while the first verb, like in NN and AN compounds, is a modifier (i.e. a manner/temporal modifier). Plag (2003) notes that VV compounds are quite difficult to classify because they resemble coordinative compounds, whose elements may have integrated into one event. For example, *stir-fry,* can mean 'to fry and stir' at the same time.

2.6 Conclusion

This chapter discussed the theory of compounding in general. Different definitions of compound words were discussed as well as the main aspects of the definition(s) and other core features. This is followed by a review of the literature on different types of compound words in section 2.4 which included endocentric compounds, exocentric compounds, synthetic compounds, root compounds and phrasal compounds. Endocentric and exocentric compounds were further analysed based on the grammatical relations between the constituents of the compounds; that is, whether they are subordinative, attributive or coordinative. Synthetic compounds were also discussed and two main theories that dominate the analysis of these compounds were reviewed. The first is the syntactic theory which

assumes the syntactic approach to synthetic compounds. In this approach, two lexical items combine first before the category changing suffix is attached. The second is the morphological theory which assumes the lexicalist approach and the nominalising suffix merges with the verb first and then combines with the noun. For the root compounds, it was noted that they have no derivational morphology. The constituent parts are free morphemes. In discussing the phrasal compounds, it was noted that the compounds are formed by joining a phrasal element and a noun. The phrasal element is the first constituent and the noun is the second and that the constituents in the compounds cannot be separated. The discussion on different types of compounding patterns revealed that NN compounds are the most productive in English.

CHAPTER 3: VN-COMPOUNDS IN CINYANJA

3. Introduction

In this chapter, I analyse a category of phrasal compounds in Cinyanja that are made up of a verb and noun which are grammatically or semantically related as exemplified below:

1. a. ci-pha-dzuwa
   7-kill-FV-5.sun
   'killer of the sun' (literal meaning)
   'a beautiful girl' (lexicalised meaning)

   b. m-pond-a ma-tiki
   1-step-FV 6-tickey
   'one who steps on tickeys' (literal meaning)
   'millionaire' (lexicalised meaning)

   c. ci-gon-a-mu-bawa
   7-sleep-FV-18.LOC-5.bar
   'it sleeps in the bar' (literal meaning)
   'a drunk' (lexicalised meaning)

In the examples above, the compounds comprise two or more words namely; *cipha* 'killer' and *dzuwa* 'sun in (1a), *mponda* 'one who steps' and *matiki* 'tickeys' in (1b) and *cigona* 'it sleeps' and *mubawa* 'in the bar' in (1c). These verb-noun (VN)-compounds as shown in (1) above, despite having a complex structure, are genuine compounds which include phrasal material and inflectional morphology. In this chapter, I claim that Cinyanja allows "reification", by which I mean, following Harley (2009), that lexicalised complex phrasal expressions can be re-introduced into the syntax as unanalysable roots and then undergo further morpho-syntactic operations.

The chapter first discusses the structure of VN-compounds in Cinyanja followed by a review of the literature on noun class (NC) prefixes and agreement. The phrasal structure of different types of VN-compounds is discussed. This is followed by a syntactic discussion of VN-compounds that shows that externally, they behave like nouns. Based on the Lexical Integrity Hypothesis (LIH), I show that, although the internal structure of the compounds is syntactic, it is not accessible by syntax. Finally, I

propose that the phrasal expressions in Cinyanja VN-compounds can be reified as roots and re-enter the syntactic derivation as unanalysable units. Finally, a conclusion is provided.

3.1 VN-compounds

VN-compounds form the bulk of phrasal compounds in Cinyanja and are particularly interesting because they share a similar structure with synthetic compound words in English. The compounds have a deverbal noun as the first member followed by an internal argument. The term 'VN-compounds' is used here as it anticipates the findings of section 3.3 where it is shown that they behave as genuine phrasal compounds. In this study, VN-compounds are classified into types 1-3. For each type, I show both the singular and plural forms. The singular forms are all classified as (a) while the plural forms are classified as (b). For the mass nouns, no plural example is provided because the structure is exactly the same as the singular one.

3.1.1 Type 1: Nominal part of the compound is a non-derived N

Type 1 VN-compounds comprise a verb and a non-derived noun in the nominal part. Generally, the verb is preceded by a noun class marker (henceforth, CM). The compounds in this category can, therefore, be schematically presented as; **Type 1: CM + V + N**

The data below is representative of Type 1 compound words.

2. a. m-tenth-a-ndevu
   3-burn-FV 9.beard
   'beard warmer' (literal meaning)
   'tea' (lexicalised meaning)

3. a. mu-sung-a cuma
   1-keep-FV 7.wealth
   'wealth keeper' (literal meaning)
   'treasurer' (lexicalised meaning)

b. a[4]-mu-sung-a cuma
2-1-keep-FV 7.wealth
'wealth keepers' (literal meaning)
'treasurers' (lexicalised meaning)

4. a. ci-ph-a-dzuwa
7-kill-FV-5.sun
'killer of the sun' (literal meaning)
'beautiful girl' (lexicalised meaning)
b. a-ci-ph-a-dzuwa
2-7-kill-FV-5.sun
'killers of the sun' (literal meaning)
'beautiful girls' (lexicalised meaning)

5. a. m-pal-a-utsi
1-scrape-FV-14.smoke
'one who scrapes smoke' (literal meaning)
'liar' (lexicalised meaning)
b. a-m-pal-a-utsi
2-1-scrape-FV-14.smoke
'smoke scrapers (literal meaning)
'liars' (lexicalised meaning)

6. a. ka-byal-a-ufa
1.a-plant-FV-14.mealie-meal
'planter of mealie-meal' (literal meaning)
'illegitimate child' (lexicalised meaning)
b. a-ka-byal-a-ufa
2-1a-plant-FV-14.mealie-meal
'planters of mealie-meal' (literal meaning)
'illegitimate children' (lexicalised meaning)

[4] In Cinyanja, the class 2 plural prefix *a-* may also be used as an honorific prefix for a person, e.g *amai*, *aBanda*, , a*Mwanida* etc., (see Scotton 1980:38)

7. a. m-tol-a-nkhani
1-collect-FV-9.story
'collector of stories' (literal meaning)
'news reporter' (lexicalised meaning)
b. a-m-tol-a-nkhani
2-1-collectFV-10.story
'collectors of stories' (literal meaning)
'news reporters' (lexicalised meaning)

In (3-7), the plural compounds categorised as (b), have a plural marker prefixed to the entire compound word. This additive plural formation (i.e. adding the class 2 prefix and preserving a class 1 prefix) is only found in compounds not in simple class 1 nouns such as *m-nyamata* 'boy' whose plural is *a-nyamata* 'boys' not **a-m-nyamata*. However, When the singular noun class prefix is subtracted from the compound as in (8b), the compound expression becomes unacceptable.

8. a. m-tol-a-nkhani
1-collect-FV-9.story
'news reporter'
b. *a-tol-a-nkhani
2-collect-FV-10.story
'news reporters'

Apart from example (2) which is a mass noun and denotes something inanimate, examples (3-7) refer to humans and have their plurals in class 2.

In this type of VN-compounds, the V constituent is made up of the verb stem and the final vowel *-a,* which Nurse (2008) refers to as a neutral vowel. Most of the nouns in Type 1 compounds do not take a class prefix in the singular or plural. This is a general property of nouns in classes 5, 9 and 14, and mass nouns such as *cuma* 'wealth' in example (3) that do not take a prefix. Further, the general structure of these compounds resembles the famous English example of *truck-driver* (see Lieber, 2009c; Padrosa-Trias, 2010) where the affix seems to be attached to the VP (i.e. [[truck drive]-er]) not to the V (i.e. [[driv(e)]-er]) as shown in Chapter 2 section 2.4.3.

3.1.2 Type 2: Nominal part has an NC prefix

In the Type 2 VN-compounds, the nominal part of the compounds includes an NC prefix. This can be summarised as; **Type 2: CM+V+FV+CM+N**

Some of the data that form this type of compound are presented below:

9.	a.	ci-dy-a-ma-kanda	
		7-eat-FV-6-baby	
		'eater of babies'	(literal meaning)
		'sugar daddy'	(lexicalised meaning)
	b.	a-ci-dy-a-ma-kanda	
		2-7-eat-FV-6-baby	
		'eaters of babies'	(literal meaning)
		'sugar daddies'	(lexicalised meaning)
10.	a.	ci-tsek-a-ma-were	
		7-close-FV-6-breast	
		'one who closes the breast'	(literal meaning)
		'last born child'	(lexicalised meaning)
	b.	a-ci-tsek-a-ma-were	
		2-7-close-FV-6-breast	
		'the ones who close the breast'	(literal meaning)
		'last born children'	(lexicalised meaning)
11.	a.	ci-ph-a-ma-so	
		7-kill-FV-6-eye	
		'killer of eyes'	(literal translation)
		'hypocrite'	(lexicalised translation)
	b.	a-ci-ph-a-ma-so	
		2-7-kill-FV-6-eye	
		'killers of eyes'	(literal translation)
		'hypocrites'	(lexicalised translation)

12. a. ka-pent-a-mi-lomo
1-paint-FV-4-lip
‘painter of lips’ (literal meaning)
‘prostitute’ (lexicalised meaning)

b. a-ka-pent-a-mi-lomo
2-1-paint-FV-4-lip
‘painters of lips’ (literal meaning)
‘prostitute’ (lexicalised meaning)

13. a. m-pond-a-ma-tiki
1-step-FV 6-tickey
‘one who steps on tickeys’ (literal meaning)
‘millionaire’ (lexicalised meaning)

b. a-m-pond-a-ma-tiki
2-1-step-FV 6-tickey
‘the ones who step on tickeys’ (literal meaning)
‘millionaires’ (lexicalised meaning)

14. a. m-pal-a-ma-tabwa
1-scrape-FV-6-plank
‘plank scraper’ (literal meaning)
‘carpenter’ (lexicalised meaning)

b. a-m-pal-a-ma-tabwa
2-1-scrape-FV-6-plank
‘plank scrapers’ (literal meaning)
‘carpenters’ (lexicalised meaning)

15. a. m-kond-a-ma-ungu
1-love-FV-6-pumpkin
‘lover of pumpkins’ (literal meaning)
‘pumpkin lover’ (lexicalised meaning)

b. a-m-kond-a-ma-ungu
2-1-love-FV-6-pumpkin
'lovers of pumpkins' (literal meaning)
'pumpkin lovers' (lexicalised meaning

16. a. mu-dob-a-ci-dzala
1-pick-FV-7-garbage
'collector of garbage' (literal meaning)
'garbage collector' (lexicalised meaning)
b. a-mu-dob-a-ci-dzala
2-1-pick-FV-7-garbage
'collectors of garbage' (literal meaning)
'garbage collectors' (lexicalised meaning)

17. a. m-tsuk-a-ma-no
1-wash-FV-6-tooth
'teeth cleaner' (literal meaning)
'name of a shrub used to brush teeth' (lexicalised meaning)
b. mi-tsuk-a-ma-no
4-wash-FV-6-tooth
'teeth cleaners' (literal meaning)
'name of shrubs used to brush teeth' (lexicalised meaning)

18. a. m-tanthauzira-ma-wu
3-translate-FV-6-word
'translator of words' (literal meaning)
'dictionary' (lexicalised meaning)
b. mi-tanthauzir -a-ma-wu
4-translate-FV-6-word
'translator of words' (literal meaning)
'dictionaries' (lexicalised meaning)

From the compounds listed above, some general observations can be drawn.The first is that the verb in the compounds is either followed by a plural noun as in (9-15 and 17-18) or a mass noun as in (16).

Concerning the interpretation of these compounds, some display mixed properties of being transparent and having lexicalised meanings as in examples (14) *mpalamatabwa* 'plank scraper' and example (15) *mkondamaungu* 'pumpkin lover'. In these compounds, the semantic meaning can be construed from the compound elements, thus, the compounds are endocentric, with 'plank scraper' being the hyponym of scraper and 'pumpkin lover' being a hyponym of lover. Further, each of the compound constituents is preceded by a NC prefix.

The V element in Type 2 VN-compounds is bimorphemic and comprises the verb root and a final vowel (FV) *–a* which is mostly attached to verbs in these compound words. In rare cases, the verb may take an applicative morpheme *–ir-* between the root and the final vowel as shown in example (18). This shows that the verb can take derivational morphology.

The noun element in VN-compounds presented above thematically corresponds to the direct object of the verb. In general, most Type 2 nouns are bimorphemic. They consist of a noun class prefix marking number/gender and a stem. Most of the nouns in this category refer to humans and objects while a small number may refer to animate organs as in (10) *mawere* 'breasts' in *citsekamawere* 'last born child'.

In VN-compounds, the noun is the internal argument of the verb. What is interesting about this type of compound words is that the noun stem is preceded by a plural noun class prefix. Carstens (2008) associates the NC prefix with number and gender features and argues that it is a functional element, hence an inflectional marker. The plural inflection marker inside the compounds only affects the noun, not the entire compound. As stated earlier, to pluralise the compound expressions such as those in (9b-16b), the plural marker is added to the entire compound word in Cinyanja. This relates to Booij's (2005) argument about the position of inflectional morphology, namely- that plural marking is externally located in compound words. This is shown in the Cinyanja examples below.

19. a. a-m-pal-a-ma-tabwa
2-1-scrape-FV-6-plank'
'carpenters'

b. a-m-pond-a-ma-tiki
2-1-step-FV-6-tickey
'millionaires'

As regards internal plural inflection, it is not accessible for elements outside the compounds as below.

20. a. ka-pent-a-mi-lomo wa-kuda w-a-thaw-a.
1-paint-FV-4-lip 1.ASS-black 1.SM-PRF-run-FV
'The dark prostitute has run away'.

b. * ka-pent-a-mi-lomo ya-kuda y-a-tupa.
1-paint-FV-4-lip 4.ASS-dark 4.SM-BE-swell

(20b) shows that agreement on modifiers of the compound cannot be determined by the noun class of the nominal part of the compound, but obligatorily reflects the noun class of the whole compound, which is marked by the noun class prefix attached to the verbal part.

3.1.3 Type 3: Nominal part of the compound is a locative

The type 3 VN-compounds are typically composed of a verb and a locative in the nominal part of the compound, and can be summarised as; **Type 3: CM +V + FV+ LOC+ (CM) N**
The data in this category is presented below.

21. a. ci-pond-a-mu-thengo
7-step-FV-18.LOC-5.bush
'it steps in the bush' (literal meaning)
'consultation fee' (lexicalised meaning)
b. zi-pond-a-mu-thengo
8-step-FV-18.LOC-5.bush
'it steps in the bush' (literal meaning)
'consultation fees' (lexicalised meaning

22. a. ci-tsengul-a-pa-kamwa
7-open-FV-16.LOC-12.mouth
'it opens on the mouth' (literal meaning)
'bride price' (lexicalised meaning)
b. zi-tsengul-a-pa-kamwa
8-open-FV-16.LOC-12.mouth
'it opens on the mouth' (literal meaning)
'bride prices' (lexicalised meaning)

23. a. ci-gon-a-mu-bawa
7-sleep-FV-18.LOC-5.bar
'it sleeps in the bar' (literal meaning)
'a drunk' (lexicalised meaning)

b. a-ci-gon-a-mu-bawa
2-7-sleep-FV-18.LOC-5.bar
'they sleep in the bar' (literal meaning)
'drunks' (lexicalised meaning

24. a. m-khal-a-pa-mw-ala
1-sit-FV-16.LOC-3-stone
'one who sits on the stone' (literal meaning)
'stone sitter' (lexicalised meaning)

b. a-m-khal-a-pa-mw-ala
2-1-sit-FV-16.LOC-3-stone
'ones who sit on stones ' (literal meaning)
'stone sitters' (lexicalised meaning)

25. a. m-khal-a-pa-ci-tsa
1-sit-FV-16.LOC-7-tree stump
'one who sits on the tree stump' (literal meaning)
'stump sitter'(i.e. idle sitting)' (lexicalised meaning)

b. a-m-khal-a-pa-ci-tsa
2-1-sit-FV-16.LOC-7-tree stump
'ones who sit on the tree stump' (literal meaning)
'stump sitters' (actual meaning)

26. a. m-ses-a-pa-khomo
1-sweep-FV-16.LOC-5.yard
'sweeper of the yard' (literal meaning)
'maid' (lexicalised meaning)

b. a-m-ses-a-pa-khomo
2-1-sweep-FV-16.LOC-5.yard
'sweepers of the yard' (literal meaning)
'maids' (lexicalised meaning)

27. a. m-khal-a-pa-m-pando
1-sit-FV-16.LOC-3-chair
'one who sits on the chair' (literal meaning)
'chairman' (lexicalised meaning)

b. a-m-khal-a-pa-m-pando
2-1-sit-FV-16.LOC-3-chair
'the ones who sit on the chair' (literal meaning)
'chairmen' (lexicalised meaning)

28. a. m-gon-a-pa-kati
1-sleep-FV-16.LOC-middle
'one who sleeps in the middle' (literal meaning)
'a punctual person' (lexicalised meaning)

b. a-m-gon-a-pa-kati
2-1-sleep-FV-16.LOC-middle
'ones who sleep in the middle' (literal meaning)
'punctual people' (lexicalised meaning)

29. a. m-khal-a-pa-kati
1-sit-FV-16.LOC-middle
'one who sits in the middle' (literal meaning)
'mediator' (lexicalised meaning)

b. a-m-khal-a-pa-kati
2-1-sit-FV-16.LOC-middle
'ones who sit in the middle' (literal meaning)
'mediators' (lexicalised meaning)

30. a. m-low-a-m-malo
1-enter-FV-18.LOC-6.place
'the thing that enters in place of' (literal meaning)
'pronoun/substitute' (lexicalised meaning)

b. a-low-a-m-malo
2-enter-FV-18.LOC-6.place
'things that enter in place of' (literal meaning)
'pronouns/substitutes' (lexicalised meaning)

The compounds in this category refer to humans, objects and things. The majority of the compounds include class 16 or class 18 locatives as shown in examples (21-30) above. While examples where a class 17 locative prefix *ku-* may precede a noun may be possible, this may potentially be an interesting finding.

As noted above, compounds with the initial class 1 prefix *m-* that refer to humans, for example, *mkhalapacitsa* in (24), *msesapakhomo* in (26), *mkhalapampando* in (27), *mgonapakati* in (28) and *mkhalapakati* in (29), have the class 2 plural prefix *–a* added to the compound as a whole, preserving the class 1 prefix, this is not the case with class 1 simple nouns. Crucially, the compounds with the class 7 prefix *ci-* that refer to humans such as *cigonamubawa* 'drunk' in (23) also add the class 2 plural prefix *a-* and preserve the singular class 7 prefix *ci-,* instead of replacing the class 7 singular prefix with the expected class 8 plural prefix *zi-* which otherwise corresponds to singulars in class 7. In this respect, the compound in (23) differs from the compound words with the initial prefix *ci-* such as *cipondamuthengo* 'consultation fee' in (21) and *citsegulapakamwa* 'bride price' in (22) that refer to things and objects rather than to humans. These latter compounds take class 8 plural prefix *zi-* and not the class 2 plural prefix *a-* in the plural.

Another noteworthy example is the compound in (30) which refers to a thing but takes the class 1 prefix, which is typically found with human nouns. Interestingly, the plural of this compound is formed with the class 2 plural prefix, which in contrast to the other examples from class 2, replaces the class 1 prefix in (30b).

The verbal element in Type 3 compounds is also bimorphemic and consists of the verb root and a final vowel (FV). The final vowel takes the shape of *-a* in all the VN-compounds just like in Type 1 and Type 2 compounds. Another notable feature of the verb in these compounds is that it does not inflect for tense, and the final vowel seems to be a categoriser for verbs. The verb is followed by a locative in the nominal part.

The noun constituent that follows the locative prefixes in Type 3 compounds either has an overt NC prefix as in (24) *m-khala-pa-**mw**-ala*, 'stone sitter', or lacks an overt NC prefix as in (23) *ci-gona-m-bawa* 'a drunk'. Most of the nouns in this category that the locative prefix attaches to refer to things and objects and are usually singular, for example, *bawa* 'bar' in *cigonamubawa*, *khomo* 'yard' in *musesapakhomo* and *thengo* 'bush' in *cipondamuthengo*.

To conclude, for Type 1, a traditional compound analysis would be available because the verb (or deverbal nominal) combines with the bare form of the noun to simply form a complex morphological entity which can then be subject to further derivations/inflections (compare e.g. English [*truck* + *drive(r)*]-s). However, to understand the properties of Type 2 and 3 VN-compounds, an understanding of the nature of NC prefixes in Bantu and Cinyanja specifically is required. This is discussed in the next section.

3.2 NC prefixes and agreement

Recall from Chapter 1 that Bantu languages are classified according to a noun class system. The noun class system includes a system of concord where concordial agreement is dependent on noun class membership. In this nominal class system, nouns are categorised according to noun class prefixes which denote the number and gender of the nouns. The gender class is sometimes associated with semantic properties of nouns such as animacy, size, plurality, quality or location.

This section discusses Bantu noun class prefixes and the question of whether they are inflectional or derivational. I argue that noun class prefixes in compounds are inflectional and that they control agreement like other noun class prefixes attached to individual nouns.

3.2.1 The nature of NC prefixes; derivational or inflectional?

This section explores the nature of the noun class prefix and determines whether it is inflectional or derivational. Generally, Bantu noun class prefixes express grammatical information of number and gender. In some Bantu languages such as isiZulu, the NC prefix is optionally preceded by an augment vowel. This augment mirrors the vowel in the NC prefix as in the isiZulu examples ***u****m****u****ntu* *'*person'*, **izi**pho* 'gifts' *(*Taraldsen, 2010; Halpert, 2012:3*)*. In most Bantu languages, compound constituents are also usually preceded by an NC marker. This is shown in the Cinyanja examples below.

31. a. **ci**-ph-a-ma-so
7-kill-FV-6.eye
'killer of the eyes (literal meaning)
'hypocrite' (lexicalised meaning)

b. **m**-pal-a-ma-tabwa
1-scrape-FV-6-plank
'plank scraper' (literal meaning)
'carpenter' (lexicalised meaning)

In (31a) above, the initial NC prefix *ci-* marks the compound expression *ciphamaso* 'hypocrite' as a member of noun class 7 while in (b), the initial NC prefix *m-* marks the compound expression *mpalamatabwa* 'carpenter' as a member of noun class 1.

Another NC prefix inside the compounds *ciph**ma**so* and *mpala**ma**tabwa* is the NC prefix *ma-* which marks the nouns *maso* and *matabwa* as members of noun class 6. The locative is another NC prefix worthy of note inside some VN-compounds in Cinyanja shown in the examples below.

32. a. m-khal-a-**pa**-mw-ala
1-sit-FV-16.LOC-3-stone
'one who sits on the stone' (literal meaning)
'stone sitter' (lexical meaning)

b. ci-pond-a-**mu**-thengo
7-step-FV-18.LOC-5.bush
'it steps in the bush' (literal meaning)
'consultation fee' (lexicalised meaning)

In (32a) above the locative prefix *pa-* is inside the compound *mkhala**pa**mwala* 'stone sitter' and in (32b) the locative prefix *mu-* is inside the compound word *ciponda**mu**thengo* 'consultation fee'. Locatives generally belong to class 16, 17 and 18 in most Bantu languages.

As regards the NC prefix in Bantu languages, it has been analysed as either inflectional, (i.e. marks number and gender of nominals and is responsive to the demands of syntax) or derivational (i.e. gives rise to new lexical items or changes the meaning of the word). The inflectional view is supported by Katamba (2003) and Katamba and Stonham (2006) who assert that in most sub-Saharan African languages, nouns are classified according to the noun class system based on the prefixes they take. These noun class prefixes indicate number and gender which are inherent features of nouns. Number is expressed through the noun class prefix to distinguish singular and plural nouns, for example*, mu-*

is noun class 1 singular whose plural is *a-* of noun class 2 and *m-* is noun class 3 singular and the plural is *mi-* of noun class 4. In the vast majority of cases, a change of NC from 1 to 2, or 3 to 4 reflects a change in number, and a change in number is an inflectional category (Mufwene, 1980; Kinyalolo, 1991).

The derivational view is supported by Msaka (2019:34) who notes that varying the noun class prefixes attached to the nominal stem *-nthu* 'being', can result in a change of class membership from *munthu* 'person' of class 1 to *umunthu* 'personhood' of class 14, or to *cinthu* 'thing' of class 7. The mixed nature of NC prefix has been noted by many linguists (Mufwene, 1980; Sproat, 1985; Myers, 1987; Bresnan and Mchombo, 1995; Schadeberg, 2001, and others). Bresnan and Mchombo (1995) show that when a particular NC prefix is attached to a stem, it can show number for the noun, and can control agreement on the verb and sentence modifiers, which is an inflection role, and that it can also change the noun class and meaning of the stem, which is a derivational role.

Basciano, et al. (2011), in their comparative analysis of Bemba, Italian, and Mandarin Chinese, note that most Bantu nominal compounds are formed by a concatenation of two nominal stems, which are usually preceded by a noun class marker. They note that initial noun class markers, for example, *ka-* and *mu-* prefixed to Bemba compound words such as *ka-tensha-mabula* 'authoritative person' and *mu-sunga-bantu* 'hospitable person', are nominalising prefixes (see also Sproat, 1985) because they derive nouns from verbs. They argue that although these prefixes are homonymous with noun class 1a and noun class 1, they are derivational prefixes.

Further, Mletshe (2010) discusses different derivations of nouns in African languages. He notes that some derivations occur within noun classes, and eventually the classes have different semantic interpretations. He postulates that the derivational nature of Bantu noun class prefixes can be attested when a class 7 prefix is attached to a noun of class 1a in isiZulu. For example when the class 7 prefix *isi-* is attached to the class 1a noun *ubhanana* 'banana', it results in a derived noun *isibhanan*a which means 'plantation'. The meaning of the noun has changed, which suggests that a class 7 noun prefix has a derivational nature.

Mletshe also indicates that the commonest method of deriving nouns from verbs is by adding a prefix and suffix to the root. Taking Venda and Sotho as examples, he notes that in Venda, the noun *mushimi* (class 1) 'worker' is derived from the verb *–shum-* 'work' after adding the prefix *mu-* and the suffix *–i* to the root (see also Poulos, 1990:68; Mugane, 1997; Bresnan and Mugane, 2006). In Northern Sotho,

the noun *moruti* 'missionary' is derived from adding the prefix *mo-* and the suffix *–i* to the verb root '*–rut* 'teach' or 'instruct'. Importantly, this claim raises some questions as regards the derivational nature of the prefix. Given that derivation also involves a particular suffix, how do we know that the prefix is involved in the derivation process at all? It could be that it is the suffixes that are derivational and form nominal stems from (verbal) roots (i.e. they may be the spell out of the categorising little n of Distributed Morphology, briefly discussed in section 1.4.4, to which I return below). From this perspective, derivation is the domain of suffixes, while the noun class prefixes express grammatical information about gender and number, and may be inflectional.

Another argument concerns the derivation of infinitives which are nouns derived from verbs. While the derived word loses some of its verbal morphosyntactic properties, it gains certain nominal properties, on the other hand (Mletshe, 2010), which is a derivational feature. In most Bantu languages, infinitive nouns are formed by attaching the NC 15 prefix *ku-* to a verb/verb phrase (VP) to nominalise it. Myers (1987:96) shows that when an infinitive *ku-* is attached to a verb phrase in Shona, a nominal phrase is formed. The *ku- + VP* structure behaves like a noun because it can be modified by a possessor and an adjective, and both modifiers show class 15 concord as shown in (33).

33. ***Ku-*kora** kw-ake ku-kuru ku-no-shamisa.
15-fatten 15-his 15-great 15-HAB-amaze
'His great stoutness is astonishing'.

(see also Fortune,1985; Bresnan and Mchombo, 1995:219)

Schadeberg (2001:13) also notes that in Swahili, the infinitive formation is a derivational process, and that there is no distinctive role the noun classes play in controlling agreement. According to Schadeberg, words that are formed by attaching a class 15 noun prefix *ku-*, for example, *kuimba* 'to sing/the singing' cannot be marked for singular or plural. If noun class expresses true number (and is inflectional), it should systematically form plurals from singulars. However, Schadeberg shows that a singular count noun is derived from a singular mass noun by changing the noun class, for example, *ukuni* in class 11 means 'a piece of firewood' while *kuni* in class 10 means 'firewood'. This argument shows that this is not a systematic inflectional number process but a derivational process.

Contrary to the arguments that the NC prefix is derivational, Kosch (2011) posits that the noun class system in Bantu languages provides an overlap between derivation and inflection. Kosch explains that the feature number, which is traditionally classified as an inflectional feature of nouns, and represented by a prefix which marks number as the singular or plural form of the noun, can exhibit both derivational

and inflectional morphology when attached to nouns in Sesotho sa Leboa. Kosch further explains that nouns such as *se-tho* 'mankind' and *bo-thu* 'humanity' are created by a shift of noun class from class 1 (*mo-tho* 'person') to class 7 and class 14 respectively. This is a derivational process because class 7 and 14 prefixes seemingly change the semantic value of the words. According to Kosch, if the shift of noun class is a formal characteristic of noun-noun (N-N) derivation, then the formation of plural or singular may also be considered a derivational process in Bantu. Consequently, Kosch suggests that the boundary between inflection and derivation should not be assumed in advance. Other linguists in consonant with Kosch argue that the formation of singular and plural forms in Bantu languages is not a matter of inflection alone but also of derivation (Bresnan and Mchombo, 1995:183; Schaderberg, 2001:15; Van der Spuy, 2010; Mletshe, 2010, 2017).

Linguists such as Mufwene (1980) argues that although noun classes are identified using particular NC prefixes, these prefixes can have both inflectional and derivational properties. Mufwene is cognisant of the fact that Bantu noun classes play a significant role in inflectional morphology. However, Mufwene observes that the NC prefix *u-* when attached to a noun stem, is a mark of nominal derivation.

The conclusion therefore is that, the NC prefix seems to be both derivational and inflectional. The generative analysis presents an alternative to the mixed view. For example, the discussin so far has shown that there are arguments to treat NC prefixes in Bantu at least in part as derivational. However, in the next section, I will discuss the current generative analysis of the NC prefix which is that, it realises the category number hence it is inflectional.

3.2.2 Bantu NC prefixes in the generative literature

As earlier stated, in Bantu languages, nouns are classified according to number and gender denoted by the noun class prefixes. Gender, though sometimes is associated with the semantic properties of nouns such as animacy, size, plurality, quality or location (Bresnan and Mchombo, 1995; Fuchs and Van der Wal, 2018), most Bantuist see the relation between gender and meaning as unsystematic and entirely arbitrary.

In this section, I analyse arguments by Carstens (2008), Fuchs and Van der Wal (2018), and others which show that nominal stems belong to particular gender classes, and that noun class prefixes are inflectional number markers which realise singular and plural of a particular NC.

According to Carstens, NC in Bantu languages is a gender system. Each nominal stem belongs to a particular gender class, and the noun classes 1-11 are reduced to six genders (Carstens, 1991, 1993, 2008) as shown below:

34. Gender A: stems of Classes 1/2
Gender B: stems of Classes 3/4
Gender C: stems of Classes 5/6
Gender D: stems of Classes 7/8
Gender E: stems of Classes 9/10
Gender F: stems of Classes 11/10

According to Carstens, noun class prefixes are gender-specific realisations of singular or number features. For example, a class 2 noun such as *atsikana* 'girls', in Cinyanja, belongs to gender A, and the class prefix *a-* realises the plural specification of a noun belonging to this gender.

As noted above, some researchers such as Mufwene (1980), Myers (1987), Bresnan and Mchombo (1995) and Kosch (2011) consider the noun class prefix as both derivational and inflectional. However, Carstens (2008) argues against analysing the NC prefix as derivational. According to Carstens, category-changing derivational processes such as those discussed above are not triggered by NC prefixes, but by a zero morpheme.

First, Carstens (2008) notes that the derivational analysis of NC prefixes implies that the gender specification of a noun is provided by its prefix. However, this cannot be correct, as not every stem can freely combine with NC prefixes of any class. For example, the stems in (35) cannot combine with class 9 or class 3 prefixes:

35. a. *n-tu
9-person
'person'

b. *mi-atu
3-shoe
'shoe'

(Carstens, 2008:138)

Rather, stems must be inherently specified for gender, which explains the combinatorial restrictions observed with nominal stems and NC prefixes. Stems belong to a particular gender and are specified for class so that the correct pairing of the stem and its NC prefixes can be guaranteed.

A second argument against the derivational status of NC prefixes provided by Carstens (2008) is based on the derivation of augmentative and diminutive nouns in Swahili. Carstens notes that when primary NC prefixes are replaced by class 7 and class 8, the resulting noun receives a diminutive interpretation. Similarly, switching to classes 5 and 6 may result in augmentative meanings (Carstens, 2008:137) as shown below;

36. a. m-sumari/mi-sumari
3-nail/4-nail
'nail/s'

b. ki-sumari/vi-sumari
7-nail/8-nail
'little nail/s'

c. su-mari/ma-sumari
5-nail/6-nail
'big (ugly/nasty) nail/

Carstens also notes that when these seemingly diminutive and augmentative prefixes are attached to the noun, they replace the primary prefix in controlling agreement as shown in the following Swahili examples adapted from Carstens (2008:137).

37. a. M-sumari m-moja u-me-anguka
3-nail 3-one 3-agr-PRF-fall
'One nail has fallen'.

b. Ki-sumari ki-moja ki-me-anguka
7-nail 7-one 7-agr- PRF-fall
'One little nail has fallen'.

c. Ma-sumari ma-wili ya-me-anguka
6-nail 6-big 6-agr- PRF-fall
'Two big nails have fallen'.

At first sight, the data in (36) and (37) support the view discussed above, namely that the diminutive and augmentative meanings of the derived nouns are provided by the NC prefixes of class 7/8 and 5/6, and that therefore, these prefixes must be regarded as derivational (see Mufwene, 1980; Myer, 1987; Schadeberg, 2001; Ferrari, 2005). However, Carstens (2008), does not agree with this view. She notes that ordinary nouns of class 7/8 (Gender D), though they bear NC prefixes *ki/vi* in Swahili, do not have a diminutive reading. Similarly, ordinary nouns of class 5/6 (Gender C) have no augmentative readings as shown in the Swahili examples.

38. a. ki-tabu/vi-tabu
7.book/8.book
'book/s'
*'little book/s'

b. gari/magari
5.car/6.car
'car/s'
*'big car/s'

(Carstens, 2008:143)

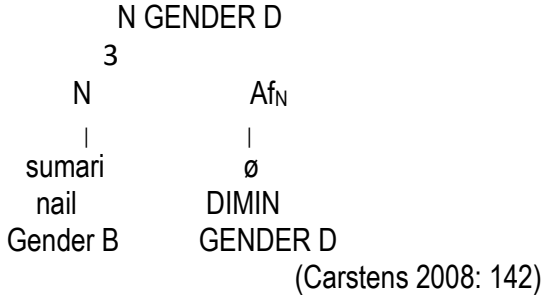
Recall from Chapter 1 that noun class 7/8 and 12/13 prefixes are used for augmentatives/ pejoratives and diminutives in Cinyanja. However, like in Swahili, some ordinary nouns of class 7/ 8 bearing *ci/zi* prefixes and 12/13 bearing *ka/ti* prefixes in Cinyanja, do not have an augmentative or diminutive reading as shown in the Cinyanja examples below.

39. a. ci-soti/ zi-soti
7-hat/8-hat
'hat/s'
*big hat/s'

d. ka-limba/ti-limba
12-musical instrument/ 13-musical instrument
'musical instrument/s'
*'small instrument'

From the examples above, it seems clear that the class 5/6 and 7/8 prefixes in Swahili, and class 7/8 and 12/13 prefixes in Cinyanja cannot be the source of the augmentative and diminutive meanings, as claimed by Mfuwene (1980), Bresnan and Mchombo (1995) and others. Carstens (2008) argues instead that these meanings are contributed by zero-affixes which are derivational and belong to particular genders. The derivation of the diminutive noun *kisumari*, 'little nail' is based on the zero-derived stem *sumari* shown in (40), and the addition of a singular number feature licenses the addition of the class 7 prefix *ki-* (gender D, singular):

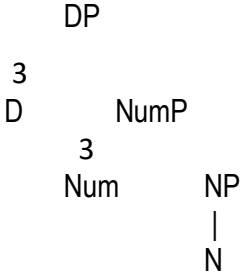
40



(Carstens 2008: 142)

In sum, Carstens argues that in Bantu languages, NC prefixes in general are inflectional number markers that realise the singular and plural of a particular gender; their derivational property is superficial, not real. Following Abney (1987) and much later work, Carstens (1993) analyses Bantu noun phrases as DPs and assume that the noun class prefix realises a functional category Num (for Number; Ritter, 1992) in the functional spine of the DP (see also Letsholo 2004; Herderson, 2006; Lusekelo, 2013).

41.



While Carstens (1993, 1997, 2008) treats NC prefixes as the gender-specific realisation of a functional category Num in Bantu, and gender as an inherent property of nouns and zero derivational affixes, Fuchs and Van der Wal (2018) extend Carstens' proposal by associating the gender feature of nouns in Bantu languages with another functional category, the so-called ("little") n, which acts as a nominaliser for categorially unspecified roots. Their analysis is based on the theory of gender articulated in Kramer (2015). Kramer posits that gender is a characteristic feature of nouns, but that

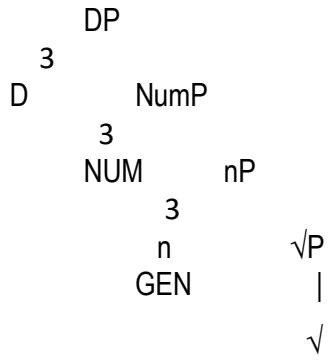
for nouns that have grammatical gender, gender is considered to be a property of n and only allotted to a root when the root combines with the nominaliser. Using Carstens' (1993) gender schema in (35), Fuchs and Van de Wal (2018:5) distinguish different "flavours" of little n, which mark the gender class of the nominal root with which it combines:

42.

n_A	stems of classes 1/2
n_B	stems of classes 3/4
n_C	stems of class 5/6
n_D	stems of class 7/8
n_E	stems of class 9/10

Fuchs and Van de Wal assume that in a nominal structure, number (Num) is generally projected above n where the gender category is marked as shown in the DP structure below.

43.



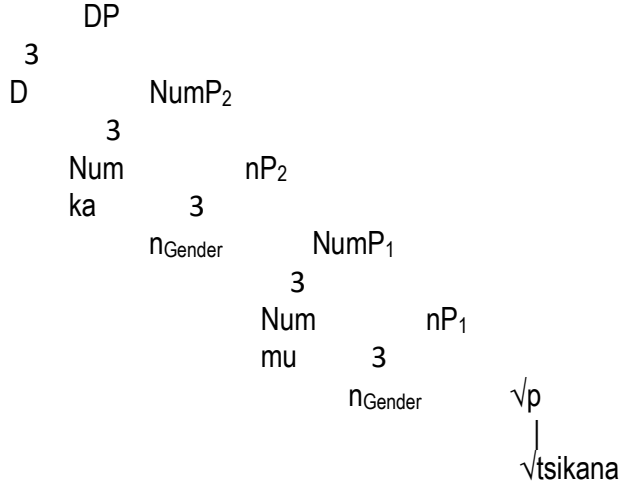
The structure above is in tandem with important aspects of Kramer's (2015) theory. Gender is located on n, and it is assigned to the root when it merges with a nominaliser. The manner in which nouns are assigned to gender categories is controlled by the exact type of n and the characteristics of each n in a language.

Fuchs and Van der Wal's (2018) proposal differs from Carstens' analysis in that they do not regard the NC-prefixes as spell-outs of the Num-category; rather, Bantu NC prefixes realise n in their theory. This phonological realisation of n is syntactically conditioned by the number feature of Num. When Num is singular, n is pronounced as the relevant singular NC of a particular gender. When Num is plural, n is realised as a plural NC. However, I show below that the nominalising head n may in fact be spelled out by derivational suffixes, rather than by NC prefixes in Bantu. I therefore continue to assume with Carstens (1993, 2008, 2011) that NC prefixes spell out Num, but according to the gender class specified in n.

The structure in (43) is based on the theory of Distributed Morphology (DM), according to which nouns are formed by an n merging with an unspecified root first (see Harley, 2009; Marantz, 2001; Borer, 2005; Fuchs and Van de Wal, 2018; for a brief review of DM, refer to Chapter 1). However, it is also possible for an n to merge with a phrase not headed by a root. Kramer (2015) adopts this idea and argues that in the formation of denominal nouns, the relevant n merges with a noun that corresponds to an already nominalised root, i.e. a root with a fully articulated nP. Fuchs and Van der Wal (2018:10) adopt this proposal and suggest that in the formation of diminutives and augmentatives in a Bantu language like Shona, one of the nominalisers in (43) first merges with a root phrase (√P), and the resulting nP then merges with the first Num which corresponds to the noun class prefix and the resulting NumP merges with the second nominaliser where the resulting nP merges with the Num which corresponds to the diminutive/augmentative to form the second NumP. The structure of diminutive and augmentative derived nouns can hence have a double n on a full nominal spine as illustrated in (45), which realises the Cinyanja diminutive in (44) (see Fuchs & Van der Wal 2018: 12).

44. Ka-m-tsikana ka-nzeru
12-1-girl 12-inteligent
'a small intelligent girl'

45.

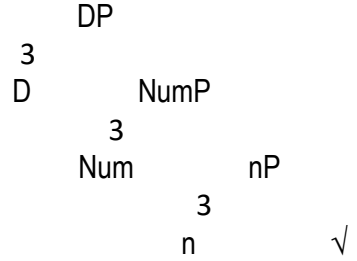


Fuchs and Van der Wal also note, following Kramer (2015), that 'the highest gender wins' in establishing agreement with DP-external elements. This is demonstrated with the Cinyanja example in (44) above, where the diminutive prefix *ka-* wins over the lower Num-head class 1 *m-* in establishing agreement on *kanzeru* 'intelligent' because it is the highest gender.

In conclusion, the discussions above show that in the current generative literature, Bantu NC prefixes are treated as spell-outs of a functional category, either Num or n. Gender may be inherently

associated with the nominal stem (Carstens, 2008), but may also be treated as being associated with a nominalising category (little n; Kramer, 2015) which attaches to acategorial roots to derive nominal stems. In languages such as Cinyanja, D may be silent, but in Bantu languages such as Zulu or Kinyarwanda, it may be realised by an augment vowel (see De Dreu, 2008)**.** So the general structure of a Bantu noun is (46), with n determining the gender class (A, B, C, D, E, F) and Num determining the number (singular or plural) of this gender class (together, gender + number = noun class):

46



In the analysis that follows, I adopt this view of the syntax of Bantu noun phrases, with a gender feature realised on n and a number feature (#) on Num. I treat NC prefixes as gender-specific spell-outs of Num. Furthermore, I also adopt the view, again based on work by Carstens (1993, 1997, 2008), that head movement of the root is up to D through Num and n. In this way, the noun class (gender and number) features are realised on the head of DP, and determine agreement and concord (see Zeller and Ngoboka, 2018 and section 3.2.3.2 below).

3.2.3 Locatives

As mentioned in section 3.2.2, many Bantu languages are characterised by a productive noun class system which includes locative prefixes. In most languages, locative class prefixes belong to class 16 *pa*, 17 *ku* and 18 *mu*.[5] Like other noun class prefixes, locative prefixes can attach directly to roots, in which case they behave like ordinary NC prefixes (Caha and Pantcheva, 2015:6) as in the Cinyanja examples in (47).

47. a. pa-kati
16.LOC-inside
'middle of'
b. ku-kati
17.LOC-inside

[5] However, Zeller and Ngoboka (2018) and many others note that other languages such as Kinyarwanda have a fourth locative noun class (class 25 *i*).

‘in the interior’

c. mu-kati
18.LOC-inside
‘inside’

However, locative prefixes are typically added to base nouns that have their primary prefixes preserved (Caha and Pantcheva, 2015; Zeller, 2017a), as shown by the Cinyanja examples in (48).

48. a. pa-ci-tseko
16.LOC-7.door
‘on the door’.

b. pa-zi-tseko
16.LOC-8.door
‘on the doors’.

c. mu-m’-tengo
18.LOC-3-tree
‘in the tree’

d. mu-mi-tengo
18.LOC-4-tree
‘in the trees’

e. ku-ci-pinda
17.LOC-7-room
‘in the interior of the room’

f. ku-zi-pinda
17.LOC-8-room
‘in the interior of the room’

There seems to be an agreement in the Bantu literature that the locatives in the above examples are syntactically derived by combining a locative element with the *phrasal* projection corresponding to the

base noun. In the following subsection, I discuss the arguments for this view presented in Bresnan and Mchombo (1995).

3.2.3.1 Bresnan and Mchombo (1995): Locatives as syntactic constructs

In this section, I show that Bresnan and Mchombo (1995) present a number of convincing arguments that locative prefixes in Chichewa are syntactically independent words which combine with the phrasal projection of the base noun.

As a starting point, I consider Bresnan and Mchombo's (1995: henceforth, B & M) arguments that locatives in Chichewa are nominal projections (NPs in their analysis). One of the reasons for this claim is that locative prefixes in Chichewa, like nouns, can occur in the subject position and control agreement on the verb as in (49a), or in the object position and follow the verb as in (49b).

49. a. *Pa* mu-dzi w-athu po-chitits-a chi-dwi
16 3-village 3-our 16.ASC INF-attract-FV 7-interest
pa-ma-sangalats-a a-lendo.
16.SM-PRS HAB-please-IND 2-visitor
'Our interesting village pleases visitors'.

b. A-lendo a-ma-pa-kond-a *pa* mu-dzi w-athu
2-vistor 2.SM-PRS HAB-16.OB-love-IND 16.3-village 3-our
p-o-chititsa chi-dwi.
16-ASC INF-attract 7-interest
'Visitors love it, our interesting village'.
(B & M, 1995:209-210)

Another argument for the nominal behaviour of locatives is that modifiers of locativised nouns can agree with modifiers and predicates in the noun class of the locative, as shown below:

50. a. pa mu-dzi p-athu p-onse
16 3-village 16-our 16-all
'at all our village'
(B & M, 1995:199)

b. Pa tebulo poyela pa-li madzi.
16 5.table 16.clean 16.be 6.madzi

'The clean top of the table is wet'.

(Carstens, 1997:385; Chichewa)

Example (50a) shows that the determiner *pathu* 'our' and the quantifier pronoun *ponse* 'all' agree with the locative noun *pa mudzi* 'at the village'. In (50b) the adjective *poyela* 'clean' modifies, and agrees with, the locative subject *pa tebulo* 'on the table'.

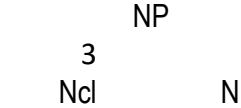
Finally, B & M (1995:211) note that locatives cannot be complements to nouns without an associative ('genitive') marker being inserted between the locative and the noun (see Myers, 1987; Bresnan, 1991).

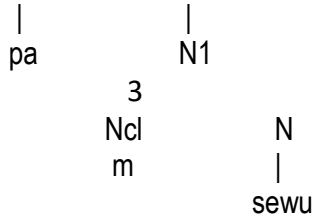
51.	a.	Mw-ana	wa	ku	mu-dzi	kw-athu
		1-child	1.ASS	17.LOC	3-village	17-our
		'a child from our village'.				
	b.	*mw-ana	ku	mu-dzi	kw-athu	
		1-child	17.LOC	3-village	17-our	
		'a child at our village'				

(51b) is ungrammatical because no associative marker links the head noun and the locative modifier. (51a) is grammatical because the associative marker intervenes between the head noun and the locative modifier. Since the function of associative markers is to assign Case to their complement (see below), the obligatoriness of the associative marker in (51b) shows that the locative phrase is a nominal projection (NP or DP).

Next, B & M provide evidence that Chichewa locatives are derived by combining the locative prefix, which they analyse as a locative noun, with the syntactic projection of the base noun. In this case, (52a) would be assigned the structure in (52b).

52. a. pa msewu
16.LOC 3-road
'on the road'

b.




The first argument in favour of treating locative prefixes as independent nouns comes from the conjoinability test. By using the conjoinability test, B & M (1995:206) show that locative prefixes are syntactically independent elements that can be conjoined as shown below.

53. Ndi-na-jambul-a zi-thunzi *mu* ndí *pa* ma-dengu.
1-REC.PST-draw-FV 8-picture 18.LOC and 16.LOC 6-basket
'I drew pictures in and on baskets'.

The example in (53) shows that the class 18 locative prefix *mu* and class 16 locative prefix *pa* can be conjoined. As only syntactically independent elements can be conjoined, (53) is an argument that the locative prefixes in Chichewa are syntactically independent elements.

Another conjoinability test involves complements of the locative prefixes. B & M show that NP complements of the locative prefixes can be conjoined under the same locative prefix in the same way that two NPs can be conjoined under one preposition (see B & M 1995: 205-206):

54. a. mu-na-chez-a ndi
2.PL/HON-REC PST-converse-FV with
[m-phunzitsi kapena m-sangalatsi]?
1-teacher or 1-entertainer
'Did you converse with the teacher or the entertainer?'

b. Mu-na-ik-a pa
2.PL/HON-REC PST-put-FV 16.LOC
[m-pando kapena m-tondo]
3-chair or 3-mortar
'Did you put (it) on the chair or the mortar?'

The example in (54a) shows that two NPs *mphunzitsi* 'teacher' and *msangalatsi* 'entertainer' are conjoined under the same preposition. Similarly, example (54b) above shows that the two NP

complements *mpando* ‘chair’ and *mtondo* ‘mortar’ can be conjoined under the same locative prefix *pa.* This further shows that locative prefixes, like prepositions, are syntactically independent elements.

Another argument by B & M that supports their claim that locative prefixes are syntactic elements in Chichewa is that the base noun following a locative prefix can be gapped just like nouns as shown in the example below:

55. a A-nyamata a-na-vin-a njerero
2-boy 2.SM-REC PST-dance-FV 9.name of dance
pa bwalo la mfumu Kapanga ndi pa
16.LOC 5.courtyard 5.ASS 9.chief Kapanga and 16.LOC
(~~bwalo~~) la mfumu Kapatuka.
(5.courtyard) 5.ASS 9.chief Kapatuka
‘The boys danced the njerero dance on chief Kapanga's courtyard and on chief Kapatuka's (courtyard)’.

b. A-nyamata a-na-low-ets-a nkhosa m’
2-boy 2.SM-REC PST-enter-CAUS-FV 10.sheep 18.LOC
khola la mfumu Kapanga kapena m’ (~~khola~~)
5.corral 5.ASS 9.chief Kapanga or 16.LOC (5.corral)
la mfumu Kapatuka.
5.ASS 1.chief Kapatuka’.
‘The boys drove the sheep either into chief Kapanga’s corral or into chief Kapatuka’s (corral)’.

c. Kodi a-na-kankh-ir-a m-pando ku chi-pinda
Q 2.SM-REC PST-push-APPL-FV 3-chair 17.LOC 7-room
cha ana kapena ku (~~chi-pinda~~) cha a-tsikana?
7.ASS 2.child or 17.LOC (7-room) 7ASS 2-girl
‘Did they push the chair to the children's room or (to) the girls’ room’?
(See B & M, 1995:207)

The fact that the nouns that combine with the locative prefixes can be gapped is further evidence that these nouns are syntactically represented as the heads of full phrasal complements of the locative prefixes.

The inbound anaphoric island test is another argument for the syntactic status of the locative prefix. B & M show that anaphoric and deictic pronouns can be used as phrasal NP complements to the locative prefixes. Example (56) is in Chichewa from B & M (1995:201).

56. a. mu iyi
18.LOC 9.this
'in this (e.g. house)'

b. pa icho
16.LOC 7.that
'on that'

c. ku iwo
17.LOC 6.them
'to them'

The examples above show that the deictic pronouns *iyi* 'this' in (56a), *icho* 'that' in (56b) and the anaphoric pronoun *iwo* 'them' can be used as NP complements of the locative prefixes within the phrase just as NPs are complements of the preposition within the prepositional phrase (PP) in English (e.g. I will give it *to them*). The fact that deictic and anaphoric pronouns can occur within the same phrase with locative prefixes and function as complements to the locative prefix, is evidence that locative prefixes are independent syntactic elements.

The final argument presented by B & M that I discuss here is based on the observation that locatives allow for what is called alternative concord. Alternative concord is concerned with the agreement relations between locatives and internal modifiers such as adjectives and possessives and refers to the fact that modifiers can agree with either the locative or the base noun (see also Myers, 1987; Carstens, 1991; Marten 2012; Zeller and Ngoboka, 2018), as illustrated by the Chichewa examples from B & M (1995:199) below.

57. a. *pa* mu-dzi *p*-athu *p*-onse
16 3-village 16-our 16-all
'at all of our village'

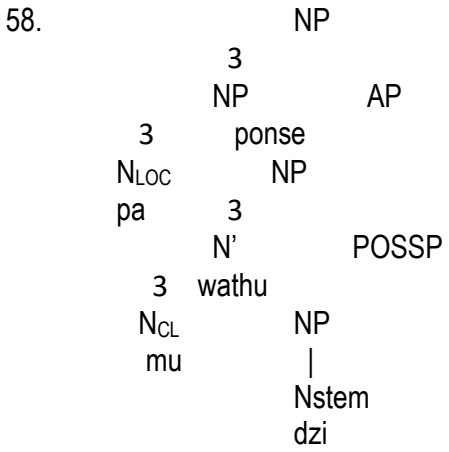
b. pa mu-dzi w-athu w-onse
16 3-village 3-our 3-all
'at all of our village'

c. pa mu-dzi w-athu p-onse
16 3-village 3-our 16-all
'at all of our village'

d. *pa mu-dzi p-athu w-onse
16 3-village 16-our 3-all
'at all of our village'

In (57a), both the possessive *pathu* 'our' and the quantifier adjective *ponse* 'all' agree with the noun class of the locative *pa* (locative concord) while in (57b), the possessive *wathu* 'our' and the quantifier adjective *wonse* 'all' agree with the noun class of the base noun (nominal concord). (57c) and (57d) show that when one modifier shows nominal, and the other one locative concord (mixed concord), the modifier showing nominal concord must precede the modifier showing locative concord.

B & M argue that the concord facts can be explained if it is assumed that a modifier showing nominal concord attaches to the projection of the base noun, while a modifier showing locative concord attaches to the projection of the locative noun (see e.g. B & M, 1995; Carstens, 1997; Marten, 2012; Zeller and Ngoboka, 2018). For example, the structure for *pa mudzi wathu ponse* 'at all our village' in (58c) above, which shows mixed concord, can be presented as follows:

58.


As (58) shows, B & M (1995) and Bresnan (1994) analyse locative prefixes as nouns which select the phrasal projection of the base noun as their NP-complement. The locative noun (= the locative prefix) projects another NP, which explains why locatives share the distributional properties and syntactic behaviour of non-locative NPs. The topmost NP licenses locative agreement on NP-external

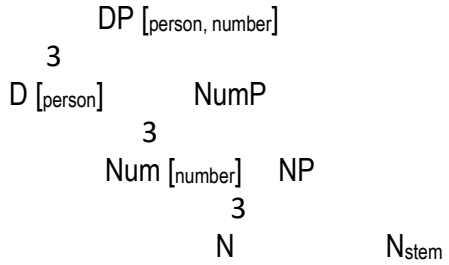
predicates, while concord with modifiers can reflect the noun class of the locative or of the base noun, depending on the attachment site of the modifier. Importantly, the possibility of nominal concord follows from the fact that the base noun projects and is represented by a full phrasal category inside the locative.

While the idea that locative NPs are headed by a nominal element that represents the locative meaning and noun class specification is widely accepted in the Bantu literature (Myers, 1987; Bresnan and Kanerva, 1989; Bresnan, 1994; Marten, 2012; Caha and Pantcheva, 2015; Zeller, 2017a), B & M's claim that this nominal head is realised by the locative prefixes is not uncontroversial. Carstens (1997) argues for a different function of locative prefixes. I discuss her theory in the next subsection.

3.2.3.2 Carstens (1997): Locative prefixes as Case-markers.

Like B & M (1995), Carstens (1997) analyses locatives as nominal projections (DPs in her theory) (see also Carstens, 2008; 2011). As shown above, according to Carstens, the DP structure in Bantu languages consists of an N-head (i.e. a lexical noun head), and two functional heads namely: the determiner (D) and Num(ber) (#) head. (As discussed above, N can be further decomposed into a little n head and an acategorial root):

59.

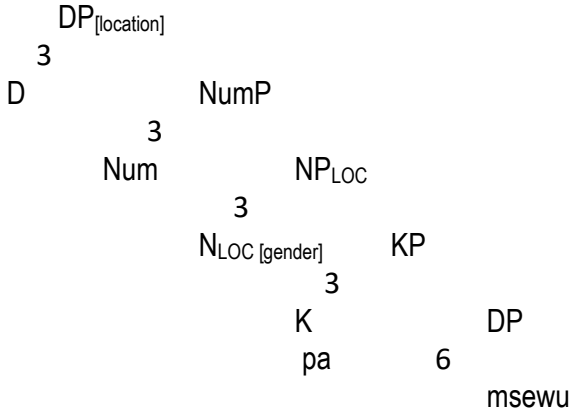


Since locatives are based on the same structure, they are analysed as locative determiner phrases (DP_{LOC}) by Carstens (1997). The locative noun, of which DP_{LOC} is the extended projection, bears locative noun class features and determines the noun class of the DP_{LOC} via head movement to D_{LOC} via Num (see Zeller & Ngoboka 2018). In Carstens' theory, the locative gender is realised on the locative noun, but adopting Fuchs & Van der Wal's theory, the locative noun can also be analysed as a locative root which combines with one of three little n-heads, which contributes to the respective locative noun class 16, 17, or 18.

In B & M's (1995) theory, the locative noun is overtly realised by a locative NC prefix. Carstens (2008, 1997) disagrees with this view, based on a problem she identifies with the structure in (58). In this

structure, the locative noun takes the NP-projection of the base noun as its complement. However, Carstens notes that the locative noun in (58) does not trigger 'of' insertion before the complement NP, which in Bantu would correspond to the insertion of an associative marker. This is unexpected – since nouns do not assign Case, the nominal complement of a noun usually has to be marked using the associative. As Carstens notes, treating locative prefixes as nouns that select an NP as a complement would imply that the associative marker appears between the locative prefix and the base noun. Since this is not the case, Carstens rejects the analysis of locative prefixes as locative nouns and instead analyses the locative prefixes themselves as Case markers which license the phrasal projection of the base noun and link this projection to the locative noun. In Carstens' analysis, the locative noun which selects the projection of the base noun is a silent nominal head of the locative DP (see Carstens, 1997). The locative prefixes like NC prefixes are gender-specific case markers realising the highest functional projection of the base noun (a KP = KaseP), but it is the gender (noun class) of the silent locative which determines the type of case marker:

60.



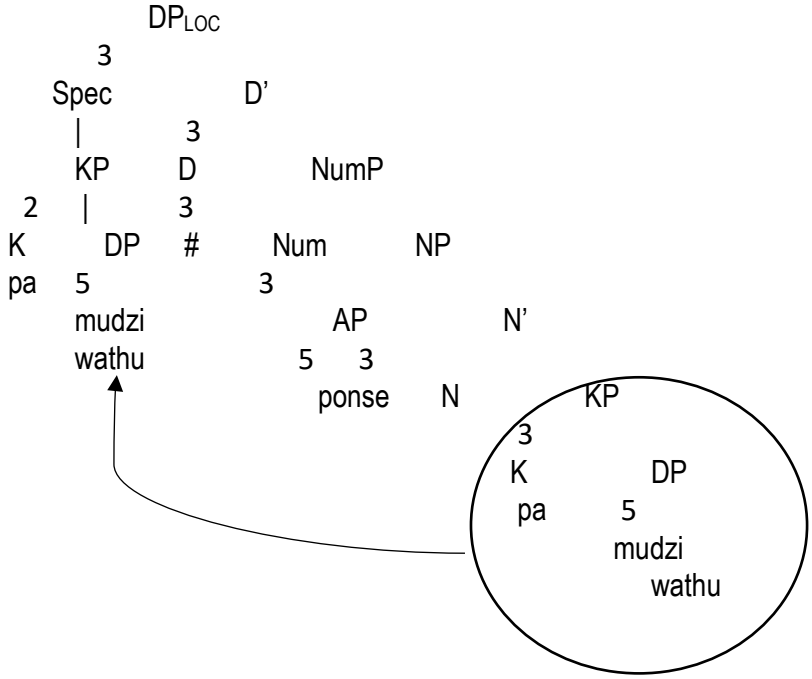
According to Carstens (1997), each locative noun class prefix acts as a specific case marker (K), selected by one of the silent 'place' nouns N_{LOC} of distinctive gender. As a result, the locative prefix is analysed as a case-marker (i.e. the head of the functional projection, Kase), and therefore must be analysed as an inflectional element.

In the remainder of this chapter, I adopt Carstens' (1997) analysis shown in (60) above, with the additional assumption that the silent locative noun is itself a root whose categorial features and locative noun class are determined by a little n which selects it (cf. Fuchs & Van de Wal 2018). However, for completeness, I end this section by illustrating one more aspect of Carstens' (1997) analysis. Recall that with DP-internal modifiers, locatives can show mixed concord, with one modifier showing locative concord and the other one nominal concord. (61) repeats example (57c) from above:

61. a. pa mu-dzi w-athu p-onse
16.LOC 3-village 9-our 16-all
'at all our village'

In Carstens' analysis, modifiers are located in left-branching specifiers of NPs, and N-to-D movement derives the order N > Mod. In order to explain that with mixed concord, the modifier showing nominal concord must precede the modifier showing locative concord. Carstens proposes that the whole locative KP raises to SpecDP_{LOC} suggesting the structure of a locative DP in (61b).

b.



According to Carstens (1997, 2008, 2011), both the N-head of the lower NP *mudzi* and the silent locative N move to D/D_{LOC} via Num. This movement is triggered because D has strong N-features to attract N to move to D (Carstens, 1997). The modifier *wathu*, located in the specifier of the base noun's NP, now follows *mudzi*. Since KP-raising is a requirement in $DP_{LOC,}$ the lower DP *mudzi wathu* raises together with the preceding Case marker, the locative prefix *pa*, i.e. the whole KP moves because D seems to have an EPP feature which attracts the KP to move to its Specifier. As a result, the modifier *ponse*, which is located in the specifier of the locative NP, now follows the KP *pa mudzi wathu*, deriving the word order in (61).

In sum, Carstens (1997) argues that locatives in Chichewa are not nouns but case markers because they fail to trigger 'of' insertion before complements. They are therefore clearly inflectional and represented in the syntax by a functional category.

3.3 The syntax of VN-compounds in Cinyanja

3.3.1 VN-compounds and synthetic compounds

As indicated earlier, VN-compounds in Cinyanja resemble English synthetic compounds of the *truck-driver* structure. In Chapter 2, I discussed the two main theories concerning the structure of synthetic compounds. One theory postulates that the compounds are built on the basis of a combination of the verb and its argument noun [N+V] to which an affix *–er* is attached (Botha, 1980; Fabb, 1984; Ackema and Neeleman, 2004; Booij, 2009:214; Harley, 2009). The other theory postulates that synthetic compounds are built on the basis of the noun [N] combining with the nominalised verb [V-er] (Selkirk, 1981; Scalise et al, 2009; Lieber, 2009c:367). In principle, these two types of analyses may also be available to account for the VN-compounds in Cinyanja. Moreover, presently, not much literature is available on phrasal compounds in Bantu languages and Cinyanja in particular.

However, a complication arises for the Type 2 and Type 3 VN-compounds. Recall that these types of VN-compounds in Cinyanja consist of a verb followed by a nominal part which includes a NC prefix (Type 2) or a nominal part with a locative (Type 3) as illustrated in the examples below:

62. a. m-pala-ma-tabwa
1-scrape-6-planks
'plank scraper'

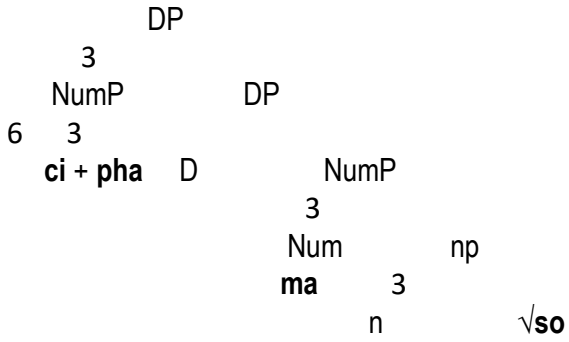
b. ci-ponda-mu-thengo
7-step-FV-18.LOC-bush
'consultation fee'

In (62) the class 6 prefix *ma-* and class 18 prefix *mu-* are part of the nominal part of the compound. As argued above, NC prefixes are realisations of a functional Num or n-category and realise number and grammatical gender of the noun *tabwa* 'plank'; the locative prefix is a Case marker of category K that licenses the DP *thengo* 'bush'. Correspondingly, Carstens (1993; 1997; 2008), Diercks (2010) and Buell, et al (2011) analyse these noun class prefixes as inflectional markers.

In this section, I show that internally, VN-compounds must have a phrasal structure because the prefixes they include (the noun class markers and locative prefixes) are inflectional. In the analysis of the internal structure of Cinyanja VN-compounds, henceforth, I adopt the representation of "ordinary" and locative noun phrases as determiner phrases (DP) and locative determiner phrases (DP_{LOC}) respectively. I argue that the VN complex expressions of Type 2 and Type 3 are based on a syntactic combination of V + DP (with "V" standing for a root and a verbalising little v), which is a VP (or vP) (Di Sciullo and William, 1987; Lieber, 1992; Mchombo, 2004). These VPs are eventually nominalised by the little n to form VN-compounds. In this case, the nominaliser attaches to the VP to form the nominal compound.

This approach is opposed to the alternative analysis of synthetic compounds as including nominalised verbs which combine with nouns (see Lieber, 2009c; Scalise et al, 2009). Given the analysis of NC-prefixes adopted here, this would involve the merger of a phrase corresponding to the nominalised verb with the DP corresponding to the verb's argument:

63.



The tree structure in (63) shows that the verb is nominalised first due to the presence of the NC prefix and this nominalised structure must at least be a NumP. This NumP would somehow combine with the DP to form a compound word *ciphamaso*. The structure in (63) is a merger of the NumP and a DP. To the best of my knowledge, this is not a structure attested anywhere else in the grammar. Furthermore, an analysis in the spirit of (63) is implausible for a Cinyanja compound like *ciphamaso*, because *cipha* by itself does not exist as a noun. In fact, in Cinyanja, the first part of the compounds (i.e., the nominalised verb), does not form an independent word with the meaning equivalent to that of the literal meaning as shown in Table 3.1 below.

Table 3.1: Nominalised verbs in VNs are not independent nouns

Compound Word	**Meaning**	**Nominalised Verb**
cipha-maso	hypocrite	*ci-pha

ka-penta-milomo	prostitute	*ka-penta
ci-dya-makanda	sugar daddy	*ci-dya
m-konda-maungu	pumpkin lover	*m-konda
m-pala-matabwa	plank scraper	*m-pala
ci-gona-mubawa	drunk	*ci-gona
m-khala-pa-mwala	stone sitter (idle person)	*m-khala
m-lowa-mu-malo	pronoun	*m-lowa
ci-ponda-mu-thengo	consultation fee	*ci-ponda

The Table in 3.1 shows that the first parts of the compounds, presented as nominalised verbs do not exist as independent nouns by themselves, and that they are devoid of independent meaning. For example, *mkhala, mkonda, mpala* and others in the table, do not exist as separate words in Cinyanja.

I therefore reject an analysis along the lines of (63) for the VN compound words in Cinyanja, and argue instead that Type 2 and Type 3 VN-compounds in Cinyanja are formed by nominalising a VP/vP consisting of a verb and a DP complement.

3.3.2 The vP structure inside VN-compounds

I now consider the question whether Cinyanja VN-compounds such as *cigonamubawa* can be derived via incorporation of the noun (e.g. *mubawa*) into the verbal part (e.g. *-gona*). This incorporation would involve head movement of the noun from inside the DP-complement of the verb to the verb position. Such an analysis based on head movement is in fact suggested in Harley (2009) for synthetic compounds in English. Harley's account is formulated in the framework of Distributed Morphology (DM) (see Chapter 1).

In this section, I focus mainly on Type 2 and Type 3 VN-compounds. Recall that Type 2 VN-compounds in Cinyanja comprise the verb and a noun object (DP) which is the complement of the verb as shown in Table 3.2.

Table 3.2: VN-compounds with the internal structure of VO word order

S/N	COMPOUND	VERB PHRASE	TRANSLATION

a.	ci-pha-ma-so	pha maso	kill the eyes
	hypocrite	V O	
b.	ka-penta-milomo	penta milomo	paint the lips
	prostitute	V O	
c.	m-konda-maungu	konda maungu	love pumpkins
	pumpkin lover	V O	
d.	m-pala-matabwa	pala matabwa	scrape the planks
	plank scrape/carpenter	V O	

The examples in Table 3.2 show that the order of constituents internal to VN-compounds resembles the order of constituents in a syntactic VP structure where the verb is the head of the VP.

Similarly, Type 3 VN-compounds consist of a verb followed by a locative in the nominal part of the compound. The locative noun phrases (i.e. DP_{LOC}) are complement or internal arguments of the verb. Internally, these VNs resemble the vP phrasal structure as shown in Table 3.3.

Table 3.3: Phrasal structure of Type 3 VN-compounds

S/N	COMPOUND	VERB PHRASE	TRANSLATION
a.	ci-gona-mu-bawa	gona mu bawa	sleep in the bar
	drunk	V LOC DP	
b.	ci-ponda-mu-thengo	ponda mu thengo	step in the bush
	consultation fee	V LOC DP	
c.	m-khala-pa-mwala	khala pa mwala	sit on the rock
	stone sitter	V LOC DP	
d.	m-lowa-mu-malo	lowa mu malo	enter in place of
	pronoun	V LOC DP	
e.	m-sesa-pa-khomo	sesa pa khomo	sweep the yard
	maid	V LOC DP	

Internally, the VN-compounds cited in Tables 3.2 and 3.3 above seem to comprise a VP word order. According to Miti (2001), VPs in most Bantu languages normally comprise an obligatory verb which may be followed by a noun phrase, a locative or an adverbial phrase. Therefore, I claim that the Cinyanja VN-compounds above are based on the phrasal structure of a V + DP, that is, they have a complete VP-structure inside them. Evidence for this view is that the canonical VO word order and the

verb + locative or adverbial phrase word order is preserved in these compound words as shown in Tables 3.2 and 3.3. Further, the verbs in these compounds all end in a vowel *–a.* This seems to suggest that the verbs comprise the root and the verbaliser *–a* (see below).

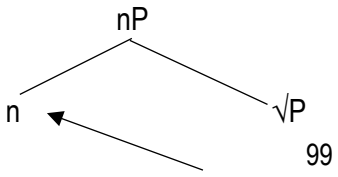
Recall that in DM, roots are considered acategorial and only realise their category when they merge with a category-creating feature bundle, represented by n, a, or v, which may also be realised by affixes (see also the discussion of Fuchs & Van der Wal's (2018) analysis of gender presented above). For example, the root √*STONE* can be a verb or a noun depending on which category-defining feature they merge with as shown in the examples below:

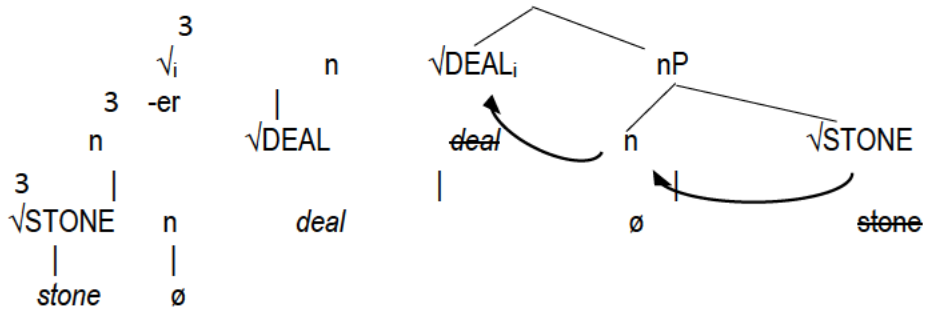
64. a. The women removed a big *stone* from the road.
    b. The women *stoned* the thief to death.

In (64a) *stone* is a noun because it appears in the context of a determiner while in (64b) *stone* is a verb because it inflects for tense and appears with a past tense morpheme, which is a functional abstract morpheme.

According to Harley (2009), a compound word such as *stone dealer* is formed when the noun *stone* incorporates into the verb *deal* and the complex head merges with a categorising head. Given the decomposition of nouns and verbs into roots and little n or v, the analysis involves a root √STONE which incorporates into a nominalising affix n**.** The complex n then incorporates into the root √DEAL. Notably, this root is not verbalised in Harley's theory; rather, the complex root phrase merges with another nominaliser n, which is pronounced as the derivational suffix *–er*. The complex head consisting of the root √DEAL and the nP *stone* then incorporates into this head to derive *stone dealer*, resulting in the structure below (see Harley, 2009:136):

65.





In (65) above, the nP *stone* acts as the internal argument of the verb *deal* which through vocabulary insertion is realised as a synthetic compound word *stone-dealer*. According to Harley, the incorporation of the complex n *stone* is motivated by the need of this noun to check a case feature.

Importantly, Harley argues that if the argument of √DEAL was a DP such as *the stone* or *stones,* incorporation would have been blocked, and the DP argument would remain to the right of the head noun *dealer*. Harley claims that what blocks incorporation out of complex structures with functional elements, to derive, for example, **stones-dealer* or **the-stone-dealer*, is that the functional heads Num and D check the case features of the nP internally, and as a result, a complex head derived by n-to-Num-to-D-movement has no internal case feature to check and therefore becomes unavailable for further incorporation into a higher root (see Padrosa-Trias, 2010). This implies rather that incorporation occurs in synthetic compounds only when the complement of the (verbal) root is an nP, and not a DP. However, as argued above, Cinyanja VN-compounds of Type 2 and 3 include a full DP, because the NC and locative prefixes are associated with functional structure in a DP. This means that there cannot be an incorporation of the nominal part of the compound into the verbal part in the derivation of VN compounds in Cinyanja.

Another critical observation regarding VN-compounds in Cinyanja is that all the verbs in the compounds end in the vowel *-a* (e.g. *gon-**a**, -low-**a***, etc.). This vowel can be analysed as a verbaliser, i.e. a head that turns an acategorial root into a verb. This view is borne out of the observation that most roots end in the suffix ***–a*** when verbalised and in the suffix ***–i,*** or ***-o*** when nominalised as shown in the examples in (66).

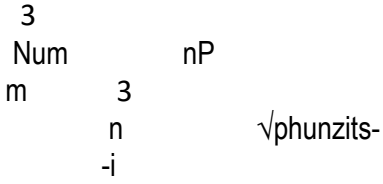
66.	**Root**	**Verb**	**Translation**	**Noun**	**Translation**
	kond-	kond-a	love	ci-kond-i	love

phunzits-	phuzits-a	teach	m-phunzits-i	teacher
low-	low-a	substitute	m-low-i	substitution
phikitsan-	phikitsan-a	compete	m-pikitsan-o	competition
phunzir-	phunzir-a	learn	phunziro	lesson
kuw-	kuw-a	shout	m-kuw-o	shout
fanizir-	fanizir-a	compare	ci-fanizir-o	comparison
funs-	funs-a	ask	funs-o	question
ganiz-	ganiz-a	think	ci-ganizo	thought
tsek-	tsek-a	close	ci-tsek-o	door.

The examples in (66) show that the roots are unspecified and only become specified when a categorising suffix is added to the terminal node (Marantz, 1997; Harley 2009, 2014; Borer, 2005; 2009; Acquaviva, 2014). For example, a root *phunzits-* has the semantic content of *teach*, and becomes a verb when the *–a* is suffixed as shown in (67a), and a noun when *-i* is suffixed and a little n is added, and NumP to nominalise it as in (67b).

67. a. phunzits-a

b. m-phunzits-i



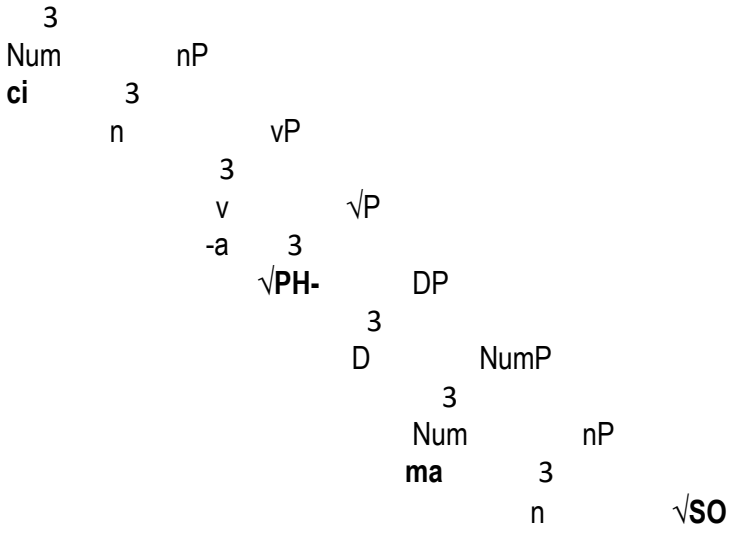
The examples in (67a) and (67b) clearly show that the suffix changes to *–a* when the root is a verb and to *–i* when it is a noun. This is one of the core assumptions of DM where roots are considered category neutral and devoid of grammatical features (Marantz, 1997; Harley, 2009; 2014). Note that the analysis of n as the location of the derivational suffix leads me to adopt the view that NC prefixes spell out Num, and not n (see Fuchs & Van der Wal 2018), as mentioned in section 3.2.2 above. With regards to the VN-compounds, (refer to Table 3.1) all the verbal roots in these compounds end in the suffix *–a.* The fact that the V-part of VN-compounds in Cinyanja seem to always end in *–a*, suggests

that there is a little v that verbalises the root and derives a verb. In contrast, in synthetic compounds in English, the "verbal" part is only represented by a root (see Harley's structure in (65)).

The view that the suffix *-a* is a reflex of a functional category v that categorises the root suggests that these verbal roots cannot be nominalised unless they combine with the object complement, and then the entire structure is verbalised. Considering the examples above, I argue that the relevant part of a compound can only be nominalised when it is a vP. VN-compounds in Cinyanja are not formed by a nominalisation of a root or a “morpho-verbalised” form of a root and the argument. Rather, I claim that the verbal root is verbalised after it has combined with its complement, and then the entire phrasal structure is nominalised. This suggests that the compounds comprise a nominalised VP.

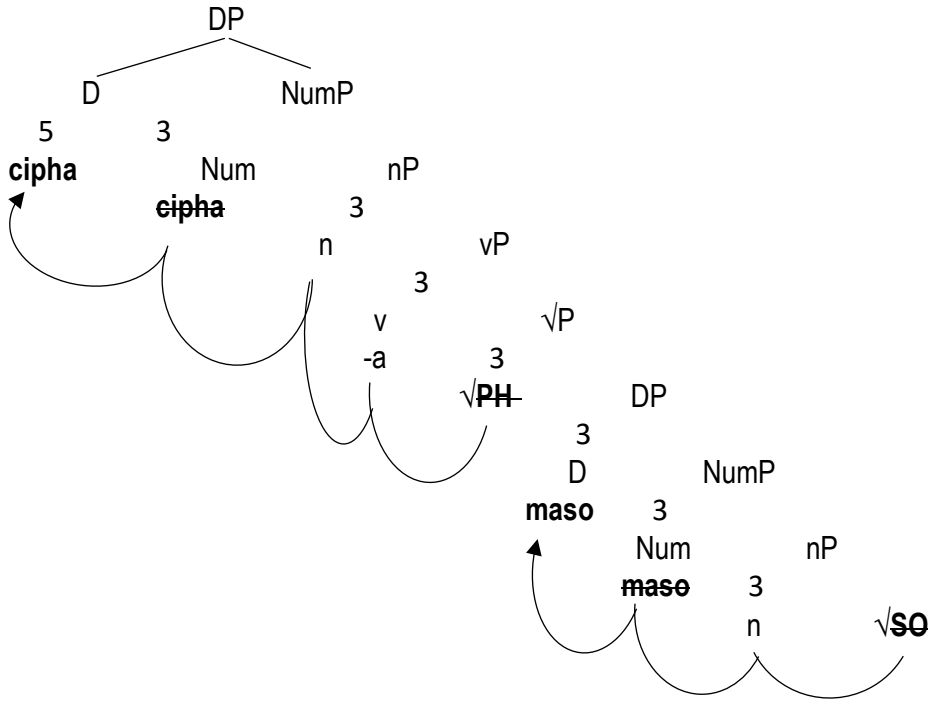
Based on these considerations, I now provide a first tentative analysis of the structure of VN-compounds in Cinyanja. (Note that some aspects of this analysis will be revised below.) In a compound word like *ciphamaso* ‘hypocrite’, the nominal part *maso* is a full DP in which the NC prefix realises the category Num (following Carstens 1993, 2008). This DP merges with the root *ph* -, which is then verbalised by the little v (pronounced by the suffix *-a*), deriving a vP. This vP then merges with the gender-determining little n, and nP merges with number inflection in Num, which is spelled out by the NC 7 prefix. Finally, D is added to yield a full-fledged DP, as shown below:

68. DP
3
D NumP



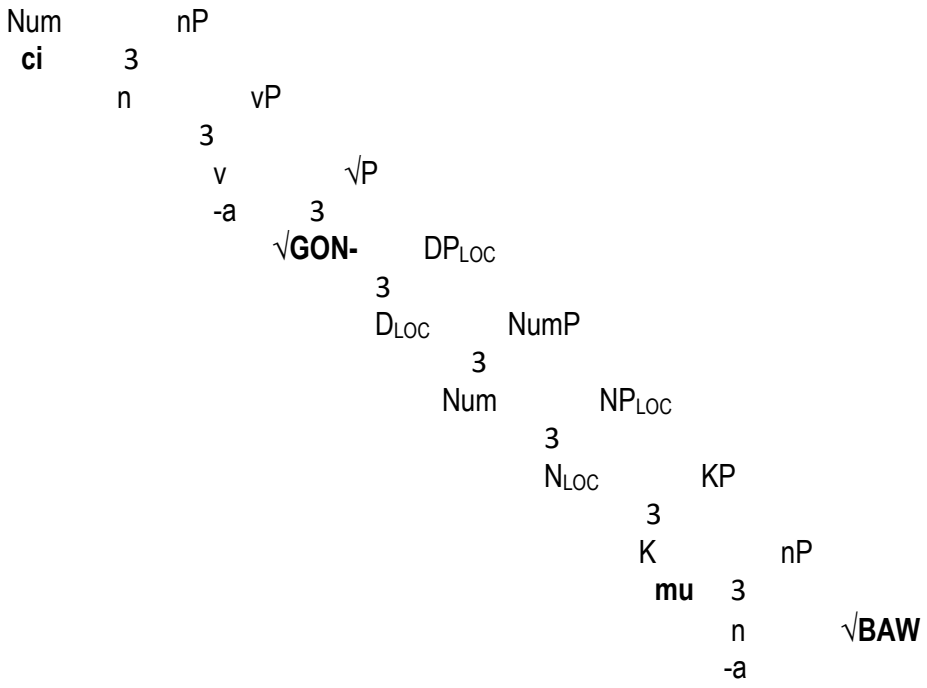
The compound expression *ciphamaso* is derived when both roots undergo head movement to the D-nodes of their projections as shown below.

69.



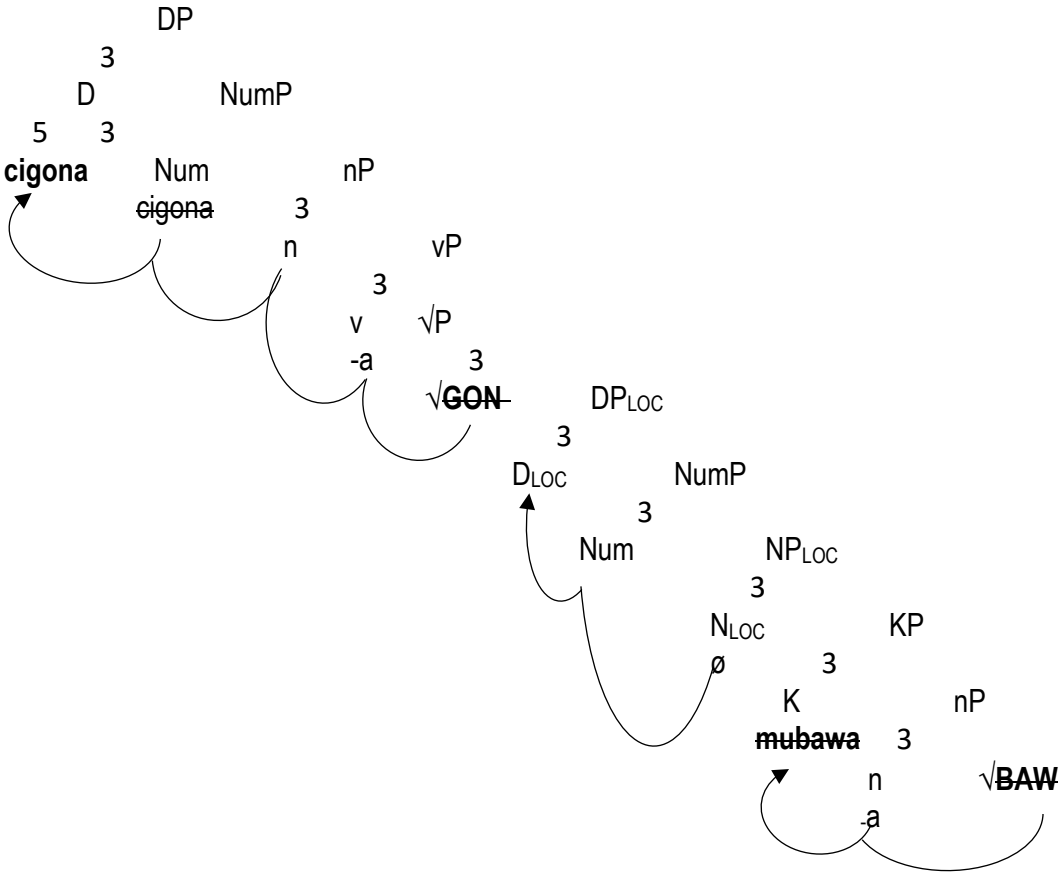
Similarly, in a VP structure of Type 3, VN-compounds such as *cigonamubawa* 'drunk', the DP_{LOC} *mubawa* merges with the verb root √*GONA* to realise the root phrase which merges with the little v, and then the little n is added to vP, followed by the number inflection, reflecting the NC 7 prefix. Finally, D is added to realise the VN compound, as shown below.

70.



After head movement, the structure looks as follows:

71



The structures in (68) and (70) capture a number of aspects of VN-compounds discussed in this section. NC and locative prefixes are functional heads, and verbalising and nominalising heads derive nPs and vPs from unspecified roots. In the next section, I demonstrate that VN-compounds do indeed share the distributional properties and syntactic behaviour of DPs in Cinyanja. However, section 3.5

shows that the internal structure of these compounds is not accessible for syntactic rules and is syntactically opaque. As I demonstrate below, this means that the structures in (68) and (70) have to be revised to capture the fact that the phrasal VN-compounds in Cinyanja, although including phrasal structure, still behave like simple nouns.

3.4 VN-compounds in Cinyanja as DPs

In this section, I discuss VN-compounds in their syntactic context and show that they appear in the same syntactic contexts as other DPs. VN-compounds can appear in the subject and object positions in simple sentences, and in relative clauses. The entire compound word can also be modified by an adjective, and be coordinated with a simple noun, suggesting that VN-compounds are indeed represented by the DP-structures suggested above.

3.4.1 VN-compounds as subjects and objects of the verb

In a basic clause structure, VN-compounds can function as subjects of a verb. Consequently, they typically agree in the class indicated by the first prefix as shown in the examples below:

72. a. **Ci-ph-a-dzuwa** c-a-mang-a nyumba.
7-kill-FV-5.sun 7.SM-PRF-build-FV 9.house
'The beautiful girl has built a house'.

b. **M-pal-a-ma-tabwa** a-na-gulits-a mi-pando. .
1-scrape-FV-6-plank 1.SM-PST-sell-FV 4-chair
'The carpenter sold the chairs'.

c. **Ka-pent-a-mi-lomo** w-a-gul-a nyumba.
1-paint-FV-4-lip 1.SM-PRF-buy-FV 9.house
'The prostitute has bought a house'.

d. **Ci-gon-a-mu-bawa** c-a-gul-a galimoto.
7-sleep-FV-18.LOC-5.bar 7.SM-PRF-buy-FV 5.car
'The drunk has bought a car'.

In (72a) the subject is *ciphadzuwa* and the initial prefix *ci-* of NC 7 is reflected in the verb by the class 7 subject agreement marker. Similarly, in (72b), *mpalamatabwa* is the subject of the clause. The compound controls agreement on the verb which is indicated by the NC 1 prefix *m-.* The other example (72c) shows that the compound word *kapentamilomo* 'prostitute' is a subject noun and agrees in the noun class of animates indicated by the NC 1 prefix. In (72d) *cigonamubawa* is the subject of the sentence and the initial class 7 prefix *ci-* is reflected in the verb as an agreement marker. The fact that these compounds can act as subjects, is consistent with their analysis as DPs presented in the preceding section.

It is interesting to note that some phrasal compounds in class 7 can also trigger agreement in class 1, as illustrated in examples (72a') and (72d') repeated hereunder:

72 a'. **Ci-ph-a-dzuwa** (S) w-a-mang-a (V) nyumba. (O)
7-kill-FV-5.sun 1.SM-PRF-build-FV 9.house
'The beautiful girl has built a house'.

d'. **Ci-gon-a-mu-bawa** (S) w-a-gul-a (V) galimoto. (O)
7-sleep-FV-18.LOC-5.bar 1.SM-PRF-buy-FV 5.car
'The drunk has bought a car'.

The details of the agreement ambiguity for class 7 phrasal compounds are discussed in section 3.6.2.

Another indication that VN-compounds are DPs is that they can function as nominal objects as shown in (73a-c); *kapentamilomo*, *mpalamatabwa,* and. *cigonamubawa.*

73. a. A-nyamata a-na-pirikits-a **ka-pent-a-mi-lomo.**
2-boy 2.SM-PST-chase-FV 1-paint-FV-4-lip
'Boys chased the prostitute'.

b. M-phunzitsi w-a-on-a **m-pal-a-ma-tabwa.**
1-teacher 1.SM-PRF-see-FV 1-scrape-FV-6-plank
'The teacher has seen the carpenter'.

c. M-silikali w-a-mang-a **ci-gon-a-mu-bawa.**
1-policeman 1.SM-PRF-arrest-FV 7-sleep-FV-18.LOC-5.bar
'A policeman has arrested the drunk'.

Noteworthy is that sometimes an object marker (OM) can sometimes be prefixed to the verb in Cinyanja. The inclusion of the OM in the verbal morphology allows right or left dislocation of the compound expressions in the clause (see Bresnan & Mchombo, 1987). To demonstrate this, example (73c) is repeated in (74).

74. a. M-silikali w-a-ci-mang-a mwankhanza
1-policeman 1.SM.PRF7.OM-arrest-FV harshly
ci-gon-a-mu-bawa.
7-sleep-FV-18.LOC-5.bar
'A policeman has arrested (him) harshly the drunk'.

b. **Ci-gon-a-mu-bawa,** m-silikali, w-a-ci-mang-a
7-sleep-FV-18.LOC-5.bar 1-policeman 1.SM.PRF-7.OM-arrest-FV
mwankhanza.
harshly
'The drunk, the policeman, has arrested him harshly'.

c. **Ci-gon-a-mu-bawa**, w-a-ci-mang-a mwankhanza
7-sleep-FV-18.LOC-5.bar 1.SM.PRF-7.OM-arrest-FV harshly
m-silikali.
1-policeman
'The drunk, the policeman has arrested him harshly.'

The examples above show that the compound word *cigonamubawa* 'drunk' can be dislocated just like other DPs. In (74a) the compound word has been right dislocated. This is opposed to (74b & c) where the compound words have been left dislocated. The examples in (74) demonstrate that compound words can be dislocated just like other "regular" DPs as long as the OM is prefixed to the verb. However, when the OM is absent in the verbal morphology, left dislocation of the compounds is impossible as shown in (75).

75. a. * **Ci-gon-a-mu-bawa**, mwankhanza, m-silikali
7-sleep-FV-18.LOC-5.bar harshly 1-policeman
w-a-mang-a.
1.SM.PRF-arrest-FV

*'The drunk, harshly, the policeman, has arrested'.

b. ***Ci-gon-a-mu-bawa***, m-silikali, w-a-mang-a
7-sleep-FV-18.LOC-5.bar 1-policeman 1.SM.PRF-arrest-FV
Mwankhanza.
harshly.
*'The drunk, the policeman, has arrested harshly'.

The examples above show that VN-compounds cannot be left dislocated when the OM is absent in the verb. Therefore, we can claim that the presence of the OM in the verbal morphology triggers dislocation of compound words in a clause in Cinyanja.

3.4.2 VN complex expression as subject or object of a relative clause

VN complex expressions can also function as head nouns of a relative clause in which they correspond to a nominal subject or object. For better clarity, *cigonamubawa* is the object of the clause in (76a) that is used as a base example from which the relative clauses in (76b) and (76c) are drawn. In (76b), the VN compound *cigonamubawa* 'drunk' is the subject of the (passivised) relative clause; in (76c), *cigonamubawa* is the object of the relative clause.

76. a. Mw-ana a-na-on-a **ci-gon-a-mu-bawa**
1-child 1.SM-PST-see-FV 7-sleep-FV-18.LOC-bar
m-nyumba.
18.LOC-house
'The child saw the drunk in the house'.

b. **Ci-gon-a-m-bawa** ci-mene ci-na-wonedw-a ndi
7-sleep-FV-18.LOC-5.bar 7-REL 7-PST-see-PASS-FV by
mw-ana mu nyumba, ci-na-kwatir-a ci-ph-a-dzuwa.
1-child 18.LOC 9.house 7.SM-PST-marry-FV 7-kill-FV-sun
'The drunk who was seen by the child in the house married a beautiful girl'.

c. **Ci-gon-a-mu-bawa** ci-mene mw-ana
7-sleep -FV-18.LOC-5.bar 7-REL 1-child
a-na-on-a m'nyumba, ci-na-kwatir-a ci-ph-a-dzuwa.

1-PST-see-FV 18.LOC-house 7.SM-PST-marry-FV 7-kill-FV-sun
'The drunk whom the child saw in the house, married a beautiful girl'.

3.4.3 Coordination

VN-compounds are also considered DPs because they can be coordinated with other DPs as illustrated in (77a-b) or objects as in (77c-d).

77. a. **Ci-gon-a-mu-bawa** ndi gogo a-a-pit-a
7-sleep-FV-18.LOC-5 and 1a.grandparent 2.PRF-go-FV
ku-dambo.
17.LOC-5.swamp
'The drunk and grandfather have gone to the swamp'.

b. **Ci-ph-a-ma-so** ndi m-phunzitsi a-a-thaw-a.
7-kill-FV-6-eye and 1-teacher 2-PRF-run-FV
'The hypocrite and the teacher have run away'.

c. M-sangalatsi kapena **m-pal-a-ma-tabwa** w-a-gul-a
1-entertainer or 1-scrape-FV-6-plank 1.SM-PRF-buy-FV
m-thuthuthu.
3-motorbike.
'The entertainer or the carpenter has bought a motor-bike'.

d. M-lendo kapena **mu-ses-a-pa-khomo** w-a-b-a
1-visitor or 1-sweep-FV-5.yard 1.SM-PRF-steal-FV
ndalama.
9.money
'The visitor or the maid has stolen the money'.

From the examples in (77)[6], it is clear that VN-compounds can be conjoined with other DPs.

[6]The examples in (77) show that when a coordinator *ndi* 'and' conjoins a class 7 compound with a class 1 noun, the verb agrees in the plural of class 2. But when a compound word and a simple noun are conjoined by a coordinative word *kapena* 'or', the verb agrees in class 1 (see Corbett and Mtenji,1987:31; Van de Velde, 2006; Zeller, 2008)

3.4.4 Modification

3.4.4.1 Adjectives

VN-compounds, like other nouns, can also be modified with adjectives. In this case, modification affects the entire compound word, not only the nominal constituent inside the compound. For example, in (78a) the adjective *wamtali* 'tall' modifies the whole compound word *kapentamilomo* 'prostitute' as demonstrated by the fact that the adjective agrees with the whole compound in class 1. Similarly, in (78b) *canzeru* 'wise' carries the class 7 concord and modifies the compound word *cigonamubawa* 'drunk'. In contrast, as examples (78a') and (78b') show, when the adjectival concords reflect the noun class of the compound-internal nominal, the constructions become ungrammatical, which is evidence that an adjective cannot modify only the nominal part of the expressions. I return to this point in section 3.5.2 below:

78. a. Ka-pent-a-mi-lomo w-a-m-tali w-a-b-a ndalama.
1a-paint-FV-4-lip 1-ASS-1-tall 1.SM-PRF-steal-FV 9.money
'The tall prostitute has stolen money'.

a'. * Ka-penta-mi-lomo y-a-i-tali y-a-tup-a.
1a-paint-FV-4-lip 4-ASS-4-tall 4.SM-COP-swell-FV
Intended: 'The prostitute's long lips are swollen'.

b. Ci-gon-a-mu-bawa c-a-nzeru c-a-gul-a
7-sleep-FV-18.LOC-5.bar 7-ASS-wise 7-SM.PRF-buy-FV
nyumba.
9.house
'A wise drunk has bought a house'.

b'. * Ci-gon-a-m-bawa mw-a-dothi mu-li nchenche.
7-sleep-FV-18.LOC-5.bar 7.ASS-dirty 7-COP 10.fly
Intended: 'The drunk's dirty bar has flies'.

3.4.4.2 Possessor modifiers

The compounds can also take possessor modifiers just like non-compound nouns. In Cinyanja, possessor modifiers canonically are post-nominal modifiers, and they agree in class with the noun class of the nouns they modify.

79. a. Ci-pond-a-mu-thengo c-anu c-a-cepa.
7-step-FV-18.LOC-5.bush 7-your 7.SM-BE.PRS-less-FV
'Your consultation fee is less'.

a'. *Ci-pond-a-mu-thengo l-anu l-a-psy-a.
7-step-FV-18.LOC-5.bush 5-yours 5-PRF-burn-FV

b. M-pal-a-ma-tabwa w-athu w-a-thaw-a.
1-scrape-FV-6-plank 1-our 1.SM-PRF-run-FV
'Our carpenter has run away'.

b'. *M-pal-a-ma-tabwa a-thu a-a-psy-a
1-scrape-FV-6-plank 6-our 6.-BE.PRS-burn-FV

In the examples above, the possessor agrees with the compound in the compound's respective class. In (79a & b), the possessor *canu* 'your' and *wathu* 'our', agree in class with the compound word *cipondamthengo* 'consultation fee' and *mapalamatabw*a 'carpenter' while in (79a' and b') the possessor *lanu* 'yours' and *athu* 'ours' reflect the noun class of the compound internal nominal and the construction is ungrammatical. This is evidence that the possessor cannot modify only the nominal part of the compound.

3.4.4.3 Demonstratives

Demonstratives/deictic pronouns can also modify compound words. These pronouns agree in class with the noun they modify. The pronominal prefix of the deictic pronouns is the same as the nominal class subject marker (Miti, 2001; 2006) as shown in the example below.

80. a. M-dob-a-ci-dzala u-yo w-a-tay-a ndalama.
1-collect-FV-7-garbage 1-that 1.SM-PRF-drop-FV 9.money
'That garbage collector has dropped the money'.

b. Ci-gon-a-mu-bawa i-ci c-a-gul-a nyumba.

7-sleep-FV-18.LOC-5.bar 7-this 7.SM-PRF-buy-FV 9.house
'This drunk has bought a house'.

c. I-ci ci-gon-a-mu-bawa c-a-gul-a nyumba.
7-this 7-sleep-FV-18.LOC-5.bar 7.SM-PRF-buy-FV 9.house
'This drunk has bought a house'.

d. *Ci-gon-a-mu-bawa li-ja li-mene ci-na-gon-a,
7-sleep-FV-18.LOC-5.bar 5-that 5-which 7-PST-sleep-FV
l-a-gw-a.
5-PRF-fall-FV
*'The drunk that bar where it slept has fallen'.

In (80a) above, the deictic pronoun *uyo* 'that' agrees in class with the nominal class of the compound word *mdobacidzala* 'garbage collector', and in (80b & c), the deictic pronoun agrees in class with *cigonamubawa* 'drunk'. Nevertheless, in (80d), where the deictic pronoun agrees with the compound internal noun *bawa* 'bar', the construction is ungrammatical. This suggests that the demonstrative cannot modify the nominal part only in VN-compounds. Note that the examples in (80b & c) show that both left- and right-adjunction are possible in Cinyanja. While some syntactic theories do not permit right-adjunction because of Kayne's (1994) Linear Correspondence Axiom (LCA), Carstens (2008) explicitly allows it in her analysis of the DP in Bantu (see also Carstens and Zeller (2020) and references therein for empirical evidence that adjuncts are exempt from LCA).

In this section, I have shown that VN-compounds are indeed DPs because they can function as subjects and objects of a clause, be modified by a relative clause, be coordinated with other DPs, and be modified by an adjective. These syntactic properties are expected from the structures for Type 2 and Type 3 compounds that I have suggested in section 3.4. However, as I discuss below, the problem with a purely syntactic analysis of these VN-compounds is that it does not explain that the internal parts of the compounds are not accessible to syntactic rules.

3.5 VN-compounds and the Lexical Integrity Hypothesis

The Lexical Integrity Hypothesis (LIH) suggests that syntax cannot manipulate or have access to the internal structure of words (Lapointe, 1980; Di Sciullo and Williams, 1987; Anderson, 1992; B & M, 1995; Ackema and Neeleman, 2004; Booij, 2009; Lieber, 2009a). The main argument behind this

principle is that morphological and syntactic rules are different. As a result, pieces of words cannot be manipulated by syntactic operations and cannot be accessible to elements outside of the word.

In this section, I use a battery of tests to show that VN-compounds with a VP structure are nominal compounds whose parts are not accessible to syntactic rules. Most of the tests are proposed by Lieber and Stekauer (2009) and B & M (1995). The tests include: pronominalisation, modification and inseparability (see Lieber and Stekauer, 2009; Haspelmath and Sims, 2010). Other tests are extraction, gapping and conjoinability (B & M, 1995).

3.5.1 Anaphoric pronominal replacement test

The anaphoric pronominal replacement test is used here to show that a part of VN compound words with a VP (or vP) internal structure cannot be replaced by a pronoun and/or deictic pronoun in the phrasal part (see also B & M, 1995 test on the inbound anaphoric island in Chichewa synthetic compounds). In a syntactic DP/NP, it is possible to substitute the head noun in the NP with a pro-form. For example, we can say a *black one* to refer to a *black bird* in a phrase. But in a compound word *blackbird,* it is impossible to replace *bird* with *one*; for example **blackone*. Further, in the examples below, the noun *watch* in the phrase *gold watch*, in (81a) has been replaced with the pro-form *one* which means that *gold watch* is a phrase. In contrast, in example (81b) the head noun, *smith*, cannot be substituted by a pronominal item *one*, which shows that *goldsmith* is a compound word. Examples (81a) and (81b) are from Haspelmath and Sims (2010:194).

81. a. My aunt has one *gold watch* and three *silver ones.*
    b. *My aunt knows one *goldsmith* and three *silver ones.*

Similarly, in Cinyanja, pronominal replacement is allowed in syntactic structures such as (82a) and (82b) while it is disallowed in compound words. For example, the noun *milomo*, 'lips', in the VP *penta milomo*, 'paint lips', can be replaced by *iyi*, 'these', as shown in example (82a'). However, in (82b'), anaphoric pronominal replacement of *–milomo*, 'lips' with *iyi* is impossible because *kapenta-milomo*, 'prostitute', is a compound word.

82. a. A-na-penta milomo.
       'She painted the lips'.
    a'. A-na-penta iyi.

‘She painted these (lips)’.

b. ka-pent-a-mi-lomo w-a-thaw-a.
1-paint-FV-4-lip 1.SM-PRF-run-FV
‘A prostitute has run away’.

b’. *ka-pent-a-iyi w-a-thaw-a.
1-paint-FV-these 1.SM-PRF-run-FV

The application of this test in (83) below is particularly interesting because the compounds in these examples include locative phrases, and locative phrases are unambiguously syntactic. As alluded to earlier, the locative prefix, unlike other noun class prefixes, can attach directly to roots or to base nouns that have primary prefixes. In (83a’) and (84a’), the locative nouns inside these sentences are substituted with a deictic pronoun, but this is not possible in (83b’) and (84b’) because the expressions are compound words.

83. a. Nandipa a-na-khala pa ci-tsa dzulo.
‘Nandipa she-PST-sit on 7.tree stump yesterday
‘Nandipa sat on a tree stump yesterday’.

a’. Nandipa a-na-khala pa ici dzulo.
‘Nandipa she-PST-sit on 7.this (i.e. tree stump) yesterday’.
‘Nandipa sat on this (i.e. tree stump) yesterday’.

b. M-khal-a-pa-citsa w-a-thyok-a mw-endo.
1-sit-FV-16.LOC-tree.stump 1.SM-PRF-break-FV 3-leg
‘The stump sitter has broken his leg’.

b’. * M-khal-a-pa-ici w-a-thyok-a mw-endo.
1-sit-FV-16.LOC-DEM 1.SM-PRF-break-FV 3-leg
Lit. *‘The on this sitter has broken his leg’.

84. a. A-na-gona mu bawa.
‘he-PST-sleep in bar’

'He slept in the bar'.

a'. A-na-gona mu iyo.
he-PST-sleep in that
'He slept in that one' (i.e. bar).

b. Ci-gon-a-mu-bawa c-a-gul-a nyumba.
7-sleep-FV-18.LOC-5.bar 7.SM-PRF-buy-FV 9.house
'The drunk has bought a house'.

b'. * Ci-gon-a-mu-iyi c-a-gul-a nyumba.
7-sleep-FV-18.LOC-this 7.SM-PRF-buy-FV 9.house
Intended: 'The drunk has bought a house'.

The examples in (83b') and (84b') show that the nominal part of the compound words, (i.e. the constituent corresponding to the DP selected by the locative prefix) cannot be substituted by a pronominal form. This provides evidence that these expressions are not syntactically transparent phrases, but compounds, whose parts are not accessible to operations such as pronominal substitution, in accordance with the lexical integrity hypothesis.

3.5.2 Modification

Another indication that VN-compounds with a VP/vP structure are indeed compound words is that a constituent part of these words cannot be modified. Although it is canonical for syntactic atoms such as nouns and verbs to be modified by adjectives and adverbs respectively, Haspelmath and Sims (2010:194) argue that constituents in a compound cannot be modified in this way. For example, in the phrase *a black bird*, the adverb *very* can modify the adjective *black* when it is used as a separate syntactic constituent but when the adjective *black* is part of a compound word such as *blackbird* to refer to a genus *Agelaius* (see Lieber and Stenkauer, 2009), the adverb *very* cannot modify the adjective *black* (compare **the very blackbird*). Similarly, in Cinyanja VN-compounds, constituent parts of a phrasal compound cannot be modified, as shown in examples (85b) and (86c, d), while phrasal elements can be modified as in (85a) and (86a, b). It is important to note that in many Bantu languages, including Cinyanja, attributive adjectives typically appear post-nominally, as illustrated below:

85. a. M-nyamata a-ku-pal-a ma-tabwa o-wuma

1-boy 1.SM-BE.PRS-scrape-FV 6-planks 6.ASS-dry a-mbiri.
6-many‘
The boy is scrapping so many dry planks’.

b. *M-pal-a-ma-tabwa o-wuma a-mbiri w-a-bwer-a.
1-scrape-FV-6-planks 6.ASS-dry 6-many 1.SM-PRF-come-FV
*’The carpenter (lit. "the many-dry-planks-scrapper") has come’.

86. a. Tombi w-a-khal-a pa mw-ala w-o-yera.
1.Tombi 1.SM-PRF-sit-FV 16LOC 3-stone 3-ASS-white
‘Tombi has sat on a white stone.’

b. Tombi w-a-khal-a pa mw-ala
1.Tombi 1.SM-PRF-sit-FV 16.LOC 3-stone
p-o-yera.
16.LOC-ASS-white
‘Tombi has sat on a white part of a stone’.

c. * M-khal-a-pa-mw-ala w-o-yera w-a-gul-a
1-sit-FV-L16.LOC-3-stone 3-ASS-white 1.SM-PRFbuy-FV
nkhuku.
9.chicken.
*’The white stone-sitter has bought a chicken’.

d. * M-khal-a-pa-mw-ala p-o-yera w-a-gul-a
1-sit-FV-L16.LOC-3-stone 16-ASS-white 1.SM-PRFbuy-FV
nkhuku.
9.chicken.
*’The white stone-sitter has bought a chicken.’

In the syntactic structure in (85a), the noun *matabwa* ‘planks’, has been modified by the adjective *owuma* ‘dry’, which agrees with the noun in class 6. However, when the same nominal element is part of the VN compound *mpalamatabwa* ‘carpenter’ in (85b), it cannot be modified by the adjective *o-wuma.* In the same way, in (86a) the noun *mwala* ‘stone’, which is part of the locative phrase that combines with the verb, is modified by the adjective *wo-yera* ‘white’, showing nominal concord. In (86b), the modifier shows locative concord. In contrast, in (86c) and (86d), the adjectives cannot modify

the base noun constituent *mwala* or the locative *pamwala* in the compound word *mkhalapamwala* 'stone sitter' because the noun together with the locative prefix and the verb form a compound word. It is evident from these examples that it is impossible to modify a noun inside a DP-projection that forms part of the subject DPs in (85b) and (86c, d). This, therefore, confirms that the phrasal constructs in these examples are indeed compound words.

3.5.3 Inseparability test

Another test, inseparability, suggests that no element can be inserted to separate the elements of a compound (Lieber and Stekauer, 2009:11; Altakhaineh, 2016a). Generally, it is easy to separate the constituents of a syntactic phrase by inserting material between them. For example, the possessor and possessee inside a DP can be separated by a prenominal adjective or a modifier which then intervenes between the elements of the phrase, as shown in (87b-c).

87. a. Ayanda's house
    b. Ayanda's big house
    c. Ayanda's big strong house

By contrast, elements of a compound are inseparable. Modifiers such as adjectives and adverbs cannot intervene between the compound members as illustrated in (88b) and (89b)

88. a. mosquito net
    b. *mosquito big net

89. a. pickpocket
    b. *pick-quickly-pocket

In example (88b) the adjective is prevented from separating the two members of the NN compound while in (89b) the adverb is equally disallowed from separating the elements of the VN compound word. This shows that parts of compound words cannot be separated by adjectives or adverbs.

In Cinyanja, I applied this test based on Lieber and Stekauer's (2009:11) claim that inseparability is the most reliable criterion for distinguishing compound words from phrases in English. The test is also highly relevant for Cinyanja compounds because of the VP structure inside them. Therefore, the adverbs *mosamala* 'carefully' in (90) and *mofulumira* 'quickly' in (91) were used. The examples in

(90a') and (91a') show that it is possible to separate the verb and its complement by modifying the verb with a postverbal adverb, which will now intervene between the verb and the complement phrase. However, the examples in (90b') and (91b') show that it is impossible to separate the same constituents inside a compound by inserting an adverb.

90. a. M-tsikana w-a-kha-la pa citsa.
1-girl 1.SM-PRF-sit-FV 16.LOC 7.stump.
'The girl sat on the stump'.

a'. M-tsikana a-na-khal-a **mosamala** pa citsa.
1-girl 1.SM-PST-sit-FV carefully 16.LOC 7.stump.
'The girl sat carefully on the stump.

b. m-khal-a-pa-citsa w-a-pit-a ku-mu-dzi.
1-sit-FV-16.LOC-7.tree stump 1.SM-PRF-go-FV 17.LOC-3-village
'The stump sitter has gone to the village'.

b'. * M-khal-a-**mosamala**-pa-ci-tsa w-a-pit-a ku-mu-dzi.
1-sit-FV-carefully-16.LOC-7-stump 1.SM-PRF-go-FV 17.LOC-3-village
Intended: 'The one who sits carefully on a stump has gone to the village'.

91. a. A-na-pala ma-tabwa y-anga y-onse.
1-PST-scrape 6-plank 6-my 6-all
'He scraped all my planks'.

a'. A-na-pala **mofulumira** ma-tabwa y-anga y-onse
1-PST-scrape fast 6-plank 6-my 6-all
'He scraped all my planks very fast'.

b. M-pal-a-ma-tabwa w-a-bwer-a.
1-srape-FV-6-plank 1.SM-PRF-come-FV
'The carpenter has come'.

b'. *M-pal-a-**mofulumira**-ma-tabwa w-a-bwer-a.
1-scrape-FV-fast-6-plank 1.SM-PRF-come-FV
Intended: 'The fast carpenter has come'.

The tests in (90a') and (91a') show that it is possible to separate phrasal constituents in a syntactic construct. However, in (90b') and (91b') the adverbs *mosamala* 'carefully' and *mofulumira* 'quickly' cannot separate the same elements when they are part of a VN-compound.

An objection that could be raised against the arguments provided here is that idiomatic complex phrasal expressions also cannot be interrupted by other material, even if they are unambiguously phrasal, and not compounds. For example, the noun inside a phrasal idiom such as *kick a bucket*, which is *unambiguously* a VP, cannot be modified with an adjective without losing the idiomatic meaning (e.g. in *kick the big bucket*, the idiomatic reading is no longer available). However, it is important to note here that the VN complex expressions used in these examples are not entirely idiomatic. Therefore, we cannot attribute the ban on modification entirely to the idiosyncratic lexical semantics of these constructions. Rather, the ungrammaticality of (90b') and (91b') is at least in part due to the LIH, which disallows the insertion of syntactic words such as adverbs into a morphological word such as a compound (Haspelmath and Sims, 2010; Lieber and Stekauer, 2009).

3.5.4 Extraction

Another test, extraction, is used to show that constituents of a compound word can neither be relativised nor topicalised (B & M, 1995). For example, we cannot extract *apple-juice* from the compound word, *apple-juice tray*, or relativise it, as in (92b).

92. a. The boy bought the *apple-juice tray* yesterday.
b. **Apple juice*, which the boy bought _____ tray yesterday.

In Cinyanja, extraction of the DP from inside a VN-compound is equally disallowed. B & M (1995) show that arguments of the verb can undergo extraction in Chichewa. However, despite the fact that the nominal part of VN-compounds constitutes an argument of the verbal part, as shown in subsection 3.3.2, the constituents that form these compounds cannot be relativised or topicalised. In example (93a) the VN-compound expression *mpalamatabwa* is the internal object argument of the verb *waona* 'has seen'. In (93b), the whole expression has been relativised. However, relativisation or topicalisation of only the nominal part of the compound expression renders the construction ungrammatical, as shown in (93c & d).

93. a. Mw-ana w-a-on-a **m-pala-ma-tabwa** m'-nyumba.
1-child 1.SM-PRF-see-FV 1.scrape-6-plank 18.LOC9.house
'The child has seen the plank scraper in the house'.

b. **M-pal-a-ma-tabwa** a-mene mw-ana w-a-on-a
1.scrape-6-plank 1-whom 1-child 1-PRF-see-FV
mu nyumba a-na-gul-a galimoto.
18.LOC 9.house 1.SM-PST-buy-FV 5.vehicle
'The plank scraper whom the child has seen in the house bought a vehicle'.

c. ***Ma-tabwa**, a-mene mw-ana w-a-on-a m-pala__
6-plank 6-REL 1-child 1-PRF see-FV 1-scrape__
m'-nyumba a-na-gul-a galimoto.
18.LOC-9-house 6.SM-PST-buy-FV 5.vehicle
*'The planks which the child has seen a scraper ____ in the house bought a car'.

d. ***Ma-tabwa**, mw-ana w-a-on-a m-pal-a_____
6-plank 1-child 1-PRF-see-FV 1-scrape
m'-nyumba a-na-gul-a. galimoto.
18.LOC-9-house 6.SM-PST-buy-FV 5.vehicle
*The planks, the child has seen a scraper ____ in the house bought a car.

The example in (93c) shows that extraction of a DP from the VN-compound in a relative clause, or left dislocation of the DP-part of the VN-compound in (93d), is not allowed. For further details about left and right dislocation refer to section 3.4.1 above.

3.5.5 Gapping

Another test, gapping or ellipsis, is based on B & M's (1995) argument that it is not possible to elide a part of a morphological word. For example, it is possible to gap the verb in (94b) without affecting the grammaticality and semantic interpretation of a sentence because the verb is an independent syntactic constituent. However, in (95b), it is impossible to elide *driver* from the compound *taxi-driver* because *taxi-driver* is a morphological word.

94. a. John sat on a chair and Mary sat on the floor.

b. John sat on a chair and Mary _____on the floor.

95. a. John is a truck-driver and Mary is a taxi-driver.
b. *John is a truck-driver and Mary is a taxi _____.

As discussed in subsection 3.2.3.2, B & M (1995) show that the noun inside the complement of a locative prefix can be gapped, which is one of the arguments that B & M provide to show that the locative is a syntactic construct. (96) repeats example (55) from above:

96. A-nyamata a-na-vin-a njerero
2-boy 2.SM-REC PST-dance-IND 9.name of dance
pa bwalo la mfumu Kapanga ndi pa
16.LOC 5.courtyard 5.ASS 9.chief Kapanga and 16.LOC
(bwalo) la mfumu Kapatuka.
5.courtyard 5.ASS 9.chief Kapatuka
'The boys danced the njerero dance on chief Kapanga's courtyard and on chief Kapatuka's (courtyard)'.

As (96) shows, the nominal part of a locative phrase can be gapped in Chichewa (the same is true in Cinyanja). However, in a VN-compound, when two locatives from the same noun class appear in two coordinated VN-compounds, gapping of identical material is impossible. Example (97b) demonstrates this for the locative prefix, and (97c) shows that the nominalised verb cannot be gapped either:

97. a. Kodi ndi m-khal-a-pa-mw-ala wa-m-tali kapena
Q is 1-sit-16.LOC-3.stone 1.ASS-1-tall or
m-khala-pa-ci-tsa wa-m-tali?
1-sit-16.LOC-7-tree.stump 1.ASS-1-tall
'Is she/he a tall stone sitter or a tall stump sitter?'

b. *Kodi ndi m-khala-pa-mwala wa-m-tali kapena
Q is 1-sit-16.LOC-3.stone 1.ASS-1-tall or
m-khala-(~~pa~~)-citsa wa-m-tali?

1-sit (16.LOC)-7.tree.stump 1.ASS-1-tall
Intended: 'Is he a tall stone sitter or a tall stump sitter'?

c. *Kodi ndi m-khala-pa-mwala wa-m-tali kapena
Q is 1-sit-16.LOC-3.stone 1.ASS-1-tall or
~~(m khala)~~-pa- citsa wa-m-tali.?
1-sit-(16.LOC- 7.tree.stump) 1.ASS-1-tall
Intended: 'Is he a tall stone sitter or tall stump sitter'?

From the examples above, it is clear that eliding the locative prefix in (97b), or the nominalised verb in (97c), renders these compounds unacceptable. This demonstrates that the VN-compounds are morphological words, hence are indeed nominal compound words.

In sum, different tests have been applied to show that the phrasal structure in VN-compounds is not accessible to syntactic operations and processes that can be applied to syntactically independent constituents. Therefore, this is evidence that VN-compounds with a phrasal structure are indeed compound words and obey the LIH.

3.6 Reification

In section 3.3, I analysed the internal structure of VN-compounds and argued that they have a phrasal structure. The DP in the object position merges with a √ROOT which is verbalised to realise a verb phrase (vP), and the entire phrase is nominalised by means of little n. The DP-projection of the nominalised structure corresponds to the VN-compound. Consequently, these VNs are phrasal compounds. In this section, building on Harley's (2009) work on phrasal compounds in English, I suggest that lexicalised phrasal constructs in Cinyanja can be reanalysed as roots and re-enter the syntactic derivation as unanalysable units. Following Harley (2009), I call this process *reification*.

Harley (2009) analyses phrasal compounds in English, which are formed by joining a syntactic phrase and a noun (Lieber, 1992; Salvini, 1984). The first element in a phrasal compound can be any kind of phrase, e.g. a DP as in ([*bikini-girls-in-trouble*] *genre*), a PP ([*over-the-fence*] *gossip*) or a sentence ([*God-is-dead*]-*theology*) (Pafel, 2017). The phrasal constituents in phrasal compounds behave like words because they cannot be separated and the elements within the modifying phrase cannot be inverted (Lieber, 2009a; Harley, 2009).

According to Harley, the structure of the phrasal XP-part of these compounds is nominalised by means of an n-head nominaliser before combining with the noun-part:

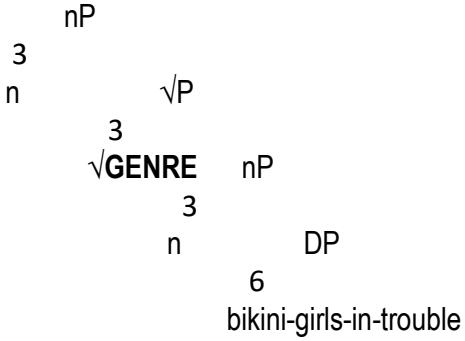
98. [[XP] n]$_{nP}$, e.g. [[$_{DP}$ bikini girls in trouble] n]$_{nP}$ (+ *genre*)

Harley notes that it is not mandatory for the nominalised XP to be part of a compound. The nP corresponding to the nominalised XP can also merge with derivational morphemes, as in the following derived adjective in example (99a), with the structure in (99b), based on the derivational suffix *-ish* of category (little) a (Harley 2009: 143):

99. a. feeling a bit **rainy-day-ish**
    b. [[$_{DP}$ rainy-day] n]$_{nP}$ -ish]$_{aP}$

While the nP formed by the merger of XP and n combines with a derivational suffix in (99), this nP combines with a root which is then itself nominalised in phrasal compounds (see (98)).

However, as Harley notes, the order of elements inside the compound raises a problem. Since n typically merges to the left of its complement, it is not clear how the order [[XP] n] in (98) is derived, given that incorporation of phrases is not an option. Similarly, the root of the nominal part of the compound is also expected to take nP as a complement to its right, but in the derived structure, the phrasal part of the compound (nP) precedes the root:

100.


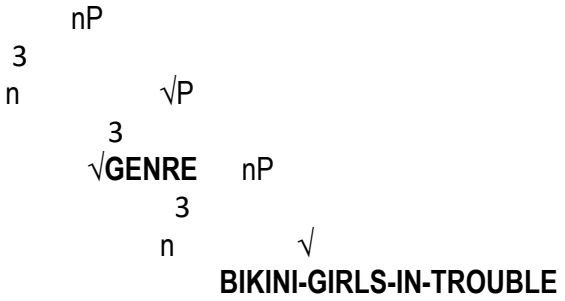
In order to explain how the XP-part of the phrasal compound ends up to the left of the higher noun, Harley argues that the phrasal part of these expressions can be "reified" in the process of merging with little n:

101. *Reification* (Harley 2009: 143)

The XP (part of a phrasal compound, N.N.) is created in a separate derivational workspace from a separate Numeration, sent off to LF for interpretation, and then 'renumerated' as a Root, in the derivation of the matrix clause – a Root denoting the abstract concept that was evoked by the previous computation of the XP's compositional meaning.

What Harley has in mind is that in phrasal compounds, the relevant phrasal structure (XP in (98)) is built by the syntax and mapped to the interface with semantics. Her important idea is that when the output of the semantic computation is lexicalised (as with phrasal idioms), the derived meaning is accessed by the computational system again, but as a root. This means that XP in (98) is not really a phrase, but a root, and (100), the basic structure of a phrasal compound like *bikini-girls-in-trouble genre*, rather looks like (102):

102.



For Harley, what is most critical is that the syntactic behaviour of the entire XP resembles that of a root instead of the internal complex XP, and she argues that it must be perceived as uttered for it to compose with the reifying n head. As I argue below, as a root, *bikini-girls-in-trouble* can incorporate into n and then combine with the higher root, solving the problem identified with (100) above.

In the following section, I suggest that only certain phrases can in principle undergo reification. Based on the concept of a phase (Chomsky 2000, 2001, 2008), I suggest that phase heads can trigger reification of their complements after Spell-out, provided that aspects such as speaker intention, culture etc. trigger a lexicalisation of the concept expressed by the complement of the phase head.

3.6.1 Phases and reification: basics

The concept of a phase as introduced in Chomsky (2000) refers to the Minimalist implementation of cyclic derivations. Radford (2004) describes a cycle as a syntactic operation (e.g. agreement and movement) that involves the merger of a head with one or more other constituents and results in a

repeat of the occurrence forming a cyclic phase pattern. According to Chomsky, in the course of the derivation, chunks of syntactic structure are transferred to the interfaces with semantics and phonology in the operation Spell-out. These chunks are determined by *phases*. What constitutes a phase has been the subject of debate since Chomsky (2000) (see e.g. Den Dikken, 2007; Gallego, 2007; Bošković, 2014 and references therein). Chomsky argues that certain heads in a clause structure are phase heads, and that the vP phrase headed by the little v and the CP phrase headed by the complementiser C are phases. When a phase is completed, the complement of the phase head is transferred to the interfaces. The derivation continues, but only the head of the phase and its "edge" (= its specifiers) remain accessible for syntactic operations; a condition which Chomsky (2001) calls the *Phase Impenetrability Condition* (PIC). The part of the phase that has undergone Transfer is no longer accessible to the syntax but is interpreted at the interfaces.

Chomsky's (2000) theorisation of phases indicates that the phrase vP headed by the little v is a phase (see also Kratzer and Selkirk, 2007). However, in DM, according to Marantz (2007) and Embick and Marantz (2008), the category-defining heads *a, n* and *v* are all phase heads and trigger Spell-out.

I now suggest that the type of phrase that can undergo reification in the sense of Harley (2009) are those phrases that correspond to phase domains and undergo Transfer, i.e. typically, complements of phase heads.[7] It has often been noted that the material transferred to the interface behaves like an unanalysable unit; it becomes a "giant lexical compound" (Uriagereka 1999: 421) that behaves like a syntactic atom (Hsieh and Sybesma 2011: 73). Reification, where the output of a complex phrase which is productively derived becomes an abstract syntactic root, is a process that can be applied to the output of phasal Transfer. A complex expression like *bikini-girls-in trouble*, which has been assembled and is merged as the complement of the head of the nP-phase, undergoes Transfer at this point and is interpreted as a syntactic atom. Reification can now apply and reanalyse this "giant lexical compound" as a root.

Reification has consequences beyond the PIC. The PIC already specifies that material inside a phase domain that has undergone Transfer is no longer accessible to syntactic operations. The constituents in the reified root therefore cannot be extracted, inverted, or moved. However, in the current analysis, the complex reified expression also behaves like a root. This means that after reification, the reified phrase can undergo syntactic operations as a unit. Since roots can undergo head movement, the

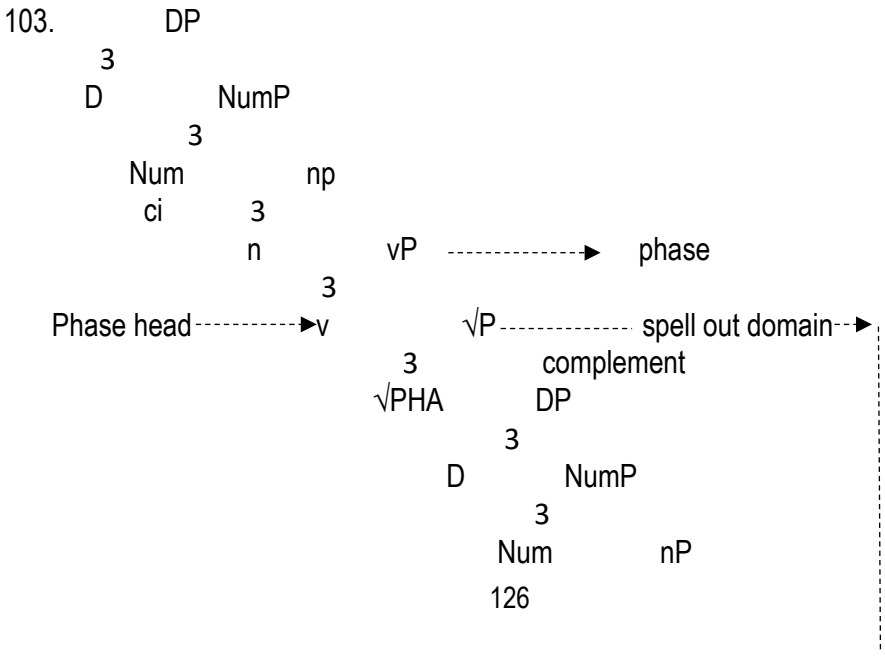
[7] Below I suggest that the root node of a structure, which also undergoes Transfer, can also be reified, even though it is not itself the complement of a phase head.

reified XP *bikini-girls-in-trouble* can undergo head movement and combine with n to form a complex head. This head can undergo further head movement and bring the whole reified phrase to the left of the noun *genre*.

The analysis of reification available to complements of little n, a and v potentially allows complements of the little n, a, and v to be reified. I leave the exploration of the broader implications of this idea for future research. In this study, I concentrate on the traditional view of considering the little v as a phase head. The complement of this phase head is "shipped off" when the little v is merged, and the unit so formed can then be targeted by reification and be inserted as a complex root.

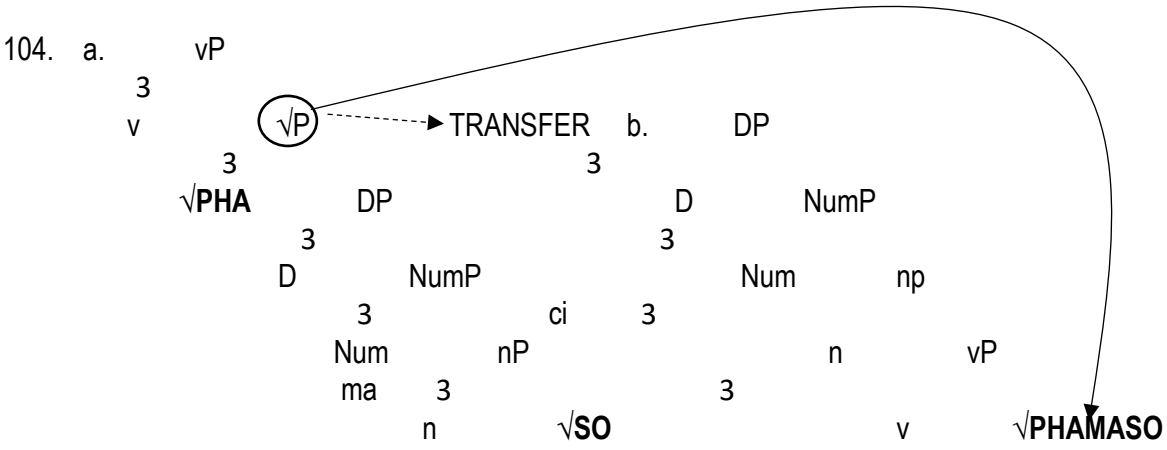
3.6.2 Proposed application of reification in Cinyanja

Applying this idea to the analysis of Cinyanja phrasal compounds, I consider the little vP as a phase and the little *v* as the phase head. I assume that a VN-compound like *ciphamaso* 'hypocrite' is derived as shown in the example below:



ma 3
n √SO ◄-------

In (103) above, the root √SO has no syntactic features and is merged with the syntactic head n, projecting nP, then the NumP and DP. The DP then merges with the root √PHA which is verbalised by the little v to realise a vP, (104a). At this stage, Transfer applies, and the complement of v, the √P, is sent to the interfaces. Now reification can apply, and the output of the syntactic computation is reanalysed as an abstract root. The √P is now treated as a new concept which is being used as a root at the same level as a root from the lexicon. The argument is that the phrasal structure of *phamaso* in (103) above is no longer a phrase but a root; when the little v triggers Transfer, the spelled-out phrase becomes completely re-rooted. The reified root is *phamaso*:



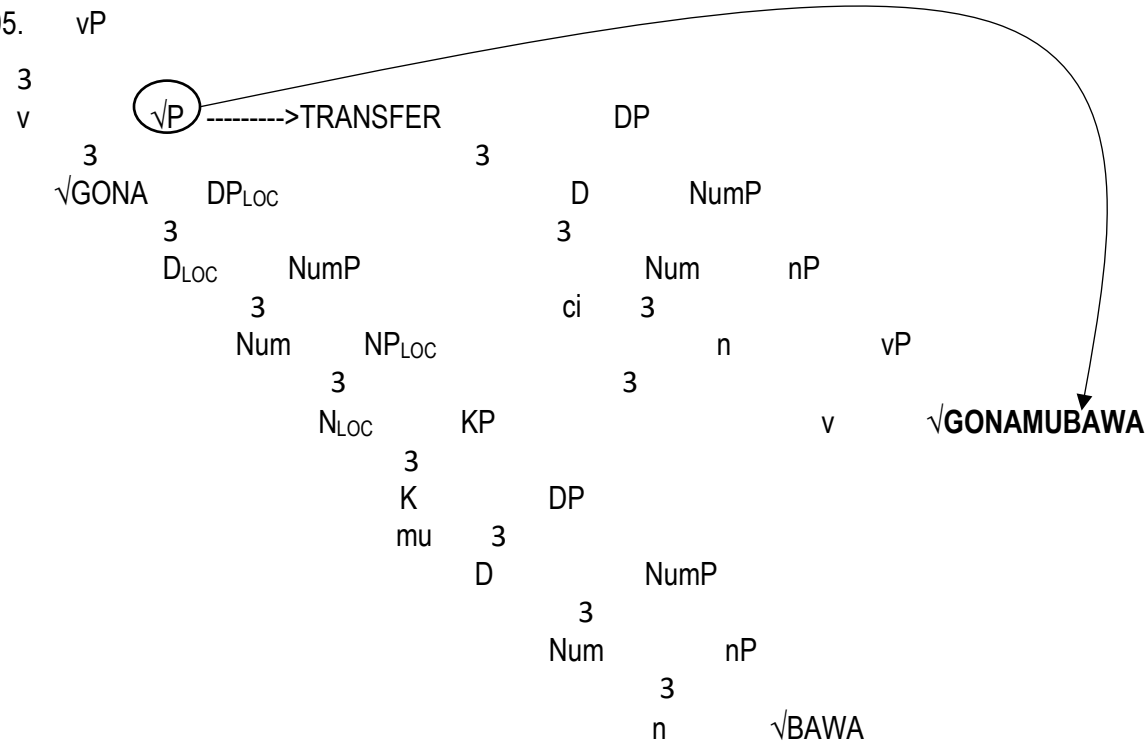
At this stage, the syntactic derivation can continue and *phamaso* can re-enter the syntax as an unanalysable root. The root is no longer accessible to any syntactic operation such as modification, extraction or conjunction. It now becomes atomic (see Di Sciullo and Williams, 1987). The reified root is the complement of the phase head v; now the little n is added, specifying gender, then Num (spelled out as the noun class prefix *ci-*), and then D projecting DP, as shown in (104b).

The root *phamaso* is nominalised by the little n in (104b), which is how the reified root becomes a noun. Even though syntactically, the reified root looks like a nominalised verb phrase, it is a root and as such part of a compound word. Consequently, none of the compound constituents can be modified, gapped, conjoined or extracted (see subsection 3.3.2 for details on modification, gapping, conjoinability and extraction).

Similarly, in Type 3 compounds such as *cigonamubawa*, 'drunk', where the nominal part includes a locative prefix, the vP is the phase and the little v is the phase head. When the phase head complement

is transferred to the interface, the entire phrase structure is re-rooted and appears in the syntax as an unanalysable root which is then merged with the new n, Num and D as shown in (105).

105.



In (105) above, the reified root √GONAMUBAWA becomes *cigonamubawa* after the vP merges with the little n, Num and D. Although syntactically the reified structure looks like a nominalised verb phrase, it is a compound word. Similar to (104) above, the constituents in the compound word cannot be modified extracted, gapped or pronominalised. Compound constituents cannot be conjoined.

3.6.3 Agreement ambiguities

In this section, I return to the observation made in section 3.1 and 3.4.3 above, that some of the VN-compounds in class 7 which refer to humans can optionally agree in class 1. This is surprising because simple human nouns in class 7 do not seem to allow this kind of noun class agreement, as shown in the examples below:

106. a. Ci-tsiru ci-ja c-a-pand-a mw-ana.
7-drink 7-that 7.SM-PRF-beat-FV 1-child
'That fool has beaten a child.

b. * Ci-tsiru u-ja w-a-pand-a mw-ana.

7-drink 1-that 1.SM-PRF-beat-FV 1-child

'That fool has beaten a child'.

107. a. Ci-gawenga ci-ja c-a-ph-a m-phuzitsi.
7-thug 7-that 7.SM-PRF-kill-FV 1-teacher
'That thug has killed the teacher'.

b. *Ci-gawenga u-ja w-a-ph-a m-phuzitsi.
7-thug 7-that 1.SM-PRF-kill-FV 1-teacher
'That thug has killed the teacher'

108. a. Ci-lema c-a-kuda c-a-b-a mw-ana.
7-lame 7-ASS-dark 7.SM-PRF-steal-FV 1-child
'The dark lame person has stolen the child'.

b. * Ci-lema w-a-kuda w-a-b-a mw-ana.
7-fool 1-ASS-dark 1.SM-PRF-steal-FV 1-child
'The dark lame person has stolen the child'.

However, it is interesting to note that VN-compounds that have a class 7 prefix and refer to humans behave differently in Cinyanja. Since these compounds are euphemistic terms that refer to humans, they do not behave like other simple [+human] words in class 7, but optionally agree in class 1, as exemplified in (109b), (110b) and (111b).

109. a. Ci-dy-a-ma-kanda ci-ja c-a-mangidw-a.
7-eat-FV-.6-baby 7-that 7.SM.PRF—arrest-FV
'That sugar daddy has been arrested'.

b. Ci-dy-a-ma-kanda u-ja w-a-mangidw-a.
7-eat-FV-.6-baby 1-that 1.SM-PRF-arrest-FV
'That sugar daddy has been arrested'.

c. *Ci-dy-a-ma-kanda ci-ja w-a-mangidw-a.
7-eat-FV-.6-baby 7-that 1.SM-PRF-arrest-FV
'That sugar daddy has been arrested'.

110. a Ci-gon-a-mu-bawa ci-ja c-a-gul-a nyumba.
7-sleep-FV-18.LOC-5.bar 7-that 7.SM-PRF-buy-FV 9.house
'That drunk has bought a house'.

b. Ci-gon-a-mu-bawa u-ja w-a-gul-a nyumba.
7-sleep-FV-18.LOC-5.bar 1-that 1.SM-PRF-buy-FV 9.house
'That drunk has bought a house'.

c. * Ci-gon-a-mu-bawa ci-ja w-a-gul-a nyumba.
7-sleep-FV-18.LOC-5.bar 7-that 1.SM-PRF-buy-FV 9.house
'That drunk has bought a house'.

111. a. Ci-ph-a-ma-so ci-ja c-a-kwatir-a m-phuzitsi.
7-kill-FV-6-eye 7-that 7.SM-PRF-marry-FV 1-teacher
'That hypocrite has married a teacher'.

b. Ci-ph-a-ma-so u-ja w-a-kwatir-a m-phunzitsi.
7-kill-FV-6-eye 1-that 1.SM.PRF-marry-FV 1-teacher
'That hypocrite has married a teacher'.

c. * Ci-ph-a-ma-so ci-ja w-a-kwatir-a m-phunzitsi.
7-kill-FV-6-eye 7-that 1.SM.PRF-marry-FV 1-teacher
'That hypocrite has married a teacher'.

d. * Ci-ph-a-ma-so u-ja c-a-kwatir-a m-phunzitsi
7-kill-kill-FV-6-eye 1-that 7.SM-PRF-marry-FV 1-teacher
'That hypocrite has married a teacher'.

Notice that VN-compounds do not allow for "mixed" agreement. In (111a), for example, both the demonstrative modifier and the verb agree in class 7; in (111b), both agree in class 1. However, in (111c) and (111d) where there is "mixed" agreement, the result is ungrammatical.

Another interesting characteristic of class 7 VN-compounds is that they take their plural in class 2. In fact, for most Cinyanja speakers in Zambia, the plural of *ciphamaso* must be *aciphamaso* in class 2,

while class 8 **ziphamaso* is not acceptable.[8] Note that in these plural forms, the noun class 2 prefix does not replace the class 7 prefix *ci-*, but is added to the whole compound word, as shown in (112a) and (113a). In this case, the class 7 noun class agreement marker then becomes opaque and cannot agree with modifiers and the verb as shown in (112b-d) and (113b-d).

112. a. A-ci-ph-a-ma-so a-ja o-nyada a-a-thaw-a.
2-7-kill-FV-6-eye 2-that 2-proud 2.SM-PRFrun-FV
'Those proud hypocrites have run away'.

b. *A-ci-ph-a-ma-so ci-ja co-nyada a-a-thaw-a.
2-7-kill-FV-6-eye 7-that 7.ASS-proud 2.SM-PRF-run-FV
Intended: 'Those proud hypocrites have run away'.

c. *A-ci-ph-a-ma-so ci-ja co-nyada c-a-thaw-a.
2-7-kill-FV-6-eye 7-that 7.ASS-proud 7.SM-PRF-run-FV
Intended: 'Those proud hypocrites have run away'.

d. *A-ci-ph-a-ma-so ci-ja o-nyada c-a-thaw-a.
2-7-kill-FV-6-eye 2-that 2-proud 7.SM-PRF-run-FV
Intended: 'Those proud hypocrites have run away'.

113 a. A-ci-gon-a-mu-bawa a-ja a-na-gul-a nyumba.
2-7-sleep-FV-18.LOC-5.bar 2-that 2.SM-PST-buy-FV 10.house
'Those drunks bought a houses'.

b. *A-ci-gon-a-mu-bawa a-ja ci-na-gul-a nyumba.
2-7-sleep-FV-18.LOC-5.bar 2-that 7.SM-PST-buy-FV 10.house
Intended: 'Those drunks bought a houses'.

c. *A-ci-gon-a-mu-bawa ci-ja ci-na-gul-a nyumba.
2-7-sleep-FV-18.LOC-5.bar 7-that 7.SM-PRF-buy-FV 10.house

[8] While in Zambia, *aciphamaso* is the most preferred plural for *ciphamaso*, this may not be true for Chichewa speakers elsewhere (e.g. in Malawi), who may use class 8 as the plural e.g. *ziphamaso*. These disparities in pluralisation strategy may be based on language variations and demographic reasons (see also below).

Intended: 'Those drunks bought a houses'.

d. *A-ci-gon-a-mu-bawa ci-ja a-na-gul-a nyumba.
2-7-sleep-FV-18.LOC-5.bar 7-that 2.SM-PST-buy-FV 10.house
Intended: 'Those drunks bought a houses'.

Example (112a) and (113a) indicate that pluralisation of noun class 7 compounds that refer to humans can be achieved by adding a noun class 2 prefix to the entire compound expression. This suggests that indeed certain VN-compounds that refer to humans can optionally be classified as belonging to class 1/2.

Further, class 7 VN-compounds that refer to humans, like class 1 nouns, can also combine with diminutive or pejorative/augmentative class 12 and 13 noun class prefixes; in which case agreement is with class 12 and 13, not class 7 as shown below.

114. a. Ka-ci-ph-a-ma-so ko-nyada k-a-bwer-a.
12-7-kill-FV-eye 12.ASS-proud 12.SM-PRF-come-FV
'A small/slim proud hypocrite has come'.

b. *Ka-ci-ph-a-ma-so co-nyada c-a-bwer-a.
12-7-kill-FV-6-eye 7.ASS-proud 7.SM.-PRF-come-FV
'A small/slim proud hypocrite has come'.

c. *Ka-ci-ph-a-ma-so o-nyada a-a-bwer-a.
12-7-kill-FV-6-eye 6.ASS-proud 6.SM-PRF-come-FV
'A small/slim proud hypocrite has come'.

115. a. Ti-ci-ph-a-ma-so to-nyada t-a-bwer-a.
13-7-kill-FV-6-eye 13.ASS-proud 13.SM-PRF-come-FV
'The proud hypocrites have come'.

b. *Ti-ci-ph-a-ma-so co-nyada c-a-bwer-a.
13-7-kill-FV-6-eye 7.ASS-proud 7.SM.-PRF-come-FV
'The proud hypocrites have come'.

c. *Ti-ci-ph-a-ma-so o-nyada a-a-bwer-a.

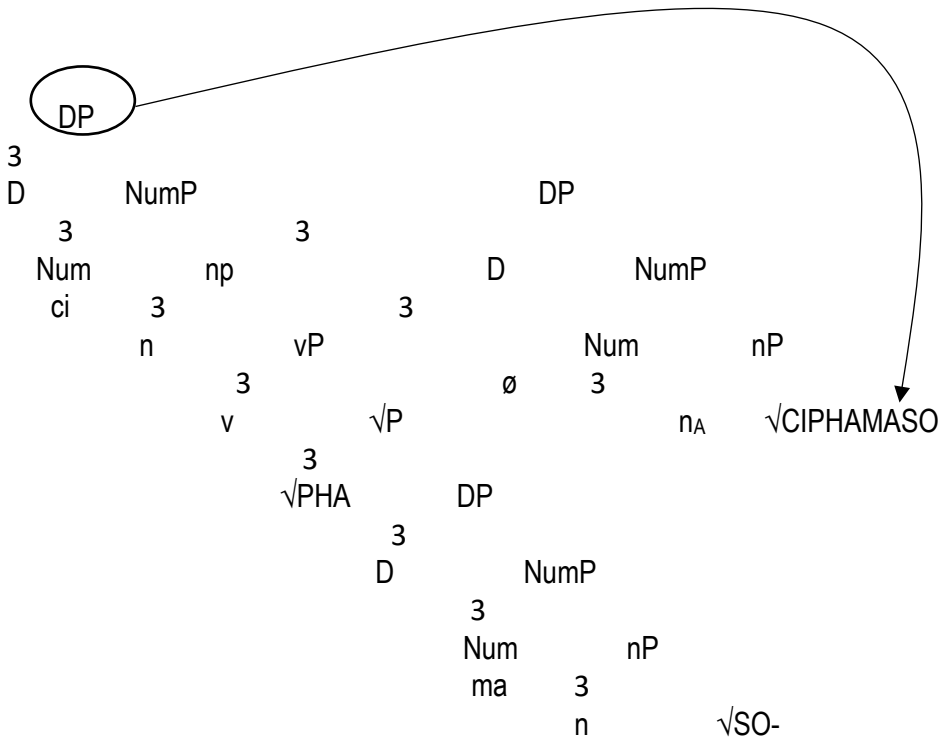
13-7-kill-FV-6-eye 6.ASS-proud 6.SM -PRF-come-FV
'The proud hypocrites have come'.

The examples above show agreement ambiguities for class 7 compounds. I claim that the ambiguities may be attributed to dialectal differences between varieties of the same language. A dialect may be interpreted as a particular form of a language spoken in a certain region or geographical area and may show some differences in pronunciation, grammar and idiomatic use of expressions from the other language considered the standard form, but not adequately distinct to be called a different language (see Haugen, 1966; Solano-Flores, 2006). A good example is Chichewa spoken in Malawi and Cinyanja spoken in Zambia. Although there are some phonological and grammatical differences between these two language varieties, the languages are mutually intelligible. The dialectal differences are mainly due to the linguistic distance between the two speech communities. That is why, when it comes to compound expressions in class 7 (e.g. *cigonamubawa*), issues pertaining to the agreement may be controversial. In Cinyanja, as shown in the examples in sections 3.1.and 3.6.3, compounds in class 7 that refer to persons may not take the formal NC 8 in the plural. Instead, NC 7 compounds take NC 2 plural. This behaviour of NC 7 compounds when it comes to the plural form may be attributed to the fact that in Bantu languages, animate interpretations generally take NC 1 (singular) and NC 2 (plural) agreement. This may elaborate why agreement options in class 1 and class 2 of class 7 compound words are possible for Cinyanja speakers in Zambia resulting in the agreement ambiguities shown in subsection 3.4.3 above. This explains why, although the reification possibility may exist for these lexicalisations in Cinyanja, it may not be possible in Chichewa NC 7 compounds. As a native speaker of Cinyanja in Zambia, my judgement is that class 7 [+human] compounds optionally take class 1 agreement and take class 2 plural for Zambian speakers while simple class 7 nouns take class 8 plural (e.g. the plural for class 7 noun *citsiru* is *zitsiru* of noun class 8 not **atsiru* of noun class 2). This may not be the same with speakers of Chichewa elsewhere.

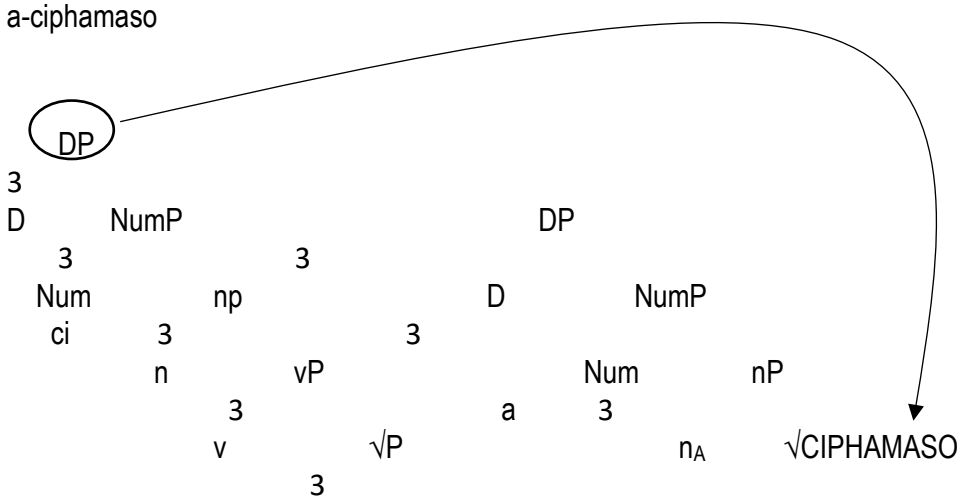
How can we explain that these [+human] VN-compounds can optionally agree in class 1, and are pluralised in class 2? I argue that the ambiguities in agreement for VN-compounds in class 7 may be a case of "delayed reification": when the vP in (104) is completed, its complement undergoes Transfer. Now suppose that, in contrast to what happens when a class 7 phrasal compound is derived, reification does not yet take place at this stage of the derivation. Rather, the vP first undergoes further derivation and derives a syntactically well-formed DP (see (105)) Importantly, when this structure is completed, the root node undergoes Transfer. Since reification can apply to the output of Transfer, this means that the whole DP, which includes not only the vP-phrase *phamaso* but also the class 7 prefix *ci-*, which realises Num, may be reified when completed. As a result, we derive a new root √CIPHAMASO,

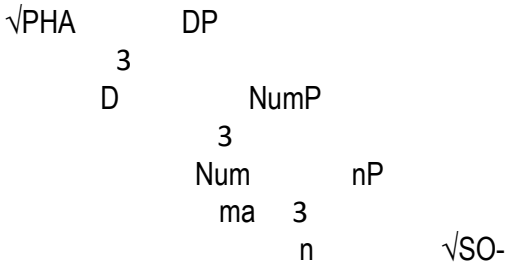
which can now merge with a new n. This new little n belongs to gender class A ([+human]). When a singular Num is added, Num is spelled out as zero, and we derive a class 1 noun *ciphamaso*, which agrees in class 1 with modifiers and predicates as shown in (111b). When a plural Num is added to nP, it is spelled out as *a-,* and we get *aciphamaso*, and agreement is in class 2 as shown in (112a). The analysis of the class 1 compound *ciphamaso* is shown in (116a), and the derivation of the class 2 plural form is shown in (116b):

116. a.



b. a-ciphamaso





(116) shows the derivation of VN-compounds which agree in class 1 or 2. However, without the delayed reification, and the additional nP-structure which contributes the gender class A, *ciphamaso* will be derived as in (104) and consequently agree in class 7, as shown in (111a).

According to this analysis, mixed agreement in (109c), (110c) and (111c) is ungrammatical due to reification. Once the little n of gender A is added to a reified noun *ciphamaso*, *ciphamaso* behaves like a new root inside this nP. It is not accessible for further derivation. This means that the noun class information realised by the class 7 prefix becomes opaque. The class 7 information is gone; it is no longer part of the compound word.

This delayed reification analysis is only available with phrasal compounds. Therefore, the simple class 7 human nouns cannot be nominalised as roots and combine with gender A little n. This explains the difference between the compounds which can agree in class 1/2, and nouns like *cigawenga* "thug" and *citsiru* " fool" which cannot.

3.7 Conclusion

The chapter discussed VN-compounds in Cinyanja and showed that these compound expressions with phrasal material are genuine compounds. The noun class prefixes and locative prefixes are part of the nominal compound word morphology. The VN-compounds in Cinyanja were analysed as having a VP/vP phrasal structure. However, the compounds at the same time are nouns and adhere to LIH. The constituents of the compound words cannot be pronominalised, modified, separated, extracted or gapped. The syntactic analysis of the compounds also shows that the phase domain is impenetrable for additional syntactic operations, as predicted by the PIC. Therefore, I claim that the complement of the VP/vP part of the compounds is a root that can be reified and can re-enter the syntactic derivation as an unanalysable unit. The reified root then merges with new n, Num and D to realise the VN compound.

CHAPTER 4: ASSC-COMPOUNDS IN CINYANJA

4. Introduction

In Chapter 3, I have looked at VN-compounds and have shown that the syntactic structure in relation to the Lexical Integrity Hypothesis commits us to the view that they are words. My analysis was guided by the idea that certain phrasal constructs can be reified and re-introduced as roots. In this chapter, I discuss a different type of phrasal compound which is formed by connecting two nouns with an associative marker in Cinyanja, and which also has a phrasal structure. However, in contrast to what I showed in Chapter 3 about VN-compounds, my evidence shows that these associative compounds

have a phrasal structure because the syntax has access to the internal structure of the words. Consequently, the complex expressions are not words but phrasal idioms for which the concept of reification is not necessary.

The compounds in Cinyanja analysed in this chapter have an associative marker inside them as shown in the examples below:

1. a. njinga ya moto
   9.bicycle 9.ASS 3.fire
   'bicycle of fire' (literal meaning)
   'motor-bike' (lexical meaning)

   b. m-misiri wa nyumba
   1-expert 1.ASS 10.house
   'expert of houses' (literal meaning)
   'builder' (lexical meaning)

In the compound words above, the two nouns in (1a) are *njinga* 'bicycle and *moto* 'fire' linked together by an associative marker *ya* 'of' while in (1b) the two nouns are *mmisiri* 'expert' and *nyumba* 'houses' linked together by the associative marker *wa* 'of'. These nominal compounds are mostly referred to as associative nominals or associative constructions in the Bantu literature (Mugane, 1997; Delfitto et al., 2011; Kula, 2012), and they seem to have the same structure as regular associative nominal phrases which are generally syntactic as shown in example (2) because, by assumption, they include the agreeing associative marker. Further, Delfitto et al. (2011) argue that such associative constructions are syntactic because they allow modification of internal constituents.

2. a. delesi la mw-ana
   5.dress 5.ASS 1-child
   'the child's dress'

   b. mphika wa ndiwo
   3-pot 3.ASS 9.relish
   'a pot of relish'

In (2) above, the regular associative nominal phrases, like the compounds with the associative marker, also include two nouns; the head noun and the possessor[9] linked together with an associative marker. Associative nominals such as (2) are discussed in detail in Delfitto et al. (2011). The authors provide evidence that they are syntactic constructions because they allow internal modification of constituent parts.

In light of the superficial similarity between the examples in (1) and those in (2), one may wonder why the former complex expressions have been analysed as "compounds" in the first place. I refer to these expressions as compounds in this study because, like the VN-compounds discussed in Chapter 3, they have some of the properties that were identified as properties of compounds in Chapter 2. First, they are generic; that is, the non-head does not refer to any specific entity (Haspelmath and Sims, 2010). For instance, in *mmisiri wa nyumba* 'builder', the noun *nyumba* 'houses' does not refer to any particular house, but the entire expression refers to a "builder". A second reason why these complex expressions are called compounds is that the constituent parts cannot be separated by an adjective because separating them would either alter the lexical meaning or conventionalised meaning of the expression (see also Lieber's (2009a) discussion of compound words in English). For example, in the compound word *mmisiri wa nyumba*, we cannot insert an adjective *wamtali* 'tall' between the head noun and the associative phrase to have *mmisiri wamtali wa nyumba* the 'tall expert of houses' because the adjective would only modify *mmisiri* 'expert' and not the entire expression which means 'builder'. Notice that with regular associative constructions such as those in (2), an adjective can be inserted in this position (compare *mphika waukulu wa ndiwo* 'a big pot of relish'). Further, *mmisiri* alone does not translate to a builder, it is a general term that means expert, for example, *mmisiri wokonza njinga* 'expert of repairing bicycles/bicycle repairman', *mmisiri wosema* 'expert of sculpture/sculptor'. Therefore, to retain the interpretation of the complex expression as 'builder', the adjective has to modify the entire complex expression, that is, *mmisiri wa nyumba wamtali '*a tall builder'. (The possibilities of modifying compounds such as those in (1) are discussed in detail in section 4.4 below.) Third, these expressions are referred to as compounds because their meaning is conventionalised and fully listed in the lexicon. This is in contrast to syntactic constructions such as *mphika wa ndiwo* 'a pot of relish', where the interpretation of the whole construction is fully determined by the meaning of its parts.

[9] I use the term "possessor" broadly here to refer to the noun following the associate marker inside the associate phrase. It should be kept in mind, however, that the semantic relationship between the two nouns is not always that of possession. In (2b), for example, the "possessor" is a content or purpose modifier of the head noun (see Defitto et al. 2011). The semantic relations that can be expressed by means of an associate phrase are discussed in section 4.2 below.

However, even though I use the term "compound" to refer to these expressions, it is my aim in this chapter to show that these lexicalised complex expressions are not words, but are syntactic constructs with special conventionalised meanings. In other words, I argue that the complex expressions in (1) are a subset of the class of associative nominals, and have the same syntactic structure. However, because their meaning is conventionalised and often idiosyncratic, I argue in this chapter that they should be analysed as phrasal idioms. This type of "compound", therefore, differs fundamentally from the VN-compounds analysed in Chapter 3 in that they are genuine syntactic constructs, rather than words built on "reified" syntactic structures.

The main focus of this chapter is to analyse the complex expressions linked with an associative marker (henceforth Assc-compounds) in Cinyanja. First, the chapter discusses the structure of Assc-compounds in Cinyanja where I identify two types of associative compounds. This is followed by a literature review of work on Bantu associative markers. In this section, I review the nature of the associative marker, its formal characteristics, morphological properties, semantic properties and how it is spread across Bantu languages. Then I discuss the syntax of complex DPs where I show how arguments and modifiers are realised within the noun phrase. This is followed by a discussion of Bantu associative nominal phrases (ANPs). The analysis of the internal morphosyntax of Bantu ANPs is presented where I argue, following Carstens (2018), that the associative (of category K, for "Kase") and the possessor DP form a KP structure. I then argue that Cinyanja Assc-compounds also have a KP internal structure because the syntax of these “compounds” is the same as that of the Bantu ANPs. The internal structure of the Cinyanja Assc-compounds is then analysed to show how the agreement is established between the head noun and the KP. This is followed by a syntactic discussion of Assc-compounds as DPs. The section on the Lexical Integrity Hypothesis shows that Assc-compounds have phrasal properties and that their parts are accessible to syntax. In the last section, I discuss phrasal idioms and their properties, and I show that Assc-compounds with a KP internal structure can be analysed as phrasal idioms.

4.1 Assc-compounds

In this section, I present the data of Ass-compounds in Cinyanja. I note that these compounds have a unique structure because they seem to resemble the structure of regular syntactically derived associative noun phrases (ANPs) which are discussed in section 4.3, except that the meaning of the compounds is lexicalised. In this study, Ass-compounds are classified into Types 1 and 2. For each type, I present both the singular and plural forms of the compounds. I present the singular forms in the (a) examples and the plural forms in the (b) examples.

4.1.1 Type 1 Assc-compound: Compositional and conventionalised

Type 1 Assc-compounds are compositional because the meaning of each "compound" is predictable from the meaning of the constituent parts (Neef, 2009). In this respect, Type 1 Assc-compounds behave like English compound words such as *schoolbook* 'a book used in a school', which is compositional because its meaning is predictable from the constituent parts *school* and *book.* As regards its semantics, Type 1 Assc-compounds can be described either as having a completely transparent meaning which means that the meaning of the compound is determined from both compound constituents, or as having a partially transparent meaning where only the head of the compound contributes to the meaning of the entire compound word (see also Altakhaineh, 2016a). Additionally, however, the compounds in this category are conventionalised which means that the compound is a very specific instance of that meaning, for example, a *schoolbook* is not any odd book that one finds in a school, but it also includes additional meaning aspects such as that it is used to teach children in school. Therefore, the fact that the meaning of these compounds is compositional is consistent with them being stored in the mental lexicon. Clark et al (1986) use conventionality to mean the use of words to express requisite meaning; that is, meaning that refers to a specified item or purpose. The data for Type 1 compounds are presented below:

3.	a.	buku	la	nyimbo	
		5.book	5.ASS	10.song	
		'the book of songs'			(literal meaning)
		'songbook'			(lexicalised meaning)
	b	ma-buku	a	nyimbo	
		6.book	6.ASS	10.song	
		'books	of	songs'	(literal meaning)
		'songbooks'			(lexicalised meaning)
4.	a.	buku	la	ma-pemphero	
		5.book	5.ASS	6-prayer	
		'book of prayers'			(literal meaning)
		'prayer book'			(lexicalised meaning)
	b.	ma-buku	a	ma-pemphero	
		6.book	6.ASS	6-prayer	
		'books of prayers'			(literal meaning)

‘prayer books’ (lexicalised meaning)

5. a. pepala la ma-yeso
5.paper 5.ASS 6.examination
‘paper of examinations’ (literal meaning)
‘examination paper’ (lexicalised meaning)

b. ma-pepala a ma-yeso
6.paper 6.ASS 6.examination
‘papers of examinations’ (literal meaning)
‘examination papers’ (lexicalised meaning)

6. a. mw-ana wa sukulu
1-child 1.ASS 5.school
‘a child of school’ (literal meaning)
‘pupil/school child’ (lexicalised meaning)

b. ana a sukulu
2.child 2.ASS 5.school
‘children of school’ (literal meaning)
‘pupils’ (lexicalised meaning)

7. a. mw-ana wa ma-siye
1-child 1.ASS 6-leave
‘a child of the deceased’ (literal meaning)
‘orphan’ (lexicalised meaning)

b. ana a ma-siye
2-child 2.ASS 6.leave
‘children of the deceased’ (literal meaning)
‘orphans’ (lexicalised meaning)

8. a. mu-lezi wa ana
1-keeper 1.ASS 2.child

		‘keeper of children’			(literal meaning)
		‘nanny’			(lexicalised meaning)
	b.	a-lezi	a	ana	
		2-keeper	2.ASS	2.child	
		‘keepers of children’			(literal meaning)
		‘nannies’			(lexicalised meaning)
9.	a.	m-misiri	wa	nyumba	
		1-expert	1.ASS	10.house	
		‘expert of houses’			(literal meaning)
		‘builder’			(lexicalised meaning)
	b.	a-misiri	a	nyumba	
		2-expert	2.ASS	10.house	
		‘experts of houses’			(literal meaning)
		‘builders’			(lexicalised meaning)
10.	a.	Njinga	ya	moto	
		9.bicycle	9.ASS	3.fire	
		‘bicycle of fire’			(literal meaning)
		‘motorbike’			(lexicalised meaning)
	b.	Njinga	za	moto	
		10.bicycle	10.ASS	3.fire	
		‘bicycles of fire’			(literal meaning)
		‘motorbikes’			(lexicalised meaning)

The examples above are all compositional and have conventionalised meaning but differ in levels of transparency. Examples (3-5) are completely transparent because the meaning of these phrasal compounds is predicted from their constituent parts. While it is easy to get the meaning of *pepala la mayeso* ‘examination pa per’ just by looking at the constituent parts, it is not easy to get the meaning of *mmisiri wa nyumba* ‘expert of ‘houses’ just by looking at the constituent parts. Therefore, although the meaning of the compounds in (6-10) is conventionalised, the compounds are not completely transparent because their meaning has to be learned by looking at the expression as a whole. In

example (10), *njinga ya moto* 'motorbike', is partially transparent because only the head noun contributes to the meaning of the entire compound word[10].

4.1.2 Type 2 Ass-compounds: non-compositional with idiosyncratic meaning.

The Type 2 Ass-compounds are non-compositional; the meaning of the compound as a whole cannot easily be construed from the meaning of the constituent parts. Normally, the sequence of the constituents that make up these constructions like compound words in general (see Nunberg et al., 1994; Kavka, 2009), is fixed. With regards to their semantics, Type 2 compounds have an idiosyncratic meaning which is not predictable from their constituent parts, as illustrated below:

11.	a.	madzi	a	moyo	
		6.water	6.ASS	3.life	
		'water of life'			(literal meaning)
		'oral-rehydration'			(lexicalised meaning)
12.	a.	mw-ana	wa	m-thengo	
		1-child	1.ASS	18.LOC-bush	
		'a child of the bush'			(literal meaning)
		'bastard'			(lexicalised meaning)
	b.	ana	a	m-thengo	
		2.child	2.ASS	18.LOC-bush	
		'children of the bush'			(literal meaning)
		'bastards'			(lexicalised meaning)
13.	a.	mw-ana	wa	leza[11]	
		1-child	1.ASS	1.god	
		'child of god'			(literal meaning)
		'rainbow'			(lexicalised meaning)
	b.	ana	a	leza	

[10] In *njinga ya moto* 'motorbike', the head noun *njinga* 'bicycle', even though it has the same shape both in the singular and plural, belongs to two different noun classes. This is also a typical characteristic of nouns in class 9 and 10 where the morphological shape of the nouns remains the same in the singular and in the plural

[11] *Mwana wa leza*, 'rainbow' is also known as *uta wa leza* 'bow of god' (i.e. rainbow) by some people.

		2.child	2.ASS	1.god	
		'children of god'			(literal meaning)
		'rainbows'			(lexicalised meaning)
14.	a.	wailesi	ya	kanema	
		9.radio	9.ASS	12.picture	
		'radio of pictures'			(literal meaning)
		'television set (TV)'			(lexicalised meaning)
	b.	ma-wailesi	a	kanema	
		6.radio	6.ASS	12.picture	
		'radios of pictures'			(literal meaning)
		'television sets (TVs)'			(lexicalised meaning)

In the Type 2 category, the Assc-compounds refer to things, objects and humans. In the examples above, the head nouns belong to various NCs. Some belong to class 1 and have their plural in class 2 as in example (12) and (13). Note that the compound in example (13) refers to an inanimate entity, even though the head noun belongs to class 1/2. Other head nouns belong to class 6, for example, *madzi* 'water' in example (11). In *madzi a moyo* 'oral-rehydration', the head noun, *madzi* is a liquid and as a mass noun, belongs to class 6 by default.

It is important to note that the borderline between Type 2 and the Type 1 compounds which are partially transparent is not rigid. For example, the Type 1 compound *njinga ya moto* 'motor-bike' and the Type 2 compound *wailesi ya kanema* 'TV' are both lexicalised and have idiomatic meaning. The only difference is that in *njinga ya moto*, the head noun partially contributes to the meaning of the compound as a 'type of bicycle' (i.e. a motorbike) whereas it is not the case with *wailesi ya kanema* 'TV' where the head noun *wailesi* 'radio' does not contribute to the meaning of the compound as 'television set'.

The two nouns in all the Assc-compounds discussed in this chapter are linked by an associative marker (AM). This means that an understanding of the role of the AM in Bantu and Cinyanja, in particular, is necessary. The AM is discussed in the next section.

4.2 The nature of the associative marker in Bantu languages

In this section, I explore the nature of the associative marker, its formal characteristics, morphological properties, semantic properties and how it is spread across Bantu languages. The AM is found in most Bantu languages, and is used in associative constructions. It is sometimes referred to as a genitive marker (Abney, 1987; Carstens, 1991; Bentley, 1998), connective (Van de Velde, 2013), or linker (Creissels et al., 2008). The AM expresses different relations in many Bantu languages (Welmers, 1963), and links two related nouns, namely the possessor and the head noun (Matamboirofa, 2000; Spreda, 2000; Kula, 2009; Carstens, 2018).

Bresnan and Mugane (2006), Carstens (2000), Delfitto et al. (2011) and many others, note that the associative marker bears a concordial prefix which agrees only with the noun class of the preceding noun, which is the head noun as shown in the Swahili examples below:

15. a. picha ya Mariamu
9.picture 9.ASS 1.Mariamu
'a picture of Mariamu'

b. picha za Mariamu
10.picture 10.ASS 1.Mariamu
'pictures of Mariamu'
[Carstens, 2000:323]

The AM consists of a root morpheme *-a* and the concordial agreement marker. For example, in (15a), the AM *ya* 'of' contains a class 9 Agr-marker *y-* attached to the root *-a* which agrees with the class 9 subject noun *picha* 'picture'. In (15b), the associative marker *za* contains a class 10 Agr-marker which agrees with the head noun '*picha'*, a class 10 noun. In both examples, the associative marker links two nouns but agrees with the head noun only. Agreement with the possessor is never possible, as demonstrated by the Swahili example adapted from Carstens (2000:334) below:

16. a. kiti cha m-toto
7chair 7.ASS 1.child
'the childs's chair'

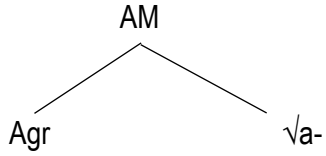
b. *kiti wa mtoto
7chair 1.ASS 1child

Intended: ‘the child's chair’

In (16a), where the agreement prefix agrees with the head noun, the expression is grammatical but when it agrees with the possessor as in (16b), the construction is ungrammatical.

The morphological structure of the AM can hence be represented as in (17), with concordial agreement represented by an agreement node (Agr-node) which combines with the root *a-*:

17.



In (17) above, the Agr-node is an inflectional element because it reflects the noun class (gender and number) features of the head noun (Carstens 1993, 1997, 2000, 2008). Therefore, if Assc-compounds were words like the VN-compounds discussed in Chapter 3, inflectional material would be part of these compound words.

Although the AM does not agree with the possessor, it forms a constituent with it. Even in languages such as Tswana, where the AM appears as an affix attached to the possessor noun, as shown in example (18), the AM functions as the head of the possessor phrase (Bentley, 1998). For example, in (18), the AM *wa-* and the possessor *motsoni* ‘hunter’ form a constituent.

18. mo-sadi w-a-mo-tsoni Tswana
1- woman 1-ASS-1-hunter
‘the wife of the hunter’
(Bentley, 1998:186)

Van de Velde (2013) notes that the AM is closely linked to the possessor in most Bantu languages as shown in the way it morphologically bonds with the possessor. This morphological link may explain why the possessor cannot move away from its position stranding the associative in Bantu languages (see section 4.3 below).

The AM does not always take the shape of *-a*. In some Bantu languages such as LoNkundo, the associative takes the form of *ka* in associative constructions. Welmers, (1963) observes that in Zulu

the associative *ka* is used with possessor nouns of NC 1a and 3a. These nouns mainly include proper names and kinship names (Bosch, 1997; Güldemann, 1999; De Dreu, 2008) as illustrated in the Zulu examples in (19). When the head noun belongs to the so-called weak noun classes (which include a nasal in their prefix), *ka* remains uninflected, (19a), but with head nouns from strong noun classes, *ka* is preceded by an agreement marker that refers back to the head noun in Zulu, (19b):

19. a. inja ka -baba
9.dog POSS father
'my father's dog'
(Bosch, 1997:2; Güldemann, 1999:160)

b. aba -fazi ba -ka -baba
2 -woman 2 -POSS -my father
'my father's wives'
(Güldemann, 1999:160)

In (19a) above, Bosch notes that the associative *ka-* is used to express personal possession, while in example (19b), Güldemann (1999) notes that the associative *ka* is an attributing element that is inserted between what he calls the nucleus (i.e. a head noun) and the satellite (i.e. the possessor), and reflects the NC of the head noun. Additionally, when the associative *ka-* comes before the possessor, the possessor has its initial vowel dropped as shown in the examples in (19) above (compare *ubaba*, 'father').

Güldemann adds that the associative *ka* is also used in other southern Bantu languages such as Tswana, where it is realised as *–ga-*. In Tswana, *–ga* is inserted between the possessive concord and the stem (see also Welmers, 1963). In another Bantu language, Mwera, *ka* occurs with the possessors of all noun classes except class 1a (Güldemann, 1999). It is important to indicate that although the associative marker *ka* is used in some Bantu languages, it is not as widespread as the associative marker *-a*.

The AM *-a* does not only express possessor relations but also conveys various other semantic associations between the preceding noun, the head noun, and the following noun in Bantu languages. According to Welmers (1963:433), in Swahili, the associative *–a* introduces several semantic aspects, some of which include material, time of use, function and possession as shown in the examples below.

20. a. Material
nyumba za mawe
'houses made of stone'

a. Function
miti ya kujengea
'sticks for building'

b. Time of use
chakula cha asubuhi
'food for morning'
'breakfast'

d Possession
kisu cha hamisi
'Hamisi's knife'

Matambirofa (2000) also indicates that the associative can express part-whole relations between the possessor and the head noun as shown in the Chichewa and Shona examples below.

21. a. mchira wa mbewa Chichewa
tail of mice'
'the mice's tail'
(Matamboirofa, 2000:72; see also Bresnan and Mchombo, 1987:774)

b. denga re-mba Shona
roof of–house
'the roof of the house'
(Matambirofa, 2000:75)

Kula (2012:436) and Delfitto et al. (2011:303-304) provide a detailed account of semantic associations introduced by the AM in Bemba. Some of the associations include possession, content and location/place as shown in some of the examples from Bemba below.

22 a. Possession

ulukasu	lwa	mfumu
11.hoe	11. ASS	9.chief

'the chief's hoe'

b. Content

ulupe	lwa	mbalala
11.basket	11.ASS	9.groundnuts

'basket of nuts/ basket for nuts'

c Location/place

ulukasu	lwa	pamunshi
11.hoe	11ASS	16.village

'hoe of the village'

(Kula, 2012:436; Delfitto, et al., 2011:306)

From the examples in (22) above, it is clear that a variety of semantic relationships between the head noun and the possessor can be expressed by the associative marker.

In conclusion, it has been shown that the AM in the associative construction links the possessor to the head noun and that the AM is widely used in Bantu languages. Further, the AM agrees in gender and number with the head noun. Therefore, the AM is an inflectional element.

4.3 The syntax of associative constructions

4.3.1 Complex DPs

Associative constructions are widespread in Bantu languages. They consist of an AM which is a linker between the possessor and the head noun. In (17) above, I represented the AM as consisting of an Agr-morpheme which bears the number and gender features of the head noun, and the root *-a*. This section discusses the syntax of complex DPs to establish the syntactic representation of associative

constructions. As my starting point, I adopt the analysis of genitive DPs presented in Abney (1987) and much subsequent work. According to Abney's (1987) DP-hypothesis, a nominal expression is analysed as a projection of a determiner, i.e. a determiner phrase (DP). A determiner is a functional element which takes the noun phrase as its complement. Abney also notes that although determiners lack descriptive content, they specify the referent of the noun phrase. While traditionally, only regular articles such as *a* and *the* were considered determiners (Sideeg 2016: 21), Abney (1987) considers articles, demonstratives, possessive pronouns and, importantly, also possessive/genitive markers as determiners that head DPs. I discuss the analysis of D-elements linked to possession in subsequent sections.

D-elements precede the head noun in English DPs. Abney (1987) notes that D-elements are heads of the DP and that they occupy the D position. Noteworthy is that in English, possessive pronouns are also D-elements and that they occupy the D position. As D-elements, possessive pronouns cannot co-occur with other D-elements as shown below:

23. a. my car
    b. *my the car
    c. *this my car
    d. *John's my car

The examples in (23b-c) above show that the possessive pronoun cannot co-occur with demonstratives and articles while (23d) shows that the possessive pronoun cannot co-occur with a possessive marker in possessive constructions. Abney (1987) argues that all these elements are determiners that are in complementary distribution (i.e. they occupy the same position) and cannot co-occur in English DPs.

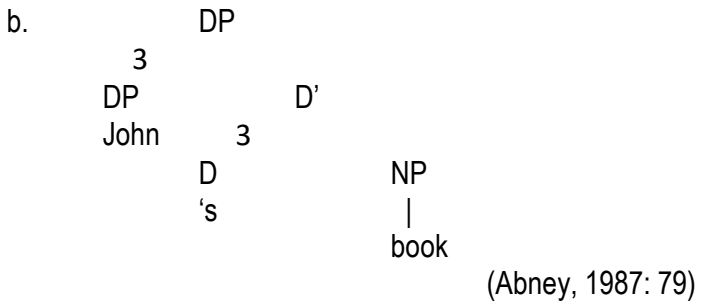
In genitive or possessive *'s* constructions in English, a noun occurs with another noun phrase (DP) that denotes possession, for example, *John's toy* or *the toy of John*. *John* is the possessor while the *toy* is the possessed noun. Abney (1987) notes that the possessive marker *'s* is a functional element that links the head noun to the possessor DP and also expresses possession.

One of the assumptions in the description of the possessive *'s* is that it is a functional element in English, thus, it is a determiner just like demonstratives and possessive pronouns. Consequently, it cannot co-occur with other determiners (Abney, 1986, 1987) as shown in (24b-d) below:

24. a. John's chair
    b. *John's the chair
    c. * the's chair
    d. *my's chair

Since according to Abney's DP-Hypothesis, determiners are heads of DPs, the possessive marker *'s* can also be analysed as the head of the DP in possessive constructions. It occupies the D position as exemplified below:

25. a. John's book

    b.


(Abney, 1987: 79)

Example (25b) above shows that the English possessive marker *'s* is the head of the whole DP. As a D-head, it precedes its complement, the possessed noun phrase, *book,* while the possessor *John* is the specifier. Abney notes that the possessive *'s* as depicted in (25b) above should not be analysed as an affix that attaches to the noun *John*, but that it is a determiner which cliticises to the right periphery of its specifier, the whole possessor DP *John*. When the possessor DP is more complex, the possessive marker attaches to the last word of the DP, as in for example, [*the man over there*]*'s hat*. The entire complex possessor DP is the specifier while the possessive *'s* is the head.

Although the analysis of the possessive *'s* as the D head looks elegant and fits into the DP-analysis, it does not easily carry over to Bantu languages where demonstratives and possessives can co-occur in possessive constructions.

4.3.2 The syntax of Bantu associative nominal phrases - Carstens (2018)

Based on the information presented above on complex DPs, I now provide an analysis of the syntactic structure of possessives/associative constructions in Bantu languages. The main aim of this section is

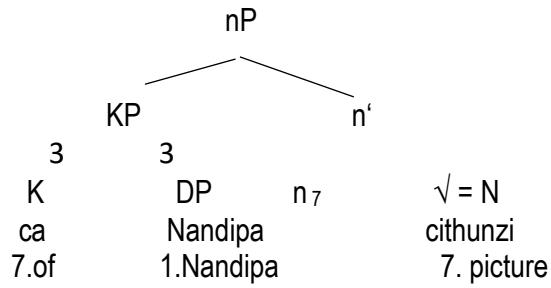
to show where the possessor originates and how the word order of the head noun and possessor is derived. I refer mainly to Carstens's (2018) analysis of associative constructions.

Carstens (2008, 2018) assumes that possessors are not base-generated in the specifier of DP (as in Abney's proposal), but that they merge as specifiers of the "little" n which takes the root of the semantic head of the DP as its complement (see Chapter 3). Now recall that possessors in Bantu are introduced by the associative marker *–a*, which agrees in noun class with the head noun, as shown in the Cinyanja example below:

26.	a.	ci-thunzi	ca	Nandipa[12]
		7-picture	7.ASS	1.Nandipa
		'Nandipa's picture'		

Carstens (2018) argues that the associative marker is the head of the possessor phrase. According to Carstens (2018), the associative marker is a case marker of category K (Kase) and assigns case to its complement, the possessor DP (see also Carstens, 1997, 2008). When the complement of the associative, the possessor DP, merges with the associative (K), they form a KP structure as shown in (27). This KP sits in the specifier of the nP which categorises the root noun as N:

27.



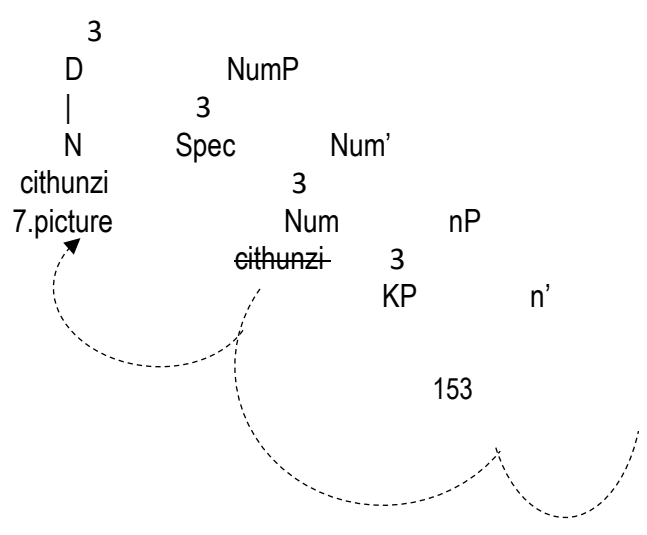
(Carstens, 2018:19)

[12] A similar example is given in Carstens (2018:2) from Chichewa (i.e. *chithunzi cha Lucy*). Chichewa and Cinyanja are sister languages, and most of the words have the same meaning and pronunciation. However, there are slight dialectal variations in the orthography. For example, in Chichewa, nouns in class 7 take the prefix *chi-* e.g. *chithunzi* 'picture', c*hipande* 'calabash', *chikho* 'cup' while in Cinyanja, class 7 nouns take the prefix *ci-* e.g. *cithunzi* 'picture', *cipande* 'calabash, *cikho* 'cup' etc. in Zambia. Therefore, some Chichewa examples will be replaced by the Cinyanja spelling system.

In (27) above, the possessor DP (PossDP) *Nandipa* merges with the associative marker. Remember that the associative marker itself is morphologically complex, consisting of a root √-*a* plus gender and number agreement features. The noun class contributed by Num and gender n is transferred to the specifier (via Spec-Head agreement), and therefore is realised on the head of KP. In (27), I ignore the internal structure and represent the inflected associative marker as the head of the KP. The syntax of KP, therefore, leads to an analysis of the associative marker which is notably different from Abney's (1987) analysis of the possessive *'s* as a determiner and head of the genitive DP. The associative in Bantu is not the determiner head of the whole construction; rather, the associative (K) is the functional head of the possessor KP, which forms a specifier of the nP of the associative construction, as shown in (27) above.

(27) does not yet provide the correct word order of associative constructions in Bantu as discussed in section 4.2 above, in which the head noun precedes the associative. This word order follows from Carstens' (2008, 2018) analysis of word order in DPs in Bantu languages that I adopted in Chapter 3. Bantu DPs are N-initial. In Carstens' analysis of complex DPs, the head noun undergoes head-movement and moves to D. Therefore, a simple DP architecture of the Cinyanja associative nominal phrase (ANP), *cithunzi ca Nandipa* 'Nandipa's picture' in (26) can be represented by the structure in (28).

28.



5 3 ca

n √N

Nandipa ~~cithunzi~~ ~~cithunzi~~

The structure in (28) shows that in Cinyanja and other Bantu languages, the head N moves to D via n and Num and eventually occupies the D position. This is different from English, where D is occupied by determiners and determiner-like elements.

Since the head noun moves to D in Bantu languages, determiners cannot appear in this position, but occupy different sites. Carstens (2000) notes that when determiners such as demonstratives appear DP-internally in associative constructions in Bantu languages, they precede KP and follow the head noun. I demonstrate this using the Cinyanja examples below.

29. ci-panda i-ci ca mowa
    7-calabash 7-this 7.ASS 3.beer
    'this calabash of beer'

The example in (29) shows that when the demonstrative precedes the associative phrase in ANPs, it modifies only the head noun, not KP, and both the demonstrative and the associative agree with the noun class of the head noun. Carstens (2008) argues that in Swahili, the demonstratives are adjuncts to NumP. However, for reasons that I make clear below, I analyse demonstratives in Cinyanja as adjuncts to nP. In whichever case, the demonstrative occupies a position between D and nP in the DP.

Another modifier, the possessive pronoun, like the demonstrative, precedes the associative KP and follows the head noun in ANPs in Cinyanja. I illustrate this word order with the example below:

30. ci-panda c-anga ca mowa
    7-calabash 7-my 7.ASS 3-mowa
    'my calabash of beer'

The example in (30) above shows that when a possessive pronoun modifies the head noun, it immediately follows it.

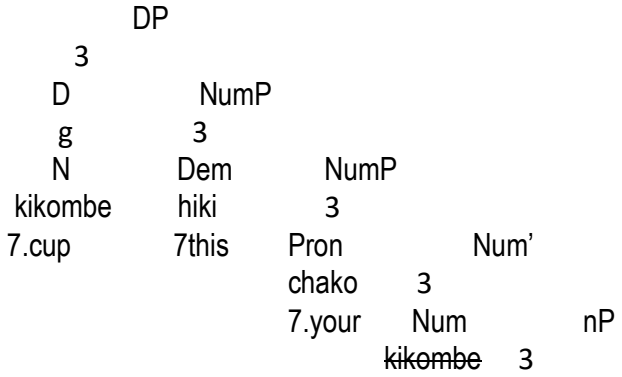
The example shows that possessive pronouns in Bantu are different from possessive pronouns in English, which cannot co-occur with full possessor phrases, and which Abney's (1987) analysis treats like other determiners, i.e. as D-heads. For Carstens (1991, 2018), possessive pronouns are not D-heads but phrasal constructions originating in a specifier. Carstens argues that this specifier is Spec, NumP in Bantu languages.

In Bantu ANPs, a possessive and demonstrative can co-occur (Carstens, 1991, 2008; Mchombo, 2004; Rugemalira, 2007; Lusekelo, 2009). In Swahili ANPs, demonstratives precede possessive pronouns as shown in the example below:

31. a. kikombe hiki chako cha kahawa
7.cup 7.this 7.your 7.ASS 9.coffee
'this cup of coffee of yours'

b. *kikombe chako hiki cha kahawa
7.cup 7.your 7.this 7.ASS 9.coffee
'this cup of coffee of yours'
(Carstens, 1991:105)

The examples in (31) above show that when the demonstrative precedes the possessive pronoun and immediately follows the head noun, the associative construction is grammatical in Swahili. But when the demonstrative follows the possessive pronoun, the expression is ungrammatical. Recall that Carstens (1991) considers demonstratives as a special type of adjective, and analyses them as adjuncts[13] to a NumP. Therefore, the word order in (31a), and the ungrammaticality of (31b), follow from this analysis:

32.


[13] In Carstens (2008), however, demonstratives are associated with a category XP between D and NumP

KP n‘
5 3
cha kahawa n √N
7.of 9.coffee ~~kikombe~~ ~~kikombe~~

Carstens (2008) also notes that demonstratives are flexible in Swahili and can precede the head noun by moving to Spec, D. When demonstratives co-occur with possessor pronouns in Swahili only demonstratives undergo optional raising to Spec D as shown in (33b) which represents example (33a).

33. a. hiki kikombe chako cha kahawa
7.this 7.cup 7.your 7.ASS 9.coffee
‘this cup of coffee of yours’

b. DP
3
Dem D'
hiki 3
7.this D NumP
g 3
N Dem NumP
kikombe ~~hiki~~ 3
7.cup 7this Pron Num’
chako 3
7.your Num nP
~~kikombe~~ 3
KP n’
5 3
cha kahawa n √N
7.of 9.coffee ~~kikombe~~ ~~kikombe~~

(see Carstens, 2008:154)

In Chichewa, on the other hand, Mchombo (2004) notes that when a possessive pronoun and a demonstrative co-occur DP-internally in an associative construction, the possessive pronoun must immediately follow the noun while the demonstrative immediately precedes KP. This shows that in Chichewa (and similarly in Cinyanja), the possessive must precede the demonstrative as shown in the Cinyanja example below:

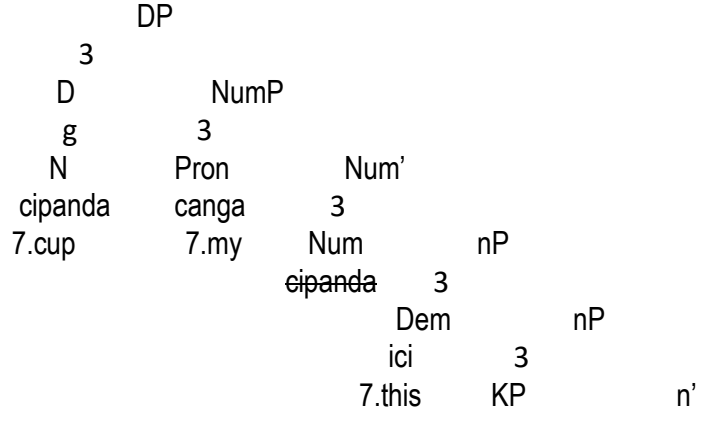
34. ci-panda c-anga i-ci ca mowa
7.calabash 7.my 7.this 7.ASS 3.beer
‘this calabash of mine of beer’

The example in (34) above, shows that the head noun precedes the possessive, which precedes the demonstrative. However, in example (35) below, where the demonstrative precedes the possessive pronoun in Cinyanja, the word order is marked and only acceptable with contrastive focus on the demonstrative which is possibly derived by focus movement.

35. ci-panda i-ci c-anga ca mowa
7.calabash 7.this 7.my 7.ASS 3.beer
'This calabash of mine of beer'

For the purpose of this study, I will follow Mchombo in assuming that the basic unmarked word order is poss > Dem in Cinyanja, and that I ignore the marked word order in (35). As noted above, Carstens (1991) derives the Swahili word order by adjoining the demonstrative to NumP, and optionally allows it to move to Spec, D. Since the head noun occupies the D-position, and the possessive pronoun is in Spec, NumP, this produces the Swahili word order. However, since the demonstrative must follow the possessive pronoun in Cinyanja, I argue that it is lower in the structure than the possessive pronoun and that it is adjoined to nP. This shows that the position of the demonstrative in Bantu languages is not fixed but varies according to the syntax of each language:

36.



5 3
ca mowa n √N
7.of 3.beer ~~cipanda~~ ~~cipanda~~

Demonstratives in Cinyanja can also precede the head noun, like in Swahili:

37. i-ci ci-panda c-anga ca mowa
7.this 7.calabash 7.my 7.ASS 3.beer
'this calabash of mine of beer'

The example in (37) above suggests that in Cinyanja, demonstratives can also take an optional position in Spec, D. Carstens (2008) posits that when the demonstrative takes the prenominal position, it has a definite interpretation similar to determiners.

Demonstratives and possessive pronouns can also follow the KP[14] in Cinyanja as shown in the examples below.

38. a. ci-panda ca mowa ici
7-calabash 7.ASS 3.beer 7. this
'this beer-calabash'

b. ci-panda ca mowa c-anga
7-calabash 7.ASS 3.beer 7-my
'my beer-calabash'

(38a) above shows that the demonstrative can follow KP, while (38b) shows the same for the possessive pronoun. Syntactically, I argue that these post-nominal modifiers are right-adjoined to DP. For example, *cipanda ca mowa ici* 'this beer-calabash' in (38a) can be represented as in (39).

39. DP
3
DP Dem(P)
6 ici
cipanda ca mowa

[14]There is a slight difference in the meaning interpretation between the modifiers preceding KP and those following KP in ANPs. When the modifiers follow KP as in (38), they modify the entire ANP but when they precede KP and immediately follow the head noun as in (29) and (30), they modify the head noun. The example in (29), i.e. *cipanda ici ca mowa* can be interpreted to mean 'this calabash which is used for beer' while *cipanda ca mowa ici* in (38a) can mean 'this beer-calabash'. Similarly, *cipanda canga ca mowa* in (30) can be interpreted as 'my calabashi which is used for beer' while *cipanda ca mowa canga* in (38b) would mean 'my beer-calabash'.

Further, demonstratives and possessives can also occur simultaneously after KP as shown below;

40. a. ci-panda ca mowa c-anga ici
7-calabash 7.ASS 3.beer 7-my 7-this
'this beer-calabash of mine'

b. *ci-panda ca mowa ici c-anga
7-calabash 7.ASS 3.beer 7-this 7-my
Intended: 'this beer-calabash of mine'

The examples in (40) show that when the demonstrative follows the possessive pronoun after KP; the expression is grammatical but when the demonstrative precedes the possessive after the possessor, the expression is ungrammatical. This seems to suggest that there are also restrictions in the ordering of the demonstrative and the possessive even when they follow the KP, which do not follow from an analysis of both of these modifiers as DP-adjuncts. However, since the precise analysis of the word order contrast in (40) is not directly relevant to my study, I leave the question of how to explain this contrast for future research.

The information above shows that demonstratives are flexible in Cinyanja; they can appear in Spec, D, they can be adjoined to nP and can also be right-adjoined to DP. Quantificational modifiers also behave just like demonstratives in terms of word order in Cinyanja as shown below:

41. a. **onse** ma-tumba a cimanga
all 6-bag 6-my 6.ASS 7.maize
'all the bags of maize'

b. ma-tumba **onse** a ci-manga
6-bag 6.all 6.ASS 7-maize
'all my bags of maize'

c. ma-tumba a cimanga **onse**
6-bag 6.ASS 7.maize 6.all
'all the bags of maize'

The examples in (41) show that the quatifier *onse* can appear in Spec D, can be adjoined to nP as well as be right adjoined to DP.

Adjectives, like demonstratives and possessives, also precede possessor KPs in ANPs when they modify the head noun in Bantu languages. In Cinyanja ANPs, the adjective occurs to the left of the associative and agrees with the head noun (Carstens, 1997, 2018). In (42a & b) below, the adjectives *cabwino* 'nice' in (42a) and *yatsopano* 'new' in (42b) appear to the left of the associative.

42.	a.	ci-thunzi	c-abwino	ca	Nandipa.
		7-picture	7-nice	7.ASS	1.Nandipa
		'a nice picture of Nandipa'.			
	b.	njinga	y-atsopano	ya	Thembi
		9.bicycle	9-new	9.ASS	1.Thembi
		'the new bicycle for Thembi'.			

Carstens analyses the adjective as a second specifier of *n* in Bantu. This specifier is higher than the specifier which hosts the possessor-KP, which explains that adjectives precede possessor phrases headed by the associative marker. It also explains that when the adjective co-occurs with the demonstrative in an associative construction in Cinyanja, the demonstrative (an nP-adjunct) precedes the adjective as illustrated below.

43.	a.	Njinga	i-ja	y-atsopano	ya	Thembi
		9.bicycle	9-that	9-new	9.ASS	1.Thembi
		y-a-sow-a.				
		9-BE-lose- FV				
		'That new bicycle of Thembi, is lost'.				
	b.	*Njinga	y-atsopano	i-ja	ya	Thembi
		9.bicycle	9-new	9-that	9.ASS	1.Thembi

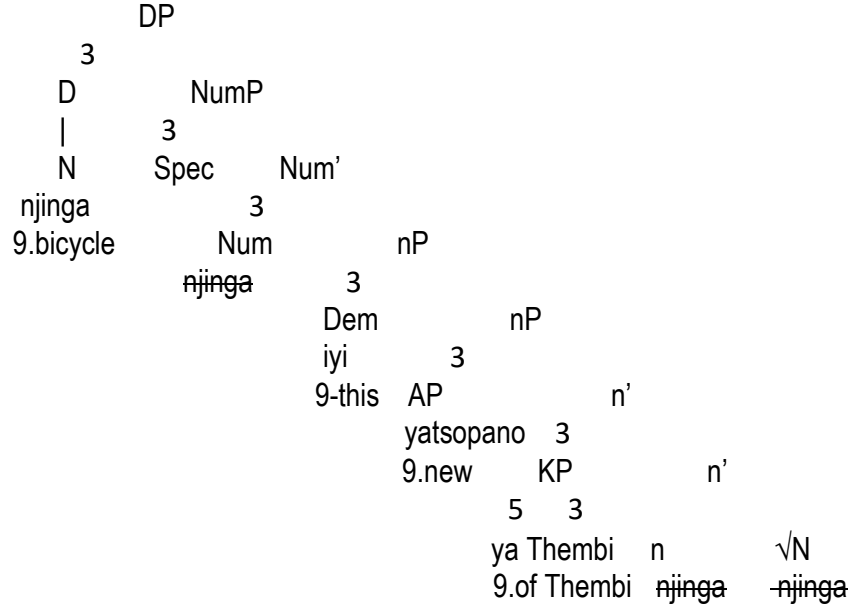
y-a-sow-a.

9-BE-lose- FV

Intended: 'That new bicycle of Thembi is lost'.

(43b) above shows that when the adjective precedes the demonstrative in ANPs, the expression is ungrammatical. This follows from the claim that in Cinyanja, the adjective is a higher specifier of n while the demonstrative is an adjunct. KP, on the other hand, is the lower specifier of the head noun:

44.



Adjectives, like demonstratives and possessives, can also follow KP. This shows that the possibility of right adjunction to DP is generally available for all types of modifiers including adjectives:

45. njinga ya Thembi y-atsopano

9.bicycle 9.ASS 1.Thembi 9-new

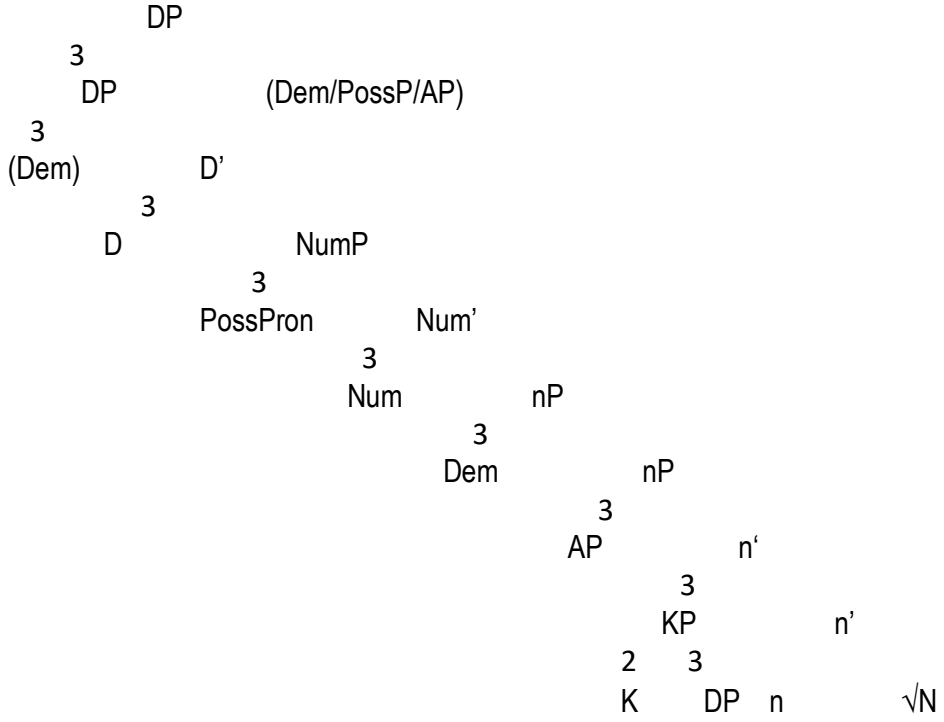
'Thembi's new bicycle'

Finally, as the above tree diagrams show, the N-initial word order of an ANP in Bantu languages is realised when N moves to D, from where it precedes demonstratives, possessive pronouns, adjectives, and possessor-KPs. This movement is triggered by the fact that in Bantu languages, the gender class of the nominal root is marked on n, not D. For gender to be visible, and for the modifiers and the associatives to obtain their phi-features, N has to move to D.

In sum, Bantu associative constructions are N-initial because of N-to-D movement. DP internally, the head noun in the ANPs can be modified by adjectives, possessives and demonstratives/quantifiers. The adjectives are higher specifiers of n, the possessive pronoun is a specifier of Num while the

demonstrative is either an adjunct to NumP (as in Swahili), or an adjunct to nP (as in Cinyanja). Demonstratives can also be moved to Spec, D, and all types of modifiers can be right-adjoined to DP. Non-pronominal possessor phrases are KPs headed by the associative marker, which takes the possessor DP as its complement. A generic tree diagram with all possible modifier positions in Cinyanja is shown in (46).

46.



4.3.3 The KP structure inside Assc-compounds

In this section, I show that the two types of Assc-compounds in Cinyanja composed of the head noun and the possessor linked together by an associative marker, form a KP phrasal structure. I also explain how the agreement between the head of the main DP and KP is established in Assc-compounds. I show that the head noun as the head of the compound shares agreement features with KP through K, the head of KP. This implies that the syntactic structure of the Cinyanja Assc-compounds is the same as that of the Bantu ANPs discussed above.

In Cinyanja associative compounds, like in other Bantu languages, the possessor is on the right and is the complement of the associative. This is shown in the Cinyanja Assc-compounds below:

47. a. mw-ana wa leza
1-child 1.ASS 1.god
'rainbow'

b. buku la nyimbo
5.book 5.ASS 9.song
'songbook'

In (47a) above, the Assc-compound *mwana wa leza* 'rainbow' consists of *mwana* 'child' as the head noun and *wa leza* 'of god' as the complement of the head noun. Similarly, in (47b) the head noun is *buku* 'book' and the complement is *la nyimbo* 'of songs'. In the examples, the associative marker (K) introduces the possessor and is the head of the associative phrase (KP). This shows that the syntactic structure of the two types of Assc-compounds in Cinyanja consists of a head noun and KP. I illustrate Type 1 Assc-compounds, in Table 4.1.

Table 4.1: Phrasal structure of Type 1 Assc-compounds

S/N	COMPOUND	HEAD NOUN	KP	TRANSLATION OF KP
a	book la mapemphero 'prayer book'	buku book	la ma-pemphero ASS PossDP	of prayers
b	mwana wa sukulu 'pupil'	mwana child	wa sukulu ASS PossDP	of the school
c	mwana wa masiye 'orphan'	mwana child	wa masiye ASS PossDP	of the deceased
d	buku la nyimbo song book	buku book	la nyimbo ASS PossDP	of songs

e	pepala la mayeso examination paper	pepala paper	la mayeso ASS PossDP	of examination
f	Njinga ya moto Motorbike	njinga bicycle	ya moto ASS PossDP	of fire

The examples in Table 4.1 show that the combination of associative (K) and PossDP forms KP. Similarly, Type 2 Assc-compounds also comprise the head noun followed by a KP structure as shown in table 4.2 below:

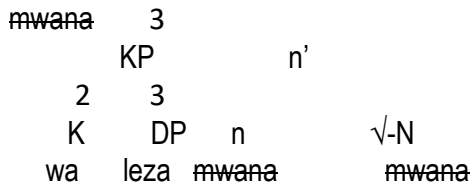
Table 4.2: Type 2 Assc- compounds with a KP inside

S/N	COMPOUND	HEAD NOUN	KP	TRANSLATION OF KP
a	madzi a moyo 'oral-rehydration'	madzi water	a moyo ASS PossDP	of life
b	wailesi ya kanema 'TV'	wailesi radio	la kanema ASS PossDP	of pictures
c	mwana wa m'thengo 'bastard'	mwana child	wa m'thengo ASS PossDP	of the bush
d	mw-ana wa leza 'rainbow'	mwana child	wa leza ASS PossDP	of god

The Assc-compounds in Tables 4.1 and 4.2 above show that the [ASS + PossDP] form KP and K is the functional head of the KP. This KP takes the specifier position of nP while the root head categorised as N is base generated and shares the feature of n (i.e. gender features) with the head of KP. Recall, that morphologically, the associative marker (K) in associative nominals is a composition of the root √-a with gender and number agreement features of the head noun. Similarly, Assc-compounds have the same structure as shown in (48b) which represents (48a):

48. a. mw-ana wa leza
1-child 1.ASS 1.god
'rainbow'

b. DP
3
D NumP
mwana 3
Num nP



The structure in (48b) shows that when the root √-N moves to n and to Num and then to D, it gets its gender and number features which are spelled out as the noun class prefix in Num. The head noun shares these features with K which KP bears as agreement features. This is the same syntactic structure discussed in subsection 4.3.2 above. Therefore, I argue that Cinyanja Assc-compounds have the same structure as other Bantu ANPs, and can also be represented by Carstens' proposal.

4.4 Assc-compounds as DPs

From the examples in section 4.3 above, the Assc-compounds in Cinyanja seem to be built on the same syntax as the ANPs. Just as with VN-compounds, which include a syntactic VP built according to syntactic rules, Assc-compounds are phrasal compounds and include a syntactic phrase, the KP. The question that arises, and which I will address in the remainder of this chapter, is whether Assc-compounds are bare nouns which include a "reified" complex associative construction (comparable to the analysis proposed for VN-compounds in Chapter 3), or whether Assc-compounds are genuine phrasal constructions, i.e. DPs.

I begin this section by discussing the Assc-compounds in their syntactic context. I show that the Assc-compounds with an internal phrasal structure as shown in section 4.3 above, appear in the same syntactic context as other nominal expressions in Cinyanja and can participate in the same syntactic operations. For example, the Assc-compounds can take the subject and object positions in simple sentences and can act as head nouns in relative clauses. I then discuss Assc-compounds in combination with other DP-internal modifiers, such as adjectives and demonstratives. As I will demonstrate, the syntactic behaviour of Assc-compounds differs from regular associative constructions in that it is not possible to interrupt the linear order of head noun and associative KP in the former, while this is possible with the latter.

4.4.1 Assc-compounds as subjects and objects of the verb

In a basic clause structure, Assc-compounds with an internal structure which includes a KP can function as subjects of a verb, and agree in the class indicated by the compound's noun class. The

compound's noun class is determined by the head noun as in (49a). The compound can also function as a nominal object as in (49b):

49. a. **Pepala** **la** **ma-yeso** l-a-sow-a.
5.paper 5.ASS 6-examination 5.SM-BE-lose-FV
'The examination paper is lost'.

b Mw-ana w-a-ng'amb-a **pepala la ma-yeso.**
1-child 1. SM-PRF-tear-FV 5.paper 5.ASS 6-examination
'The child has torn the examination paper'.

In (49a) the subject is *pepala la mayeso* 'examination paper' of NC 5, and the NC of the compound is reflected in the verb by the class 5 subject agreement marker. In (49b), *pepala la mayeso* is the object of the verb.

Similar to the VN-compounds discussed in Chapter 3, an OM corresponding to an Assc-compound can also be prefixed to the verb. The inclusion of the OM allows right and left dislocation of the entire compound in the clause (see Bresnan and Mchombo, 1987; Mchombo, 2004) as demonstrated in the following examples.

50. a. Njoka ya-kuda i-na-**mu**-lum-a **dzulo**
9.snake 9.ASS-black 9.SM-PST-1.OM-bite-FV yesterday
m-misiri **wa** **nyumba.**
1-expert 1.ASS 10.house
'A black snake bit him yesterday, the builder'.

b. Njoka ya-kuda, **m-misiri** **wa** **nyumba,**
9.snake 9.ASS-black 1-expert 1.ASS 10.house
i-na-**mu**-lum-a dzulo.
9.SM-PST-1.OM-bite-FV yesterday
'A black snake, the builder, it bit him yesterday'.

c. **M-misiri** **wa** **nyumba,** njoka ya-kuda
1.expert 1.ASS 10.house 9.snake 9.ASS- black
i-na-**mu**-lum-a dzulo.
9.SM-PST-1.OM-bite-FV yesterday

'The builder, a black snake bit him yesterday'.

From the examples above, it is clear that the Assc-compounds can be object marked and dislocated like other objects. It is, however, not possible to separate the associative KP from the head noun:

51. a. ***M-misiri**, njoka ya-kuda **i**-na-**mu**-lum-a
1-expert 9.njoka 9.ASS-black 9.SM-PST-1.OM-bite-FV
wa **nyumba** dzulo**.**
1.ASS 10.house yesterday
.*'The expert, a black snake bit him, of the houses, yesterday.

b. ***Wa** **nyumba,** dzulo, njoka ya-kuda
1.ASS 10.house yesterday 9.snake 9.ASS-black
i-na-mu-lum-a **m-misiri.**
9.SM-PST-1.OM-bite-FV 1-expert
*'Of the house, yesterday, the black snake bit him the expert.

(51) shows that part of the compound, the head noun *mmisiri*, cannot be left dislocated in (51a) or right dislocated in (51b). This is also not possible for regular associative constructions:

52. a. **M-phika** **wa** **ndiwo** mw-ana wa-m-ng'ono
3-pot 3.ASS 9.relish 1-child 1.ASS-1-small
a-na-**wu**-pwany-a mwangozi.
1.SM-PST-3.OM-break-FV accidentally
'The pot of relish, a small child broke it accidentally.

b. ***M-phika,** mw-ana wa-m-ng'ono a-na-**wu**-pwany-a
3-clay-pot 1-chid 1.ASS-1-small 1.SM-PST-3.OM-break-FV
wa **ndiwo** mwangozi.
3.ASS 9.relish accidentally
*'The pot, the small child broke it of relish accidentally'.

c ***Wa** **ndiwo** mw-ana wa-m-ng'ono
3.ASS 9.relish 1-child 1.ASS-1-small
a-na-wu-pwany-a **m-phika** mwangozi.
1.SM-PST-3.OM-break-FV 3-clay-pot accidentally

*'Of relish, the small child broke it the clay-pot accidentally'.

The impossibility of separating the head noun from the associative KP follows from the syntax of associative constructions motivated above, according to which the head noun does not correspond to a separate phrasal constituent that could be independently moved.

The above examples show that Assc-compounds behave like regular associative DPs in these examples. What they do not show is whether the compound is represented by the phrasal syntax of regular associative constructions, or whether it is a simple noun which includes a reified DP and which has then projected its own DP. I address this question in section 4.5 below.

4.4.2 Assc-compounds as subject and object of a relative clause

Assc-compounds can also function as head nouns of subject or object relative clauses. In (53a) the compound word *njinga ya moto* 'motor-bike' is the subject of the clause that has been used as a base example from which the relative clauses in (53b) and (53c) are drawn. In (53b), the Ass-compound *njinga ya moto* 'motor-bike' is the subject of the (passivised) relative clause; in (53c) *njinga ya moto* is the object of the clause.

53. a. Njinga ya moto y-a-gwa pa-bwalo.
9. bicyle 9.ASS 3.fire 9.SM-PRF-fall-FV 16.LOC -5.outside
M-limi a-na-gul-a njinga ya moto.
1-farmer 1.SM-PST-buy-FV 9.bicycle 9.ASS 5.fire
'The motor-bike has fallen outside. The farmer bought the motor-bike.

b. **Njinga ya moto** i-mene i-na-gulidw-a
9.bicyle 9.ASS 3.fire 9-REL 9.SM-PST-buy-PASS-FV
ndi m-limi y-a-gw-a pa-bwalo.
by 1.farmer 9.SM-PRF-fall-FV 16.LOC-5.outside
'The motor-bike which was bought by the farmer has fallen outside'.

c. **Nyinga ya moto** i-mene m-limi a-na-gul-a
9.bicycle 9.ASS 3.fire 9-REL 1-farmer 1.SM.PST-buy-FV
y-a-gw-a pa-bwalo.
9-PRF-fall-FV 16.LOC-5.outside.

'The motor-bike which the farmer bought has fallen outside'.

The following examples show that regular associative constructions can also be modified by relative clauses:

54. a. **Thumba** **la** **nshawa** li-mene li-na-gulidw-a
5.bag 5.ASS 5.groundnuts 5-REL 5-PST-buy-PASS-FV
ndi m-phunzitsi l-a-ng'ambik-a.
by 1.teacher 5.SM.BE.PRS-tear-FV
'The bag of groundnuts which was bought by the teacher is torn'.

b. **Thumba** **la** **nshawa** li-mene m-phunzitsi
5.bag 5.ASS 5.groundnuts 5-REL 1-teacher
a-na-gul-a l-a-ng'ambik-a.
1-PST-buy-FV 5.SM-BE.PRS-tear-FV
'The bag of groundnuts which the teacher bought is torn'.

Therefore, there is again nothing special about the syntactic environment in which Assc-compounds can occur that would distinguish them from regular associative constructions.

4.4.3 Coordination

Assc-compounds with an associative phrase are considered DPs because they can be coordinated with other DPs as subjects and objects as shown in (55a) and (55b).

55. a. **M-lezi** **wa** **ana** ndi m-silikali a-na-gul-a
1-keeper 1.ASS 2.child and 1-soldier 2.SM-PST-buy-FV
nkhuku dzulo.
9.chicken yesterday
'The nanny and the soldier bought chickens yesterday'.

b. Baibulo kapena **buku** **la** **nyimbo**
5.bible or 5. book 5.ASS 10.song
l-a-sow-a.
5.SM-BE.PRS-lose-FV

'The bible or songbook is lost'.

The examples in (55) above show that the Assc- compounds can be coordinated with other DPs. This is evidence that the Assc-compounds with a KP internal structure are also DPs.

4.4.4 Modification by adjectives

Assc-compounds, like other nouns, can also be modified with adjectives. As compounds, modification affects the entire compound word, not only one of the nouns as shown below:

56.	a.	Njinga	ya	moto	ya-i-kulu	y-a-gw-a
		9.bicycle	9.ASS	3.fire	9-ASS-9-big	9.SM-PRF-fall-FV
		pa-nsi.				
		16-LOC-down				
		'The big motor-bike has fallen'.				
	b.	*Njinga	ya	moto	wa-u-kulu[15]	ya-gw-a.
		9.bicycle	9.ASS	3.fire	3.ASS-3-big	3.SM-PRF-fall-FV
		Intended: 'The big motor-bike has fallen'.				

(56b) is ungrammatical with the intended compound reading because the adjective cannot modify the PossDP which is a part of the compound word. It can only modify the entire compound word, in which case, the agreement must be with the head noun. This example contrasts with the behaviour of adjectives that follow regular associative constructions. In an example like (57), the adjective can modify the preceding PossDP, and does not have to refer to the whole associative DP:

57.	a.	Mi-pando	ya	ana	a-a-ng'ono	y-a-psy-a.
		4.chair	4.ASS	2-child	2.ASS-2-small	4.SM-BE.PRS-burn-FV
		'The chairs of small children are burnt'.				
	b.	Ma-tumba	a	cimanga	co-fiira	
		6.bag	6.ASS	7.maize	7.ASS-red	
		a-a-sow-a.				
		6.SM-BE.PRS-lose-FV				

[15] I assume (56b) would be accepted with the literal but nonsensical meaning of "the bicycle of big fire'

‘Bags of red maize are lost’.

At first sight, the difference between (56b) and (57) may be interpreted as evidence that the Assc-compound is actually a complex N with a reified DP because the internal structure of the DP can no longer be accessed by a modifier, in contrast to regular DPs. However , I believe that (56b) should be ruled out on semantic rather than syntactic grounds. In this example, modification of the KP-part of the compound simply makes it impossible to retrieve the idiomatic interpretation of the compound. Importantly, in section 4.5 below I present examples where an adjective can, in fact, modify the KP-part of an idiomatic Assc-compound. Therefore, even though the impossibility of (56b) seems to provide evidence for a reification analysis of Assc-compounds, we have to wait and see the full picture before we can draw a conclusion about the syntax of these compounds.

4.4.5 Possessor modifiers

I showed in section 4.3.2 above that regular associative constructions in Cinyanja can also be modified by possessive pronouns. As shown in (30), repeated here for convenience, the possessive pronoun intervenes between the head noun (in D) and KP, or follows the whole construction:

58.	a.	ci-panda	c-anga	ca	mowa
		7-calabash	7-my	7.ASS	3-beer
		‘my calabash of beer’			

	b	ci-panda	ca	mowa	c-anga
		7-calabash	7.ASS	3-beer	7-my
		‘the beer-calabash’			

Assc-compounds can also be modified by possessive modifiers, which are also post-nominal and agree in noun class with the head noun. However, as (59) shows, with Assc-compounds, these modifiers must appear after the KP and cannot intervene between the head noun and KP:

59.	a.	Wailesi	ya	kanema	**y-anga**
		9. radio	9.ASS	12.picture	9.my
		y-a-pwanyik-a.			

9.SM.-BE.PRS-break
'My television set (TV) is broken'.

b. Wailesi **y-anga** ya kanema y-a-pwanyik-a.
9. radio 9-my 9.ASS 12.picture 9.SM.-BE.PRS-break
NOT: 'My television set (TV) is broken'
Possible only with literal meaning: 'My radio of pictures is broken.'

In (59a) above, the possessive pronoun *yanga* 'my' agrees in class with the compound word *wailesi ya kanema* 'television set'. This word order, with the possessive pronoun following the KP is the only possible one for Assc-compounds. However, when the possessive pronoun is located in the specifier of NumP and therefore appears between *wailesi* 'radio' and *ya kanema* 'of pictures' as in (59b), it does not modify the entire associative expression but modifies only the head noun. As a result, the idiomatic interpretation of 'television' is no longer available.

4.4.6 Demonstratives

A similar observation to the above can be made regarding demonstratives. Examples (29) and (38a) from section 4.3.2, repeated below in (60), show that demonstratives can intervene between the head noun and the KP, (60a), but also follow the KP (60b):

60. a. ci-panda i-ci ca mowa
7-calabash 7-this 7.ASS 3.beer
'this calabash of beer'
b. ci-panda ca mowa i-ci
7-calabash 7.ASS 3.beer 7-this
'this beer-calabash

Assc-compounds can also be modified by demonstrative pronouns, but as with adjectives and possessor pronouns, the demonstrative cannot interrupt the order head noun-KP. Instead, it must follow the KP:

61 a. Mw-ana wa sukulu **u-yu** ndi wa-ulesi.
1-child 1.ASS 5.school 1-this COP 1.ASS-lazy
'This pupil is lazy'.

b. *Mw-ana **u-yu** wa sukulu ndi wa-ulesi.
1-child 1-this 1.ASS 5.school COP 1.ASS-lazy
'This child of school is lazy'.

62. a. Buku la nyimbo **i-li** ndi lo-dula
5.book 5.ASS 10.song 5-this COP 5-expensive
'This songbook is expensive'.

b. *Buku **i-li** la nyimbo ndi lo-dul
5.book 5-this 5.ASS 10.song COP 5-expensive
'This book of songs is expensive'.

The examples in (61b) and (62b) above are not ungrammatical but do not allow idiomatic reading. They are grammatical with a literal interpretation. For example, (61b) can mean 'This child of school is lazy' but not 'This pupil is lazy' as in (61a). Similarly, (62b) can mean this book of songs is expensive' not 'This songbook is expensive' as in (62a). This difference in interpretation caused by the position of the demonstrative seems to suggest that the associative constructions with a KP internal structure are words because the combination of head noun and associative KP cannot be interrupted by demonstratives, in contrast to the word order of regular associative constructions shown in example (29). Instead, in Assc-compounds, the only acceptable word order with demonstratives is the one in (39) where the modifier follows the KP. The question that arises then is whether Assc-compounds should be analysed as words which are built based on a reified syntactic associative phrase, or whether the differences between these phrasal compounds and regular associative constructions illustrated in (59)-(62) are the result of their lexical semantics. The next section on the lexical integrity hypothesis attempts to answer this question.

4.5 Assc-compounds and the lexical integrity hypothesis

As indicated in section 3.5 of Chapter 3, the LIH stipulates that morphological and syntactic rules are different, hence pieces of words cannot be manipulated by syntactic operations and are not accessible to elements external to the word (Lapointe, 1980; Di Sciullo and Williams, 1987; Anderson, 1992; B & M, 1995; Ackema and Neeleman, 2004; Booij, 2009; Lieber, 2009b).

In this section, I use some Lexical Integrity tests to show that parts of Assc-compounds are in fact accessible by syntax; they can be pronominalised, and modified by adjectives, demonstratives, quantifiers, possessor modifiers and relative clauses. The compound constituents can also be gapped. The main exception is their behaviour with respect to extraction of constituent parts, where my results show that parts of these complex expressions cannot be extracted**.** However, given that these complex expressions allow syntax to have access to the internal structure, I conclude that their phrasal structure is "real", and has not undergone reification as unanalysable roots. I, therefore, suggest that Assc-compounds should be analysed as phrasal idioms and that the ban on extraction, as well as the word order facts discussed in section 4.4.5 and 4.4.6, has to be explained based on their idiomatic status.

4.5.1 Pronominalisation

The test on anaphoric pronominal replacement is used here to show that a part of an Assc-compound with a KP internal structure, specifically the PossDP, can be anaphorically referenced by a pronoun in a syntactic construction as shown in (63a & b).

63. a. M-misiri wa nyumba sa-na-mang-e
1-expert 1.ASS 10.house 1.SM.NEG-PST-build-SBJV
bwino. ***I***-dza-gw-a.
good 9.FUT-fall-FV
'The builder did not build properly. It will fall'.

b. Mw-ana wa sukulu w-a-thyok-a mw-endo.
1-child 1.ASS .5.school 1-PRF-break-FV 3-leg .
Li-dza-mu-lipir-a.
5-FUT-SM.pay-FV
'The pupil has broken her/his leg. It will pay her'.

In the examples in (63) above, the subject markers *i-* 'it' in *idzagwa* 'it will fall' in (63a) and *li-* 'it' in *lidzamulipira* 'it will pay him' in (63b) either are pronominals themselves or are agreement markers that signal the presence of null pronominals in the subject position of the second clause. Importantly, these subject pronominals refer anaphorically to the associative parts *nyumba* 'house' and *sukulu* 'school' respectively of the Assc-compounds in the preceding sentence. This suggests that the internal structure of Assc-compounds is visible to syntactic rules and can be accessed by syntax.

4.5.2 Modification test

Another indication that the associative constructions with an internal KP phrasal structure are syntactically transparent is that a constituent part of these words can be modified. Lieber and Stekauer (2009) indicate that in compound words, this is typically not possible, as modification cannot affect a constituent part of a compound, but must apply to the entire compound word.

4.5.2.1 Adjectives

In section 4.4.4, I showed that an adjective can modify the entire Assc-compound word. In the example I presented there, the adjective following the Assc-compound could not agree with the PossDP inside the associative phrase. However, in the examples below, I show that a part of Assc-compounds, namely the possessor, can be modified. For clarity, (64a), (65a) and (66a) are used here as base examples from which (64b), (65b) and (66b) are respectively drawn . These examples are syntactically like the examples in (56) above, where this type of modification was not possible. This shows that the impossibility of (56b) is because of semantic reasons, not because of syntactic reasons.

64. a. Pepala la ma-yeso li-dza-lemb-edw-a
5.paper 5.ASS 6.examination 5.SM.FUT-write-PASS-FV
mawa.
tomorrow
'The examination paper will be written tomorrow'.

b. Pepala la ma-yeso **o-limba**
5.paper 5.ASS 6.examination 6.ASS-difficult
li-dza-lemb-edw-a mawa.
5.SM.FUT-write-PASS-FV tomorrow
'The examination paper of difficult examinations will be written tomorrow.'

65. a. M-lezi wa ana w-a-phik-a mowa.
1-keeper 1.ASS 2.child 1.SM-PRF-cook-FV 3.beer
'The nanny has brewed beer'.

b. M-lezi wa ana **a-a-ng'ono** w-a-phik-a

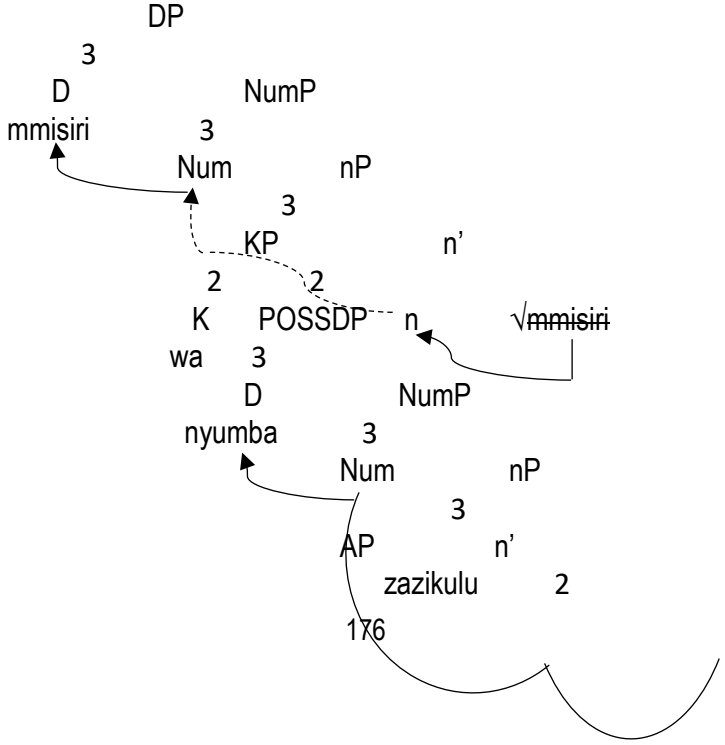
1-keeper 1.ASS 2.child 2.ASS-2-small 1.SM-PRF-cook-FV mowa.
3.beer
'The nanny of small children has brewed beer'.

66. a. M-misiri wa nyumba w-a-thyok-a mw-endo.
1-expert 1.ASS 10.house 1.SM-PRF-break-FV 1-leg
'The builder has fractured his leg'.

b. M-misiri wa nyumba **za-zi-kulu**
1-expert 1.ASS 10.house 10ASS-10-big
w-a-thyok-a mw-endo.
1.SM-PRF-break-FV 1-leg
'The builder of big houses has fractured his leg'.

In (64b) the adjective *olimba* 'difficult' modifies the possessor *mayeso* 'examinations'; in (65b), the adjective *aang'ono* 'small' modifies the possessor *ana* 'children', and in (66b) the adjective *zazikulu* 'big' modifies the noun *nyumba* 'house'. The APs in these examples are located inside the PossDP (in Spec, n) and therefore follow the possessor noun, which has moved to D. In (67), I provide a tree diagram for (66b) to illustrate:

67.



n √~~nyumba~~

It is important to note that when the PossDP is modified as in (64b), (65b) and (66b), the lexical meaning of the compound is preserved. For example, (65b) still maintains the idiomatic meaning that the subject is not just any keeper of children, but a 'nanny', and also adds the information that the children this nanny looks after are small. The regular meaning of the head noun is also activated here, in addition to the idiomatic reading.

The fact that the possessor part of the compound alone can be modified by an adjective shows that the Assc-compounds are not ordinary compounds. For example, consider the English compound *housekeeper*. Adding a modifier as in *the beautiful housekeeper* can only mean that the referent of the whole compound, the housekeeper, is beautiful, but not that it is a keeper of a beautiful house. The Assc-compounds in (64), (65) and (66) differ in this respect, which supports the view that their structure is different.

4.5.2.2 Relative clauses

Relative clause modifiers behave like adjectives in that they can also modify the possessor inside the compound by adjoining to nP (compare (67)). For the sake of clarity (68a) and (69a) are used as base examples from which (68b) and (69b) are derived. The example in (68b) shows that *nyumba* 'houses' can be modified by a relative clause, and (69b) shows that *nyimbo* 'songs' can be modified by a relative clause.

68. a. M-misiri wa nyumba w-a-thaw-a.
1-expert 1ASS 10.house 1.SM-PRF-run-FV
'The builder has run away'.

b. M-misiri wa nyumba zi-mene zi-na-gw-a
1-expert 1.ASS 10.house 10-REL 10-PST-fall-FV
dzulo, w-a-thaw-a.
yesterday 1.SM-PRF-run- FV
'The builder of the houses which fell yesterday, has run away'.

69. a. buku la nyimbo l-a-sow-a.
5.book 5.ASS 10.song 5.SM.BE-lose-FV
'The songbook is lost'.

b. buku la nyimbo zi-mene zi-na-sankh-idw-a
5.book 5.ASS 10.song 10.REL 10-PST-select- PASS-FV
dzulo l-a-sow-a.
yesterday 5.SM.BE-lose-FV
'The songbook of songs which were selected yesterday is lost'

As with the examples with adjectival modifiers, the modification using relative clauses shows that the syntax has access to parts of the associative expression. This is contrary to what the lexical integrity advocates claim to be possible with genuine compounds (see Lapointe, 1980, Di Sciullo and Williams, 1987, Bresnan & Mchombo 1995).

4.5.2.3 Possessive pronouns

Assc-compounds can also take possessive pronouns as modifiers. As indicated in subsection 4.3 above, possessor modifiers are canonically post-nominal modifiers in Cinyanja; they are located in Spec, Num and generally agree in class with the nouns they modify. In example (70a) and (71a), the possessive pronoun agrees with the head noun and hence modifies the entire Assc-compound while in (70b) and (71b) only the possessor inside the associative construction is modified by the possessive pronoun.

70. a. Buku la nyimbo **la-nga** la-a-sow-a.
5.book 5.ASS 10.song 5-my 5.SM-BE.PRS-lose-FV
'My songbook is lost.'

b. Buku la nyimbo **za-nga** la-a-sow-a.
5.book 5.ASS 10.song 10-my 5.SM-BE.PRS-lose-FV
'The songbook with my songs is lost'.

71. a. Mw-ana wa sukulu **w-anga** w-a-dwala
1-child 1.ASS 5.school 1-my 1.SM- BE.PRS-sick
'My pupil is sick'.

b. Mw-ana wa sukulu **l-anga** w-a-dwala

1-child 1.ASS 5.school 5.my 1.SM.BE.PRS-sick
'The pupil of my school is sick'.

From the examples above, we observe that the modifier *zanga* 'my' in (70b) agrees in class with the possessor *nyimbo* 'songs' while *langa* 'my' in (71b) agrees in class with *sukulu* 'school'. In both examples, the lexical meaning of the compound expression is preserved, but the regular meaning of the head noun is also activated and modified by the possessive pronoun. The fact that the possessive pronoun modifier can modify the PossDP again shows that syntax has access to specific parts of Assc-compounds, violating lexical integrity.

4.5.2.4 Demonstratives

In Cinyanja, demonstratives can occur on either side of the noun they modify. In unmarked constructions, demonstratives are post-nominal (adjoined to nP) while in marked constructions they can be pre-nominal (in Spec, D; see section 4.3). The example in (72a) is used as a base example on which (72b-c) are built. In (72b), the post-nominal demonstrative *aka* 'this' modifies the possessor *kanema* 'picture' while in (72c), the demonstrative still modifies *kanema*, but appears pre-nominally (in Spec, PossD). In (73b), the post-nominal demonstrative *iyi* 'this' modifies the possessor *nyumba* 'houses' while in (73c) *iyi* is pre-nominal to *nyumba*. Note that the agreement properties of the demonstrative show that it modifies the possessor in all examples:

72. a. Wailesi ya kanema y-a-pwany-ik-a.
9. radio 9.ASS 12.picture 9.SM-BE-break-PASS-FV.
'The TV is broken.'

b. Wailesi ya kanema k**a-ja** ka-nyama
9.radio 9.ASS 12.picture 12-this 12-10.animal-
za-m-thengo y-a-pwany-ik-a.
10-18.LOC 9.SM-BE-break-PASS-FV
'The TV with that picture/image of wild animals is broken'.

c. Wailesi ya **ka-ja** kanema ka-nyama za-m-thengo
9.radio 9.ASS 12.that 12.picture 12-10.animal 10-18.LOC

y-a-pwany-ik-a.
9.SM-BE-break-PASS-FV
'The radio of/with that picture of wild animals is broken[16].

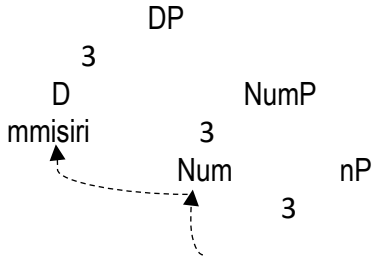
73. a. M-misiri wa nyumba w-a-thyok-a mw-endo.
1-expert 1.ASS 10.house 1.SM-PRF-break-FV 3-leg
'The builder has fractured his leg'.

b. M-misiri wa nyumba **i-zi** w-a-thyok-a
1-expert 1.ASS 10.house 10-these 1.SM-PRF-break-FV
mw-endo.
3-leg
'The builder of these houses has fractured his leg'.

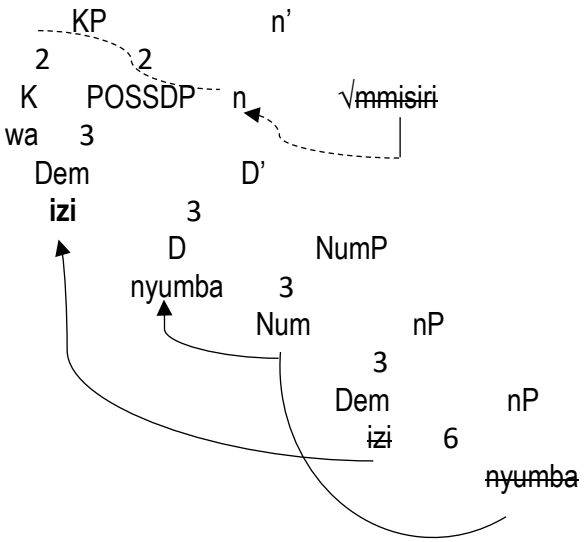
c. M-misiri wa **i-zi** nyumba w-a-thyok-a
1-expert 1.ASS 10-these 10.house 1.SM-PRF-break-
FV mw-endo.
1-leg
'The expert for these houses has fractured his leg'.

In the (b) and (c) examples, the demonstratives agree with the possessor and hence modify the possessor DP inside the Assc-compounds. Importantly, the position of the demonstrative affects the interpretation of the compound. In (72b) and (73b) above, where the demonstrative is post-nominal and occupies a specifier of nP in the PossDP (similar to the AP in the representation in (67) above), the idiomatic meaning of the compound is maintained. In contrast, in (72c) and (73c), the demonstrative is located in the specifier of PossDP and therefore between the associative marker in K and the possessor:

74.



[16] This may imply that there is a TV with a picture/image of wild animals stuck on it or it may also imply that the picture is used as a screensaver for that TV.



After moving to Spec, PossDP, the demonstrative interrupts the linear order of the words that make up the Assc-compound and separates the associative marker in K from the possessor noun in D. As a result, the Assc-compound loses its idiomatic meaning, and the demonstrative puts emphasis on the PossDP. While the latter interpretation is expected from the perspective that Assc-compounds are genuine compounds whose parts cannot be independently modified (and so the compound interpretation gets lost with a DP-internal modifier), the fact that the demonstratives in (72b) and (73b), which have not moved and are located in Spec, n, can modify a part of the compound without a loss of the idiomatic interpretation, suggests that compound formation has not affected the accessibility of the internal phrasal structure with these constructions.

4.5.2.5 Quantifiers

In Cinyanja, quantifiers are a very restrictive category of modifiers and include words such as *-onse*, 'all' and *-ena* 'some'. Quantifiers, like demonstratives, follow the noun they quantify in unmarked constructions as in (75a) and precede the noun they quantify in marked constructions as in (75b).

75	a.	M-misiri	wa	nyumba	**z-onse**	w-a-thaw-a.
		1-expert	1-ASS	10.house	10-all	1.SM-PRF-run-FV
		'The builder of all houses has run away'.				
	b.	M-misiri	wa	**z-onse**	nyumba	w-a-thaw-a.
		1-expert	1-ASS	10-all	10.house	1.SM-PRF-run-FV
		'The expert of all houses has run away'.				

In example (75a) above, the quantifier *zonse* 'all' is post-nominal (in Spec, n of the possessor DP) and agrees in noun class with the possessor *nyumba* 'house', which shows that it modifies the PossDP, not the entire associative expression. In this example, the lexical meaning of the compound word is maintained. However, when the quantifier is prenominal (in Spec, PossDP), and the associative marker and the possessor DP are therefore no longer adjacent in the compound as in (75b), the compound loses its idiomatic meaning. The quantifier expresses emphasis on the possessor, just like the demonstratives in the above examples in (72c) and (73c). The fact that PossDP can be quantified as opposed to the entire complex expression, shows that the internal syntax of Assc-compounds is accessible to modification.

Another interesting characteristic of demonstratives and quantifiers in Assc-compounds is that they can co-occur before or after the possessor, in any order. In the examples below, only the PossDP is modified by the quantifier and the demonstrative, while the head noun of the compound remains the controller of agreement on the verb as exemplified.

76. a. M-misiri wa nyumba **z-onse** **i-zi**
1-expert 1-ASS 10.house 10-all 10-DEM
w-a-thaw-a.
1.SM-PRF-run-FV
'The builder of all these houses has run away'.

b. M-misiri wa nyumba **i-zi** **z-onse**
1-expert 1-ASS 10.house 10-DEM 10-all
w-a-thaw-a.
1.SM-PRF-run-FV
'The builder of all these houses has run away'.

c. M-misiri wa **z-onse** **i-zi** nyumba
1-expert 1-ASS 10-all 10-DEM 10.house
w-a-thaw-a.
1.SM- PRF-run-FV
'The expert of all these houses has run away'.

d. M-misiri wa **i-zi** **z-onse** nyumba

1-expert 1-ASS 10-DEM 10-all 10.house
w-a-thaw-a.
1.SM- PRF-run-FV
'The expert of all these houses has run away'.

e. M-misiri wa **i-zi** nyumba **z-onse.**
1-expert 1-ASS 10-DEM 10.house 10-all
w-a-thawa.
1.SM-PRF-run- FV
'The expert of all these houses has run away'.

f. M-misiri wa **z-onse** nyumba **i-zi.**
1-expert 1-ASS 10-all 10.house 10-DEM
w-a-thawa.
1.SM-PRF-run- FV
'The expert of all these houses has run away'.

In (76a & b) above, both the quantifier and the demonstrative are adjoined to nP, and they co-occur after the possessor noun, which has moved to PossD. In these examples, the idiomatic compound meaning of the complex expression is retained. In contrast, in (76c & d) where the demonstrative and the quantifier simultaneously co-occur before the PossDP[17] and are also used interchangeably, they show emphasis on the possessor, and the lexical meaning of the complex expression is lost. Similarly, in (76e) where the demonstrative is pre-nominal and the quantifier is post-nominal, the associative construction does not retain the lexical meaning of the compound word, and the demonstrative puts emphasis on the PossDP. Further, in (76f) the quantifier is pre-nominal; it also shows emphasis on the PossDP while the demonstrative just modifies the possessor. As in (76e), the lexical meaning of the complex expression in (76f) is lost. Although the demonstratives and the quantifiers may co-occur, used simultaneously or interchangeably, the sentences are grammatical.

To summarise the section on modification, evidence has shown that it is possible to modify the possessor-DP inside the Assc-compounds with adjectives, relative clauses, possessive pronouns,

[17] I assume that the DP in Cinyanja allows multiple specifiers. The word order in (76c) and (76d) is derived when both the demonstrative and the quantifier are located in different Spec, PossD positions.

demonstratives and quantifiers, in violation of lexical integrity. In addition, an interesting generalisation emerges from the data with demonstrative and quantificational modifiers. When these modifiers destroy the linear order inside the Assc-compound to put emphasis on the possessor of the Assc-compound, the lexical compound meaning is lost. I claim that it is the emphasis on the PossDP which is achieved through a pre-nominal modification that causes the loss of lexical meaning in Assc-compounds. What this means then is that the position of the PossDP modifier plays a critical role in analysing associative constructions.

4.5.3 Extraction

Another test that shows whether or not a complex expression is transparent, is extraction. For example, lexical integrity predicts that constituent parts of a compound cannot be relativised. The example in (77b) shows that it is indeed not possible to extract a part of the compound in a relative clause:

77. a. Mw-ana wa sukulu a-mene zi-gawenga
1-child 1.ASS 5.school 1-REL 8-thugs
zi-na-pand-a dzulo w-a-psy-a.
8-PST-beat-FV yesterday 1.SM-BE.PRS-burn-FV
'The pupil whom the thugs beat yesterday is burnt'.

b. ***Sukulu** li-mene zi-gawenga zi-na-pand-a mw-ana
5.school 5.REL 8-thugs 8-PST-beat-FV 1-child
w-ace _____ dzulo l-a-psy-a.
1.PRON yesterday 5.SM-BE.PRS-burn-FV
*'The School which the thugs beat the child of yesterday is burnt'.

In (77b), the possessor DP inside the Assc-compound *mwana wa sukulu* 'pupil', has been relativised, stranding the head noun and the associative marker (which combines with a resumptive pronoun inside the relative clause). This seems to show that lexical integrity cannot be violated through extraction of part of a compound. However, the example in (78b) shows that it is not possible to relativise the possessor of a regular associative construction either:

78**.** a. **Mu-nda wa fodya** u-mene a-nyamata

3-field	3.ASS	5.tobacco	3.REL	2-boy
a-na-sakulir-a		dzulo	w-a-psy-a.	
2-PST-weed-FV		yesterday	3.SM-PRF-burn-FV	

'The field of tobacco which the boys weeded yesterday is burnt'.

b.

***Fodya**	li-mene	a-nyamata	a-na-sakulir-a
5.tobacco	5.REL	2-boy	2-PST-weed-FV
mu-nda	w-ace	dzulo	l-a-psy-a.
3-field	3. PRON	yesterday	5.SM-BE.PRS-burn-FV

*'The tobacco which the boys weeded the field of yesterday, is burnt'.

The examples in (77) and (78) above show that it is generally not possible to extract a PossDP from within an associative construction. The two types of associative constructions behave exactly the same way. The extraction test therefore does not tell us much about the transparency of the internal structure of an Assc-compound.

4.5.4. Gapping

Recall from chapter three that it is impossible to gap part of a morphological word as shown in the examples below:

79. a. John sat on a chair and Mary sat on the floor.
    b. John sat on a chair and Mary _____on the floor.

80. a. John is a truck-driver and Mary is a taxi-driver.
    b. *John is a truck-driver and Mary is a taxi _____.

The example in (79b) above shows that it is possible to elide a verb in the second conjunct because it is a syntactically independent element. In contrast, example (80b) shows that it is impossible to elide *driver* from the compound word *taxi-driver* because *taxi-driver* is a morphological word.

Regular complex associative constructions in Cinyanja allow gapping of identical head nouns:

81. a. Mu-nda wa cimanga ndi mu-nda wa
3.field 3.ASS 7.maize and 3.filed 3.ASS
fodya ya-a-psy-a.
5.tobacco 4.SM-BE.PRS-burn-FV.
'The field of maize and the field of tobacco are burnt'.

b Mu-nda wa cimanga ndi (~~mu-nda~~) wa
3.field 3.ASS 7.maize and ~~3.field~~ 3.ASS
fodya ya-a-psy-a.
5.tobacco 4.SM-BE.PRS-burn-FV.
'The field of maize and of tobacco are burnt'.

c Mu-nda wa cimanga ndi (~~mu-nda~~ ~~wa~~)
3.field 3.ASS 7.maize and ~~3.filed~~ ~~3.ASS~~
fodya ya-a-ps-a.
5.tobacco 4.SM-BE.PRS-burn-FV.
'The field of maize and tobacco are burnt'.

The grammaticality and interpretation of the coordinated associatives in (81a) are not affected when the head noun of the conjoined phrase is gapped in the second conjunct, as shown in (81b). In (81c) where the head noun and the associative marker are gapped in the conjoined phrase, the meaning of the phrase is not lost because it is inferred from the context, and the sentence is grammatical.
Similar to regular associative constructions, gapping of head nouns in Assc-compounds is possible in the second conjunct of coordinated associatives when two lexicalised Assc-compounds are coordinated, provided they are based on identical head nouns. Such an example is provided in (82):

82. a. M-phunzitsi a-na-on-a mw-ana wa ma-siye
1-teacher 1.SM-PST-see-FV 1-child 1.ASS 6.deceased
ndi mw-ana wa m'-thengo.
and 1-child 1.ASS 18.LOC.5.bush
'The teacher saw the orphan and the bastard'.

b. M-phunzitsi a-na-on-a mw-ana wa ma-siye
1-teacher 1.SM-PST-see-FV 1-child 1.ASS 6.deceased

ndi (~~mw-ana~~) wa m'-thengo.
and 1-child 1.ASS 18.LOC-5.bush
'The teacher saw the orphan and the bastard'.

In (82a), both Assc-compounds are based on the head noun *mwana* 'child'. Example (82b) above shows that when two identical head nouns appear in two coordinated Assc-compounds, gapping of identical material is possible, and the idiomatic meaning of the conjoined compound is not lost. This is further evidence that parts of Assc-compounds are accessible to rules such as gapping, and that Assc-compounds do not behave like compound words and do not respect lexical integrity.

Given the acceptability of (81c) above, we expect that Assc-compounds should also allow gapping of the head noun and the associative in the second conjunct. However, this is not possible, in contrast to regular associative constructions:

83. *M-phunzitsi a-na-on-a mw-ana wa ma-siye
1-teacher 1.SM-PST-see-FV 1-child 1.ASS 6-deceased
ndi (~~mw-ana~~ ~~wa~~) m'-thengo.
and 1-child 1.ASS 18.LOC-5.bush
Intended: 'The teacher saw the orphan and bastard'.

In (83) where both the head noun and the associative are gapped, the prediction is that this sentence is grammatical, but with a non-idiomatic meaning: "the teacher saw the child of the deceased and the bush". As regards the idiomatic meaning of the conjoined compound, it is completely lost and the sentence is ungrammatical. This shows that it is not possible to separate the associative from the PossDP through gapping.

The Lexical Integrity tests that I applied to Assc-compounds in this section have yielded mixed results. Most importantly, however, I have shown that elements internal to Assc-constructions are visible to syntax. Assc-compounds are not anaphoric islands; their parts can function as antecedents for pronouns. Furthermore, the possessor inside an Assc-compound can be modified by adjectives, relative clauses, possessive pronouns, demonstratives and quantifiers, provided the modifier does not interrupt the adjacency of an associative marker and Poss-DP. While individual elements cannot be extracted from Assc-compounds, I showed that this is also the case with regular ANPs. Finally, with

regards to gapping, the results show that the head noun of an Assc-compound can be gapped, but the combination of head noun and associative marker cannot.

The conclusion I draw from these results is the following. The fact that modifiers have access to constituent parts of the associative compounds shows that these complex expressions have not been formed via reification of their phrasal parts as unanalysable roots. At the same time, the fact that the extent to which these phrasal parts can be accessed by syntax is limited, suggests that certain licensing conditions govern the formation of these compounds, which explains why the phrasal parts of the compounds cannot be extracted or gapped. In the next section, I argue that the Assc-compounds analysed in this study are syntactically transparent phrasal idioms with lexicalised meaning or idiosyncratic meaning, which, however, are licensed only when the possessor-DP remains adjacent to the associative marker. When this adjacency is destroyed, either by an intervening modifier or via gapping of the associative marker together with the head noun, the idiomatic meaning is no longer available.

4.6 Assc-compounds as phrasal idioms

In the previous sections, I have discussed Assc-compounds and have shown that they allow syntax to have access to individual parts of the associative construction. In this section, I provide evidence based mainly on Jackendoff (1997) that the Assc-compounds discussed in the previous sections are phrasal idioms. This is achieved by comparing the distributional properties of phrasal idioms (Nunberg et al; 1994; Jackendoff, 1997) with those of the Assc-compounds. I show that the Assc-compounds have the same properties as phrasal idioms; they are more complex than X^0, they have fixed semantics, fixed syntactic structure and may allow modification of individual parts. Based on these similarities, I argue that the Assc-compounds are not compound words per se but that they are phrasal idioms.

4.6.1 Phrasal idioms

I begin by discussing some phrasal idioms from English. Phrasal idioms have a syntactic structure of an ordinary phrase. Specifically, the examples below comprise a verb and a noun phrase/DP (and hence form a VP):

84.	kick the bucket	'to die'
	shoot the breeze	'chat'
	spill the beans	'divulge information'

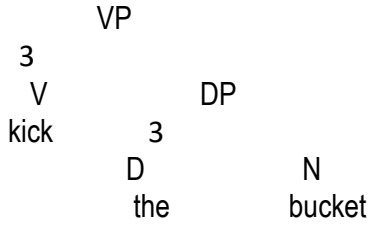
let the cat out of the bag	'reveal a secret'
bury the hatchet	'reconcile an agreement'
break the ice	'break down a barrier to social interaction'

(Jackendoff, 1997:154 &168)

The examples in (84) above are called idioms and are generally considered non-compositional, that is, the meaning of the entire expression cannot be deduced from its constituent parts (Jackendoff, 1997). Culicover and Jackendoff (2005) define an idiom as a fixed syntactic construction which is made up of words that are already in existence in the lexicon. Idioms generally have a syntactic structure that resembles regular syntactic phrases, but have a typically idiosyncratic meaning, that is, meaning which is unpredictable from the sub-parts. For instance, an idiomatic expression such as *kick the bucket* does not refer to the actual act of "kicking" or the physical object, "the bucket". Rather, the expression *kick the bucket* has the syntactic structure in (85b) but means 'to die'.

85. a. kick the bucket

b.



Nunberg et al. (1994) observe that the words that form these idioms are lexically fixed, and cannot be substituted with other words with similar meanings. This is shown in the examples below.

86. a. John's father kicked the bucket yesterday.
'John's father died yesterday'

b. John's father kicked the **pail**
'John's father kicked the pail (i.e. the physical object)

c. John's father **hit** the bucket
'John's father hit the bucket (i.e. the physical object)

87. a. Mary spilled the beans
'Mary revealed a secret or confidential information to someone'

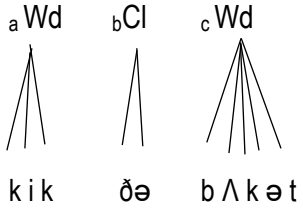
b. Mary spilled the **vegetables**
'Mary caused the vegetables to flow over'

c. Mary **dropped** the beans.
'Mary caused the beans to fall'.

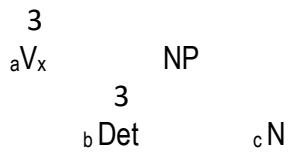
The examples in (86b-c) and (87b-c) show that when a part of the idiom is replaced by another word with similar meaning, the idiomatic meaning is lost. This shows that the special meaning of a phrasal idiom relies on the choice of the right words, both in terms of their syntax and their phonology.

The theory of Lexical Licensing provided in Jackendoff (1997) allows for a representation of phrasal idioms that captures these properties of idioms. According to Jackendoff's (1997) analysis of idioms, the lexical representation of an idiomatic expression such as *kick the bucket* looks as in (88) below. According to this representation, the idiom is lexically specified as a chunk of syntactic structure (its lexical syntactic structure (LSS)) which is a VP. The meaning associated with this VP that is stored in the mental lexicon is represented as a lexical conceptual structure (LCS) "to die" (for a detailed discussion of LCSs, see Jackendoff 1983; 1990). This idiosyncratic LCS is linked to the syntactic representation of the idiom via a lexical linking index ("x" in (88b)). Furthermore, each terminal node in the LSS in (88b) is linked (via the linking indices a, b and c) to the lexical phonological structure (LPS) of the regular words *kick*, *the* and *bucket* in (88a). Note that in Jackendoff's (1997) system, noun phrases are represented as NPs, not as DPs:

88. a. LPS



b. LSS



c. LCS [DIE([ ]$_A$)]$_x$

(Jackendoff, 1997:169)

The representation in (88) captures the fact that the idiom *kick the bucket* is syntactically a VP which is linked to a special meaning and must consist of exactly these three words, which cannot be replaced by synonyms. The syntactic terminals carry lexical indices (subscripts) b on Det, c on N, and a on V in LSS, which are linked to phonological elements in the LPS but do not correspond to any of the features in the LCS. This is an indication that they do not have an isolated meaning interpretation from the rest of the parts of the idiom. According to Jackendoff (1997), the interpretation of "die" is marked by the subscript x. The subscript *x* on the whole LCS indicates that this maps to the verb because the idiom has a verbal meaning. The empty subscript A in (88c) maps to the conceptual structure of the external argument of the sentence, but it is not possible to link the DP *the bucket* (an NP in Jackendoff's representation) to an argument in the LCS.

Jackendoff also observes that idioms such as *kick the bucket* are syntactically fixed and that their constituent parts may not undergo movement like other ordinary syntactic phrases because this may affect their idiomatic interpretation as shown in (89b).

89. a. John kicked the bucket.
    b. *The bucket was kicked by John.

In the idiomatic expression *kick the bucket* in (89a) above, there is no separation of constituents and the idiomatic meaning is maintained. However, in (89b), where the idiomatic expression has been passivised, and the object of the verb has been promoted to the subject position, the idiomatic meaning is lost. (89b) is only possible if the VP *kick the bucket* is interpreted literally; it cannot be interpreted as having an idiomatic meaning.

Furthermore, in opaque idioms such as *kick the bucket*, it is also impossible to modify constituent parts of the idiom. Jackendoff (1997) notes that modification of the object-NP in the idiom *kick the bucket* by an adjective will cause the idiomatic meaning to be lost. Therefore, modification can only apply to the entire idiomatic expression as shown in (90b).

90. a. John kicked the bucket.
    b. John certainly kicked the bucket. (idiomatic meaning retained)
    c. *John kicked the big bucket. (ungrammatical with the idiomatic reading)

The example in (90c) above does not have the same interpretation as (90a & b). The idiomatic reading in this expression is lost. This shows that the modification of constituent parts in opaque idioms is impossible. (90b) modifies the sense of the entire idiom, therefore, it is acceptable. (90c) is wrong because the noun *bucket* cannot be modified independently. According to Jackendoff, the DP *the bucket* has no theta role in the idiom, and the elements of the DP are connected to the verb *kick* syntactically. As a result, the elements cannot be modified or separated by the movement of the DP. If the DP is allowed to move, then it has a theta role and is not attached to the verb any longer. The result is the loss of idiomatic meaning.

However, Jackendoff also notes that some idioms are syntactically more transparent than idioms of the *kick the bucket*-variety. Consider:

91. a. draw a line
    'make/enforce a distinction'
    b. bury a hatchet
    'reconcile a disagreement'

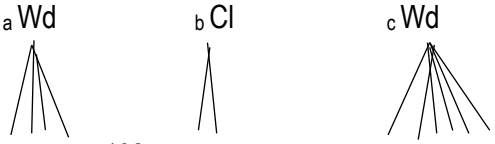
Importantly, with these transparent idioms, their constituent parts can be moved and modified, as shown in the following examples.

92. a. They buried a hatchet.
    b. They buried a proverbial hatchet.
    c. The hatchet was buried.

(Jackendoff, 1997:169)

The examples above show that it is possible to modify the DP within the idiomatic expression and move the constituent parts when the idiom is transparent and still retain the idiomatic interpretation. Jackendoff argues that parts of a transparent idiom can be modified and moved because the constituent parts of the idiom can map to the subcomponents of the idiomatic interpretation at LCS. For example, *bury* means 'reconcile' and *the hatchet* means 'disagreement'. The lexical entry for the idiom *bury the hatchet* can be represented as in (93) below.

93 a. LPS



b e r i ð ə h æ t ʃ ə t

b. LSS $_{a}V_{x}$ NP_{y}

3

$_{b}$Det $_{c}$N

c. LCS [RECONCILE ([ ]$_{A}$ [DISAGREEMENT] $_{y}$)]$_{x}$

(Jackendoff, 1997:169)

In (93) above, the V and NP are not syntactically connected. According to Jackendoff (1997:169), the NP is linked to V via its theta role. The subscript *x* in the LCS indicates that the meaning 'reconcile' is linked to the verb, and the argument slot indexed by the subscript A is to be filled by the external argument of the sentence. The LCS shows that the other semantic argument of the verbal meaning is mapped into a specified constituent in the syntax (the NP comprising the Det + N, i.e. *the hatchet*) by the subscript *y*.

The idiom in (93) is different from the one in (88) because the DP *the hatchet* is linked to *bury* only through the theta role (the index y) while in (88) the DP *the bucket* is linked to *kick* syntactically. Additionally, the parts of the idiom in (93) add to the independent meaning which contributes to the overall interpretation of the idiom. The LSS is linked to two separate parts in the LCS, that is, *bury* matches with 'reconcile' and *the hatchet* matches with 'disagreement'. These independent meanings are expressed by the subscripts x and y. Jackendoff posits that the syntactic separation between parts allows for the syntactic mobility and modification of the NP in transparent idioms. Even if the NP is realised not as the object of the verb, but in a different syntactic position as a result of movement, the idiomatic interpretation can be activated via the lexical entry in (93). This is in contrast to (88), where the entire syntactic VP, not individual constituent parts, is linked to the idiomatic interpretation. Since the VP with the verb *kick* and the NP *the bucket* is stored as one unit in the lexical entry which licenses the idiomatic reading, this interpretation is no longer available when the VP-structure specified in (88) is destroyed as a result of movement.

In sum, it has been shown that movement and modification of constituent parts are not possible with opaque idioms such as *kick the bucket*, while modification and movement of constituent parts are possible with transparent idioms. In the next section, I show that the Assc-compounds discussed in this chapter are phrasal idioms because they have a complex syntactic structure which is linked to a

fixed semantics. While parts of the fixed syntax of the idiom can be modified, I argue that this is only possible for as long as the linear order of the words that make up the idiom is preserved.

4.6.2 Assc-compounds as phrasal idioms.

In section 4.4, I showed that Assc-compounds have the same syntactic distribution as DPs and that they appear in the same syntactic context as other DPs. In section 4.5 I showed that syntax has access to parts of the compounds because these parts can be anaphorically referred to by pronouns, modified, and gapped. However, it was shown that modification cannot apply if the modifier destroys the fixed linear order of the Assc-compound. Concerning the extraction of constituent parts of the Assc-compounds, it has been shown that parts of the compound cannot be extracted. For gapping, it has been shown that gapping of the associative marker together with the head noun results in loss of idiomatic meaning for the associative constructions. In this section, I show that these properties of Assc-compounds are consistent with an analysis of these constructions as phrasal idioms because Assc-compounds, like phrasal idioms, have a complex transparent structure with a lexicalised meaning which is licensed only when a fixed order of the words that make up the idiom is preserved, which is specified in the LSS of the idiom's lexical entry.

I begin with Type 1 Assc-compounds which are syntactically transparent. In examples such as (94), the meaning of the whole compound can be predicted from the compound constituents.

94. a. buku la nyimbo
'songbook'

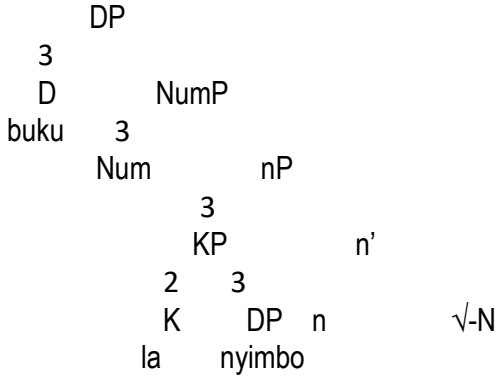
b. buku la nyimbo
'book of songs'

95. a. buku la mapemhero
'prayer book'

b. buku la mapemhero
'book of prayers'

The examples in (94 & 95) above show that some Assc-compounds are transparent and indeed closely resemble their phrasal counterparts, i.e. regular productive associative constructions. I argue that Assc-compounds such as these are represented by a regular syntactic structure associated with associative constructions as shown in (96b).

96. a. buku la nyimbo
'songbook'

b.


It is possible that the meaning of these compounds is composed in an entirely regular fashion, but that it has become conventionalised by frequent use, and that the compound with the structure in (96) is stored in the mental lexicon as one transparent phrasal idiom.

However, for Assc-compounds that are opaque, there is no one-to-one mapping from syntax to semantics. Recall that one of the defining properties of phrasal idioms mentioned above is that they are more complex than X^0 categories, but are lexicalised and have idiosyncratic meanings like words. These properties also apply to Type 2 Assc-compounds such as the ones below, because, like phrasal idioms, they comprise more than one word and their meaning may be non-compositional, that is, the meaning of the expression as a whole cannot be predicted from the constituent parts as shown in the examples below:

97. a. madzi a moyo
6.water 6.ASS 3.life
'oral-rehydration'
b 'water of life'

98. a. mwana wa leza

1.child of 1.god
'rainbow'

b. 'child of god'

99. a. wailesi ya kanema
9.radio 9.ASS 12.picture
'TV'

b. mw-ana wa m'-thengo
1-child 1.ASS 18.LOC- bush
'bastard'

c. mw-ana wa leza
1-child 1.ASS 1.god
'rainbow'

From the examples in (99) above, it is clear that these Assc-compounds are more complex than single words, and that they are phrasal. At the same time, their meaning cannot be deduced from the individual compound elements. For example, in (99c), there is no relationship between the constituent parts that form *mwana wa leza* 'child of god' and the compound meaning 'rainbow'. This shows that indeed these Assc-compounds have idiomatic properties. These properties are also illustrated by the fact that the words that make up the Assc-compound cannot be substituted with other words with similar meanings without losing the idiomatic meaning. This is illustrated in the examples below.

100. a. mwana wa leza
child of god
'rainbow'

b. mwana wa cauta
child of god
'child of god'

101. a. wailesi ya kanema
radio of picture

'TV'

b. wailesi ya chithunzi
radio of picture
'radio of picture'

The examples in (100) show that when a part of the Assc-compound is replaced with another word with a similar meaning, the idiomatic interpretation is lost. In example (100a) above, the idiomatic meaning for *mwana wa leza* is 'rainbow'. However, when the possessor DP *leza* 'god' is replaced with another DP *cauta* 'god' in (100b), it can no longer mean 'rainbow'. The replaced version of the compound becomes an ordinary associative construction which means 'child of god'. Similarly, in (101a) *wailesi ya kanema* is an Assc-compound which means 'TV'. However, when the DP *kanema* 'picture' in (101b) is replaced with the synonym *chithunzi* 'picture', the idiomatic meaning is lost. In both examples, when a part of the Assc-compound is replaced by another DP, the substitute version of the complex expression becomes an ordinary syntactic phrase. As discussed in section 4.6.1 above, phrasal idioms behave in the same way.

These properties suggest that we are dealing with idiomatic compounds with a phrasal structure. However, I have argued above that, in contrast to the phrasal compounds I discussed in Chapter 3, Assc-compounds do not include reified phrasal structures. The main argument comes from the fact, discussed in subsection 4.5.2, that individual parts of the Assc-compounds are accessible to syntactic rules and operations, such as pronominalisation, modification, and gapping. In contrast to VN-compounds, Assc-compounds allow violations of lexical integrity, which is clear evidence that the phrasal material they are built from does not behave like a reified lexical root, but like a genuine syntactic phrase.

Let me review the relevant observations from section 4.5. In Assc–compounds, pronominalisation of the PossDP which can be anaphorically referenced by a pronoun is possible. As was shown in example (64) in subsection 4.5.1, repeated in (102) below, it is possible to pronominalise the PossDP within the compound and still retain the idiomatic interpretation:

102. M-misiri wa nyumba sa-na-mang-e bwino.
1-expert 1.ASS 10.house 1.SM.NEG-PST-build-SBJV good
I-dza-gw-a.

9.FUT-fall-FV

'The builder did not build properly. It will fall'.

Furthermore, it was shown that the possessor inside an Assc-compound can be modified. For example, a demonstrative that follows the possessor it modifies is licensed without loss of the idiomatic reading. Examples (103) and (104) repeat examples (73b) and (74b) from section 4.5.2 above:

103.	a.	Wailesi	ya	kanema	k**a-ja**	ka-nyama
		9.radio	9.ASS	12.picture	12-this	12-10.animal-
		za-m-thengo	y-a-pwany-ik-a.			
		10-18.LOC	9.SM-BE-break-PASS-FV			
		'The TV with that picture/image of wild animals is broken'.				

104.	b.	M-misiri	wa	nyumba	**i-zi**	w-a-thyok-a
		1-expert	1.ASS	10.house	10-these	1.SM-PRF-break-FV
		mw-endo.				
		1-leg				
		'The builder of these houses has fractured his leg'.				

Finally, it was also shown that the head noun of the Assc-compound can be gapped. (105) repeats example (82b) from above:

105.	M-phunzitsi	a-na-on-a		mw-ana	wa	ma-siye	ndi
	1-teacher	1.SM-PST-see-FV		1-child	1.ASS	6.deceased	and
	(~~mw-ana~~)	wa	m'-thengo.				
	1-child	1.ASS	18.LOC-5.bush				
	'The teacher saw the orphan and the bastard'						

With respect to the possibility of modifying or gapping a part of the Assc-compound, Assc-compounds are similar to regular associative constructions. However, recall that there is an important difference. In contrast to regular associative constructions, which allow certain modifiers of the possessor to appear between the associative marker and the possessor, this is not possible with Assc-compounds. For example, when the demonstrative in (106) appears between the associative marker and the possessor, the idiomatic reading of the compound as meaning 'TV' is no longer available (compare example (73c) in section 4.5.2):

106. Wailesi ya **a-ka** kanema y-a-sow-a.
9.radio 9.ASS 12-this 12.picture 9.SM-BE-lose-FV.
'This radio of this picture is lost'.

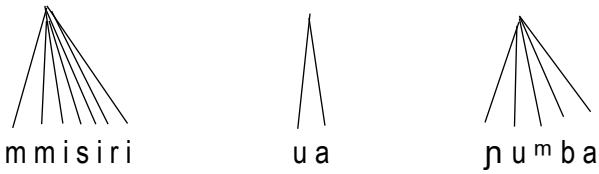
Similarly, while regular associative constructions allow gapping of the head noun together with the associative marker, this is not possible with Assc-compounds:

107. *M-phunzitsi a-na-on-a mw-ana wa ma-siye
1-teacher 1.SM-PST-see-FV 1-child 1.ASS 6-deceased
ndi (~~mw-ana~~ ~~wa~~) m'-thengo.
and 1-child 1.ASS 18.LOC-5.bush
Intended: 'The teacher saw the orphan and the bastard.'

These examples show that Assc-compounds are to some extent transparent phrasal idioms, as parts of their internal syntax are visible to processes like pronominlisation, modification and gapping. At the same time, the combination of associative marker and possessor-DP cannot be interrupted by these processes, suggesting that this aspect of the structure of Assc-compounds is lexically fixed.

Given these observations, I will now present an analysis of Type 2 Assc-compounds which, following Jackendoff's (1997) theory of lexical licensing introduced above, captures these properties. I offer the following lexical entry for an Assc-compound such as *mmisiri wa nyumba* 'builder':

108. a. LPS $_{a}$Wd $_{b}$K $_{c}$Wd



b. LSS D_x NumP

$_{a}N_i$ Num nP

KP n'

$_{b}$K DP n NP

D NumP N_i

|
$_{c}$ N

c LCS [$_{Person}$ TYPE: BUILDER]$_{X}$

The lexical entry in (108) specifies the LPS, the LSS and the LCS of the Assc-compound *mmisiri wa nyumba* 'builder'. The LSS-part of the lexical entry includes phrasal material. This classifies the Assc-compound as a phrasal idiom.

The LSS in (108b) includes two unconnected parts. The phrasal part is the NumP of the associative construction, which includes the KP headed by the associative marker and thereby marks this aspect as a fixed part of the syntax of Assc-compounds. The lexical representation of the Assc-compound also includes a regular "chunk" of the syntax that is accessible for pronominalisation, namely the possessor-DP (the complement of K). The D-head of the possessor-DP includes the possessor noun (recall that DPs in Cinyanja exhibit N-to-D movement). The terminal node K and the possessor N-head in the head position of the possessor-DP are linked to the relevant LPS entries in (108a) via the subscripts b and c. However, the subscripts b and c in the LSS do not correspond to any element in the LCS in (108c) (but see section 4.6.3 below).

Importantly, the combination of K and its complement PossDP is a fixed aspect of the LSS of the Assc-compound. As a result, the lexical entry in (108) specifies that the associative marker and possessor noun must be adjacent. This explains why no pre-nominal demonstrative or quantifier can appear with these idioms. Since these pre-nominal modifiers are located in the specifier of the possessor-DP (see section 4.5.2), a syntactic structure with these modifiers cannot be unified with the lexical entry in (108), in which the order of associative and possessor-DP is not interrupted by additional material. As a result, a syntactic structure in which the possessor-DP includes a specifier cannot license the LCS in (108c), and the idiomatic reading is therefore not available. In contrast, the LSS in (108b) does not specify the precise structure of the syntax below the possessor NumP. This means that post-nominal modifiers, which are realised as specifiers or adjuncts inside NumP and hence follow the possessor noun in D, do not interrupt the adjacency of the associative and the possessor, and are compatible with the conditions on lexical licensing specified by the LSS in (108b). With these modifiers, lexical licensing of the idiomatic interpretation is still possible.

The second part of the LSS in (108b) represents the D-head of the Assc-compound. This D-head is marked as including the head noun *mmisiri*, which is linked to the N-head in the fixed NumP-structure

via the movement index i. The LPS of the head noun in (108a) is linked to the N-head in the D-position via the subscript a. Since all three LPS-entries in (108a) are linked to parts of the LSS, the idiom consists of fixed lexical material, which explains why none of the words that make up the idiom can be replaced by synonyms with a different LPS.

The link between the idiomatic interpretation, specified in the LCS in (108c), and the LSS in (108b) is marked by the subscript x which connects the meaning represented in the LCS to the D-head of the Assc-compound, thereby specifying the meaning of the whole Assc-compound DP as 'builder'. This meaning is preserved even when post-nominal modifiers are added in the syntax.

I suggest that, because the D-head of the Assc-compound in the LSS in (108b) and its associated LPS in (108a) are not structurally connected to the fixed NumP-part of the compound and its associated LPS-entries, it can be accessed independently by a process such as gapping. However, gapping will only be able to access the D-head (including the head noun) by itself since the D-head is not connected to the associative marker, which itself is part of a fixed structure in the LSS of (108); gapping the head noun of the compound together with the associative marker will not be possible. The reason for this is that gapping, which corresponds to the deletion of a part of the idiom's LPS, must be licensed by the presence of another terminal node in the syntax whose LPS is identical to the elided content of the idiom. We can then rule out examples such as (107) above by assuming that this type of licensing cannot apply to two separate parts of the LSS of a lexical entry.

Finally, the LSS in (108b) also explains the observation made in section 4.4 that modifiers of the whole Assc-compound can only appear in a compound-final position, following the associative-KP, but not between the head noun and the KP. Recall that this is a difference between Assc-compounds and regular associative constructions, which allow modifiers to intervene between the head noun and the associative phrase. The fixed LSS in (108b) explains why this latter possibility is not licensed with Assc-compounds. Modifiers that intervene between the head noun (in D) and the associative-KP (in Spec, n) in Cinyanja are located in Spec, Num, or adjoined to nP (see section 4.3.2). However, a syntax with modifiers in these positions could no longer be unified with (108b), which specifies the whole NumP as a fixed structure, without a specifier, and without adjuncts inside nP. As a result, an idiomatic reading can only be licensed when modifiers are attached to a position which is not part of the LSS in (108b). Crucially, I suggested in section 4.3.2 above that modifiers that follow the associative phrase in Cinyanja are right-adjoined to DP. Importantly, since the D-head and the NumP of the Assc-compound are not connected in the LSS in (108b), the DP-node of the syntax of the whole

Assc-compound is not lexically fixed. Therefore, a DP which includes modifiers which are adjoined to DP can be unified with the lexical entry in (108) and license an idiomatic reading. This also applies to DPs with prenominal modifiers in Spec, D:

109.	**u-yu**	m-misiri	wa	nyumba
	1-this	1-expert	1.ASS	10.house
	'this builder'			

Since prenominal demonstratives are in Spec, D and DP is not part of (108b), example (109) is correctly predicted to be possible with an idiomatic reading.

In sum, the lexical entry in (108), which I assume represents the key properties of all Type-2 Assc-compounds, captures the empirical properties of these phrasal idioms that were discussed in this chapter. The key insight is that the LSS of these idioms combines aspects of Jackendoff's representations for opaque and for transparent idioms, as it includes both fixed and unconnected parts of the syntax of associative constructions. This explains why there are lexically defined limits to the extent to which the syntax of these phrasal compounds is transparent to processes such as modification and gapping.

4.6.3 Lexical licensing and modification

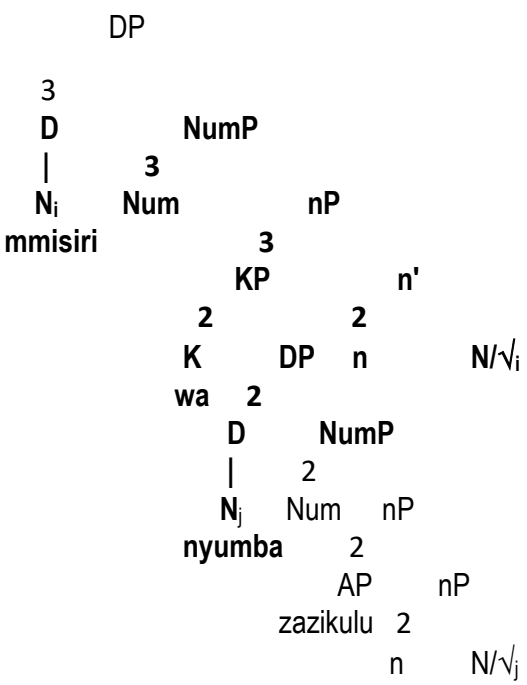
Before ending this chapter, I want to return to an interesting observation with regards to the modified idioms that was made in section 4.5. I demonstrated that with some idioms when a constituent part of the idiom is modified, the idiom preserves the idiomatic reading of the whole idiom while at the same time it seems to activate the *regular* meaning of the individual parts (i.e. 'expert of big houses' = 'builder of big houses').

110.	M-misiri	wa	nyumba	**za-zi-kulu**
	1-expert	1.ASS	10.house	10ASS-9-big
	'The expert of big houses'			
	'The builder of big houses'			

The associative DP in (110) receives its idiomatic interpretation as 'builder', but at the same time incorporates the literal meaning of the modified possessor noun. I now show that this interpretation follows from the analysis of these idioms I have provided in (108) above.

Modification of the possessor in (110) above is implemented by means of a postnominal adjective, which is located in the specifier of the (possessor) nP. The syntactic structure corresponding to (110) is given in (111):

111.



As can be easily verified, the structure in (111) can be unified with the LSS in (108). The relevant parts of the structure which match the LSS in (108b) are marked in bold in (111). The structure, therefore, licenses the meaning 'builder' by linking the LCS in (108c) to the D-head *mmisiri* in (111), as specified by the linking index x in (108). However, the syntax in (111) also includes an adjective, which modifies the noun *nyumba*, 'houses'. Note that this terminal node is not linked to any part of the LCS in (108); it does not feature in the LCS of the idiom. However, it seems that the presence of the adjective allows for the terminal node with the LPS of *nyumba* to access the regular meaning of this noun and, in fact, the whole KP-phrase. In other words, the syntax in (111) seems to simultaneously be unified with more than one lexical entry. While the boldfaced part of the structure in (111) licenses the idiom and assigns the idiomatic meaning to the head of the construction, the regular meaning of *nyumba* is also activated. The regular lexical entry of the noun *nyumba* provides the interpretation which is then modified by the adjective. As a result, the syntax in (111) is interpreted as 'the builder of big houses'.

The idea that the parts of an Assc-compound can simultaneously activate the meaning of the idiom and the literal meaning of its parts also potentially solves a problem that arises for examples such as (102) above. Since the 'house'-part of the idiom is not linked to the LCS in the idiom's lexical entry, the question that arises is how the pronoun in the sentence following the Assc-compound can refer to a

house when it follows the compound meaning ‘builder’ (whose LCS does not include a reference to a house). This suggests that the following pronoun, like an adjective, can activate the literal meaning of the part of the idiom simultaneously to the idiomatic LCS.

As far as I am aware, the ability of idioms to simultaneously license idiosyncratic as well as a regular meaning of their parts has not been observed before. This possibility is a natural consequence of the representation of phrasal idioms in terms of Jackendoff's (1997) theory of Lexical Licensing that I have adopted for my analysis of Assc-compounds in Cinyanja.

4.7 Conclusion

In this chapter, I have analysed two types of Assc-compounds. One type is almost entirely compositional and perhaps only special in that it is stored as one entry in the lexicon, and; the other type is not entirely compositional; it has more idiosyncratic meanings. The AM which links the two nouns conveys different semantic realations between the head noun and the possessor. I have claimed that Assc-compounds, like ANPs, have an internal structure that includes a KP. The empirical analysis of Cinyanja Assc-compounds has revealed that the syntax has access to its constituent parts. Particularly, the possessor-DP can be pronominalised, and the possessor noun can be modified by adjectives, possessive pronouns, demonstratives, quantifiers, and relative clauses. The lexical representation of Assc-compounds combines aspects of lexical entries that have been proposed for opaque idioms which have a fixed phrasal structure in LSS, and for transparent idioms where not all pieces of the LSS are connected. As a result, we get idioms with fixed parts which nevertheless allow modification, but this modification is not unrestricted but constrained by the way the fixed parts of the idiom are lexicalised. Some Assc-compounds may allow activation of not only the idiomatic meaning but also, and simultaneously, activate the regular meaning of its parts. Given these observations, I claim that the complex expressions discussed in this chapter are phrasal idioms.

Therefore, I claim in this chapter that the associative constructions with compound characterisation are not compound words but phrasal idioms. This is so because none of the words that make up the Assc-compounds can be replaced by synonyms of a different LPS. The Assc-compounds also have a fixed LSS which has phrasal material. In the Assc-compounds, like Jackendoffs representation of idioms, all three entries of the LPS are linked to parts of the LSS but not to the LCS. The idiosyncratic LCS is linked to the LSS through a lexical linking index. Given these observations, I claim that the complex expressions discussed in this chapter are phrasal idioms.

CHAPTER 5: CONCLUSION

5. Introduction

The previous chapters analysed two types of phrasal compounds in Cinyanja, namely; VN-compounds and Assc-compounds, and provided insights into the structure of these complex expressions as it relates to the Lexical Integrity Hypothesis.

This chapter is a summary of the investigation into these two types of phrasal compound. The chapter highlights some of the core issues that are discussed about the two types of phrasal compounds in Cinyanja. The organisation of the chapter is presented in such a way that a summary of the research findings is provided to determine the extent to which the set aim and objectives have been achieved. This is followed by what the current research on compounding has contributed to the fields of morphology and syntax. The recommendations for future research are provided and finally, the conclusion is drawn.

5.1 Summary of the findings

As indicated in Chapter 1, the aim of the study emanated from the general question, what is the status of the phrasal lexical items for the theory of compounds as regards the lexical–syntax-morphology interface? This question has been answered by analysing the structure of two types of phrasal compounds in Cinyanja; the VN-compounds and Assc-compounds. In analysing the VN-compounds three types were identified, but only two types were analysed because these included phrasal material. In one type, the VN-compounds consisted of a verb and an object complement with a NC prefix, and in the other type, the VN-compounds comprised a verb followed by a nominal part with a locative prefix. The analysis of the internal structure revealed that VN-compounds have a VP phrasal structure inside the compounds. It was shown that syntactically, these VN-compounds are DPs, because, just like other DPs, they can function as subject and object in a simple clause, can be modified, and can also be coordinated with other DPs. However, the application of the Lexical Integrity tests revealed that the phrasal part cannot be accessed by syntactic operations and behaves like a syntactic atom. To account for this, I adopted a proposal by Harley (2009) and argued that the phrasal structure can be reified, i.e. can re-enter the derivation as a root after phasal spell-out. I concluded that the VN-compounds are genuine phrasal compounds, thus, they are syntactic words.

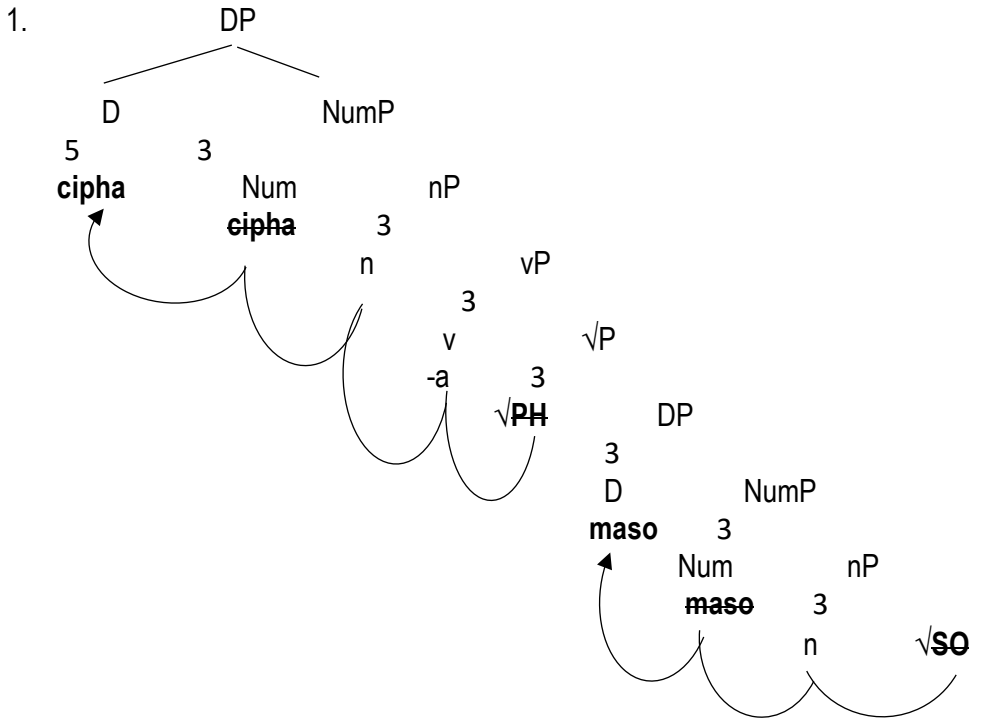
In the second type of phrasal compound, two types of Assc-compounds were examined. In one type, the Assc-compound is almost entirely compositional and perhaps only special in that it is stored as one entry in the lexicon. In the other type, where the Assc-compounds have been analysed in some detail, it is not entirely compositional and has more idiosyncratic meanings. In both types of Assc-compounds, the associative marker links the two nouns together. The study has revealed that the Assc-compounds have an associative phrasal structure (KP) inside comprising the associative (K) and PossDP, and KP shares the gender and number features with the head noun. Syntactically, it has been shown that the associative constructions are DPs because, like the VN-compounds, the entire construction can take the subject and object position, and can be modified and coordinated with other simple DPs. However, the application of the Lexical Integrity tests revealed that the complex expressions can be accessed by syntactic operations such as pronominalisation, gapping and modification. I therefore, concluded that Assc-compounds are not words, but phrasal idioms.

With regards to the morphosyntactic properties of the phrasal compounds, it has emerged in Chapter 3 and Chapter 4 that the phrasal compounds that I analysed in my thesis include inflectional elements. In chapter three, it was shown that the noun class prefixes and the locative prefixes in VN-compounds are inflectional. The noun class prefixes have number features which are realised on a functional category Num (Carstens, 1993, 1997, 2008), and gender features that are realised on the “little” n,

another functional category (Fuchs and Van der Wal, 2018), which acts as a nominaliser for unspecified categorial roots. As regards the locative prefixes, the study has shown in Chapter 3 that these prefixes have been analysed as case markers that head the functional projection Kase (K) (Carstens, 1997, 2008). It has also been shown in Chapter 4 that the associative marker which links the head noun and the PossDP reflects number and gender features of the head noun, and is also a case marker realised as the functional head of a KP projection. This study therefore, has concluded that the NC prefixes and locative prefixes in VN-compounds, and the associative markers in Assc-compounds are all inflectional elements that belong to the functional category Num and Kase.

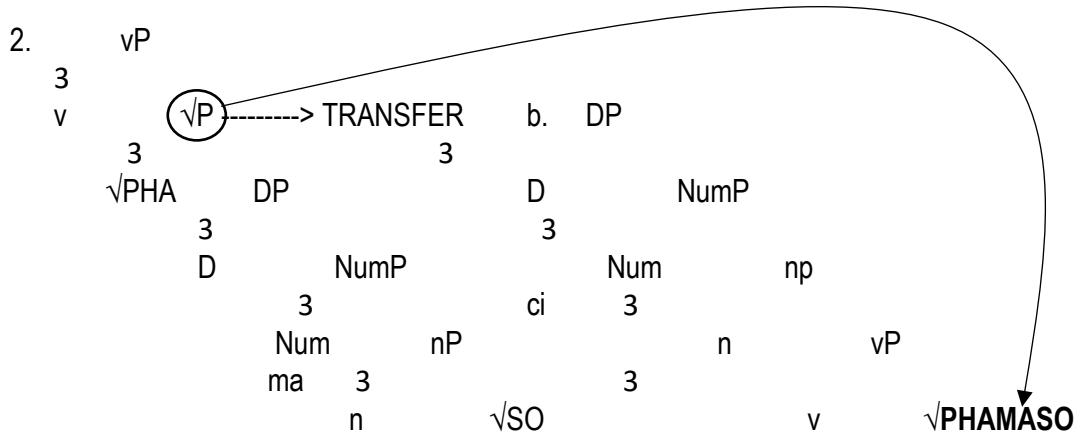
One of the claims that was defended in this thesis is that the nominal part of the VN-compounds is a full DP which merges with an acategorial root which is then verbalised to form the vP. When the vP is then nominalised and Num and D are added we realise a VN-compound which is a full-fledged DP as shown in the structure for *ciphamaso* 'hypocrite'.

1.



The example in (1) above shows that after head movement, the DP **maso** and the root √PHA form the √P which after nominalisation and number is added forms a VN-compound **ciphamaso**. The analysis proposed in this thesis shows that the phrasal part of the compounds which is the vP is a phase, and the head of the vP is a phase head. The analysis of these compounds based on the theory of phases

shows that when a derivation is over, the complement of the phase head v undergoes transfer to the interfaces. Therefore, the root phrase √P, which is PHAMASO in (1) above, is no longer accessible to syntactic operations. Chomsky (2001) calls this the Phase Impenetrability Condition (PIC). I have argued that the transferred phrasal structure becomes a syntactic root which is reified and can re-enter the syntactic derivation as an unanalysable unit. In order to realise a full-fledged VN compound word, the study has shown that the reified root has to merge with little v, the little n, then Num and D as illustrated by the structure for *ciphamaso* 'hypocrite':



In Chapter 3, I have argued that although the reified root looks like a nominalised verb phrase, it is a root and a part of the phrasal compound. The phrasal constituent parts cannot be modified, gapped, extracted or conjoined, because they are represented as a complex reified root. I have concluded that the VN-compounds are genuine phrasal compounds and syntactic words.

It has also been noted that the VN-compounds in NC 7 which refer to humans have an ambiguous behaviour of belonging to class 7 but also optionally agreeing in noun class 1. In the plural, the noun class 2 prefix is attached, and agreement is in class 2:

3. a Ci-gon-a-mu-bawa ci-ja c-a-gul-a nyumba.

	7-sleep-FV-18.LOC-5.bar	7-that	7.SM-PRF-buy-FV	9.house
	'That drunk has bought a house'.			
b.	Ci-gon-a-mu-bawa	u-ja	w-a-gul-a	nyumba
	7-sleep-FV-18.LOC-5.bar	1-that	1.SM-PRF-buy-FV	9.house
	'That drunk has bought a house'.			
c.	A-ci-gon-a-mu-bawa	a-ja	a-na-gul-a	nyumba
	2-7-sleep-FV-18.LOC-5.bar	2-that	2.SM-PST-buy-FV	10.house
	'Those drunks bought a houses'.			

The analysis in (2) is the basis for the analysis of the NC 7 VN-compounds and it can be extended to the agreement ambiguities in (3). It has been shown in Chapter 3 that the ambiguity in the agreement is attributed to delayed reification where the derivation is not complete at vP phase but continues until a syntactically well-formed class 7 DP is realised. When this DP undergoes transfer and it is reified, a new root is formed which now merges with the other n which belongs to gender class A ([+human]). When a singular Num is added, Num is spelled out as zero but when a plural Num is added to n it is spelled out as class 2 plural *a-*. The result is a compound belonging to noun class 1 or 2, but which includes a reified phrase structure nominalised by a class 7 prefix.

With regards to the Assc-compounds, it was concluded that they have the same structure as regular associative compounds. In both associative constructions, K merges with the PossDP to realise the KP phrasal structure while the head noun moves to D. The main difference between these two associative constructions relates to modification. In regular associative constructions, a modifier of the head noun, for example, a demonstrative, can appear between the head noun and the associative phrase (KP) and the associative construction is grammatical as in (4a). However, when the demonstrative appears between the head noun and KP in Assc-compounds as in (4b), the compound loses its idiomatic reading. In contrast, when the demonstrative follows the possessor noun, the idiomatic reading of the Assc-compound is maintained as in (4c).

4.	a.	ci-panda	i-ci	ca	mowa
		7-calabash	7-this	7.ASS	3.beer
		'this calabash of beer'			

b. M-misiri u-yu wa nyumba
1-expert 1-this 1.ASS 10.house
'this expert of houses'; not: 'this builder'

c. M-misiri wa nyumba u-yu
1-expert 1.ASS 10.house 1-this
'this builder'

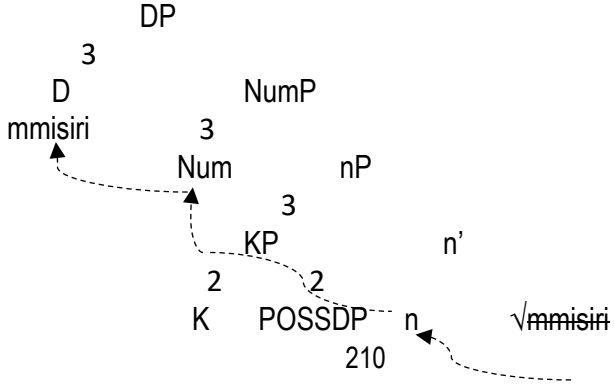
While the contrast between (3) and (4) would suggest at first that Assc-compounds are words, and that the data in (4) demonstrate Lexical Integrity, it was also shown in Chapter 4 that a constituent part of the KP inside the compound can be modified, and the idiomatic reading of the Assc-compound is still maintained. However, when a demonstrative modifier appears between the associative marker and the possessor-DP, the idiomatic reading is lost. For the idiomatic reading to be retained with modification, the demonstrative must appear after the possessor, and only the PossDP is modified:

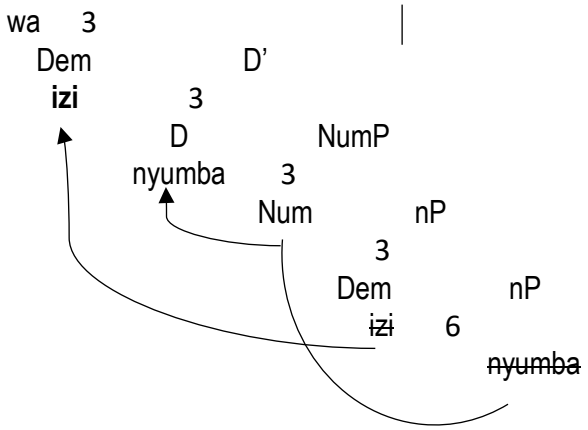
5. a. M-misiri wa nyumba **i-zi** w-a-thyok-a mw-endo.
1-expert 1.ASS 10.house 10-these 1.SM-PRF-break-FV 3-leg
'The builder of these houses has fractured his leg'.

b. M-misiri wa **i-zi** nyumba w-a-thyok-a
1-expert 1.ASS 10-these 10.house 1.SM-PRF-break-FV
mw-endo.
3-leg
'The expert of these huses has fractured his leg'.

The syntactic analysis revealed that the demonstrative modifier in (5a) is located inside the PossDP (i.e. adjoined to nP). As a result it follows the possessor noun which has moved to D:

6.





The associative and the possessor noun are adjacent, and cannot be separated by a demonstrative modifier in Spec, D. I claimed in the thesis that the modifiers following the possessor are adjoined below NumP. In this position, the idiomatic interpretation is preserved despite the occurrence of modifiers.

I also argued in Chapter 4 that the impossibility of modifiers to separate the associative and the possessor was due to the fixed structure of the KP which is specified in the lexical entry of the Assc-compounds. The lexical representation of Assc-compounds showed that the Assc-compounds combine the lexical entries proposed for opaque idioms with a fixed phrasal structure in LSS and that of transparent idioms with some pieces of the LSS not connected. The KP-part of an Assc-compound was analysed as having a fixed structure which includes the DP and NumP. As a result, Assc-compounds allow modification, but only with modifiers that are attached below NumP. A demonstrative modifier in Spec, D would interrupt the lexically fixed structure of the idiom and destroy the idiomatic reading.

The syntactic analysis of Assc-compounds that allow internal modification was then extended in Chapter 4 to show that Assc-compounds may allow activation of not only the idiomatic meaning but also, at the same time, of the regular meaning of its parts:

7.	a.	M-misiri	wa	nyumba	za-zi-kulu
		1-expert	1.ASS	10.house	10ASS-9-big
		'The expert of big houses'			
		'The builder of big houses'			

In (7) above, a part of the idiom, namely the possessor *nyumba* is modified and the idiom preserves the idiomatic reading and at the same time it seems to activate the regular meaning of individual parts

of the idiom. The analysis of these Assc-compounds revealed that this simultaneous meaning is activated because the terminal node for the modifier is not linked to any part of the LCS of the Assc-compound but allows for the terminal nodes corresponding to the parts of the idiom to have access to the regular meaning of the head noun and the entire KP-phrase. Therefore, I concluded in Chapter 4 that the fact that syntactic operations such as modification can have access to constituent parts of the Assc-compounds, shows that these constructions are not genuine words that can be reified but that they are idiomatic phrases with lexicalised meaning.

The parallel between the VN-compounds and the Assc-compounds is that they both have inflectional material inside them. In the VN-compounds, the NC prefixes have number features and the locative prefixes are case markers and thus they are inflectional and belong to a functional category. Additionally, both phrasal compounds include phrasal structure: a VP in the case of VN-compounds and a KP in the case of Assc-compounds.

The major difference that has been noted in the thesis between the two types of phrasal compounds is that in VN-compounds, the phrasal part is a syntactic phrase which can undergo reification as an unanalysable unit and be re-introduced into the derivation as a root. It has been concluded that compounds built on these roots are syntactic words. On the other hand, in Assc-compounds, the phrasal part allows syntax to see inside the phrasal compounds and Lexical Integrity is violated. It has been concluded that the phrasal part is not a root that can be reified but that the Assc-compounds are phrasal idioms with lexicalised meaning.

From a theoretical standpoint, two main theories have been used to account for the properties of the two types of phrasal compounds namely: Distributed Morphology, and Lexical Licensing theory. The DM framework accounts for the root analysis of VN-compounds (Halle and Marantz, 1993; Marantnz, 1997; Harley, 2009). In DM, a root phrase √P is acategorial until it merges with a functional head (e.g. little v or little n) which is a category defining feature to derive vPs or nPs. The theory of Lexical Licensing was applied to explain the lexical representation of Assc-compounds as phrasal idioms and how their lexical entries are linked to the three grammatical systems; phonology, syntax and semantics (Jackendoff, 1997). The analysis of complex DPs in Bantu was based on the work of Carstens (1991, 2008, 2018). The theories showed a high level of compatibility in the analysis of the two types of phrasal compounds because the three systems of grammar in Lexical Licensing theory showed some relatedness to the three lists in DM as follows: the phonological system is linked to vocabulary items, the syntactic system to syntactic terminals and the semantic system is linked to the encyclopedia.

5.2 Contributions to the field of linguistics

As indicated in Chapter 1, one of the main reasons that necessitated this study was the need for a detailed analysis of phrasal compounds in Cinyanja, which have not received much attention. As noted, there is a niche as regards the internal structure of compound words and the role of the inflectional material internal to compound words in Cinyanja. Therefore, this study has contributed to the syntax and morphology of Cinyanja nominal compounds. Chapter 1 has presented a record of Cinyanja noun class prefixes which are a prominent feature in Cinyanja phrasal compounds. Chapter 2 is a critical review of the theory of compounding. The main contribution of this chapter to the study is the identification of the definition that best suits the compound words in Cinyanja from the many definitions provided by scholars. Additionally, the discussion of phrasal compounds provided some insight into a better understanding of phrasal compounds in Cinyanja. Further, the presentation of the morphological and the syntactic theories of synthetic compounds provided the basis on which VN-compounds were analysed.

In Chapter 3, the main contribution is the analysis of lexical phrases as reified roots (Harley, 2009) which has not so far taken ground in Bantu linguistics. I have emphasised how lexicalised phrasal structures can re-enter the syntax as unanalysable roots and realise their compound status after nominalisation. One of the main empirical contributions of this chapter which is quite intriguing is the observation that class 7 VN-compounds can agree in two genders, gender A and gender D.

In Chapter 4, the main contribution is the analysis of Assc-compounds as phrasal idioms. Here, the major empirical contribution is the observation that some of the Assc-compounds can express idiomatic meaning even when an internal part is modified and at the same time retain the literal meaning of their parts. The lexical analysis of Assc-compounds in this chapter was based on the theory of lexical licensing of idioms (Jackendoff, 1997).

5.3 Recommendations for future research

In Cinyanja, when the demonstrative and the possessive co-occur DP internal, the possessive precedes the demonstrative. Even when they come after the possessor noun, the same word order seems to apply. This suggests some restrictions in the ordering of these words and needs further

investigation. It would be good to establish why the word order of the possessive pronoun and the demonstrative seems inflexible in Cinyanja/Chichewa.

In Chapter 3, it was shown that the complement of the phase head, the root phrase √P in a vP can be reified as unanalysable root. However, it would be good to establish if the little a and n can also be considered phase heads whose complements can be reified as unanalysable roots in Cinyanja or other Bantu languages. It is not clear what exactly constrains reification (other than the phase theory). Why is reification not possible with every other DP, for example Assc-compounds? It would be good to establish if any complement of the phase head can in principle be used as a root.

Another aspect concerns noun class 7 VN-compounds which refer to humans. These compounds can also behave like phrasal roots when they assume NC 1 status and are pluralised in class 2. I claimed that in the formation of these compounds the derivation is not completed at vP but extended until it derives a syntactically well-formed DP. Only then does the root node undergo transfer and since reification can apply to the output of the transfer, the entire DP with NC 7 prefix is reified and reanalysed as a new root. The root is nominalised, then Num (i.e. a null NC 1 prefix or a NC 2 prefix) and D is added to derive the class 1 or class 2 DP. One open aim of further investigation could be to establish why only DPs based on VN-compounds can be reified into class 1, and not other class 7 human nouns. Furthermore, it would be an interesting question for future research to establish whether the derivation of diminutives and similar secondary derivations can also be explained in terms of reification of the structure corresponding to the underlying base. This would challenge B & M's (1995) claim that the opaqueness of the base nouns in these derivations is an argument for their claim that noun class prefixes are morphological. I consider it possible that the reason why "inner" nouns are no longer accessible for modification as in (8b) below is because of the reification process.

8. a. Ka-m-pal-a-ma-tabwa ka-ka-fupi k-a-psy-a.
12-1-scrape-FV-6-plank 12.ASS-6-short 1-BE.PRS-burn-FV
'A small/slim short carpenter is burnt'.

b. *Ka-m-pal-a-ma-tabwa a-a-fupi a-a-psy-a.
12-1-scrape-FV-6-plank 6.ASS-6-short 6-BE.PRS-burn-FV

5.4 Summary of the chapter

In conclusion, the analysis of Cinyanja phrasal compounds has revealed a lot about the structure of phrasal compounds, namely: that the elements that form these compounds are ether inflectional markers or roots. It has also been concluded that vP phrases in VN compounds are also roots. Further, my study has shown that not all complex phrasal expressions with lexicalised meaning are phrasal idioms, but that some of them are compound words that include phrases reified as roots. It has been emphasised in this study that the Lexical Integrity Hypothesis plays a big role in determining the status of complex expressions as words or phrasal constructs.

References

Abney S. (1986). Functional Elements and Licensing. Paper presented to GLOW. Gerona, Spain.

Abney, S. P. (1987). *The English Noun Phrase in Its Sentential Aspect*. PhD Thesis. Cambridge, MIT Mass.

Ackema, P & Neeleman, A (2001), Competition between Syntax and Morphology. In Legendre, G, Grimshaw, J. and Vikner, S. (eds). *Optimality Theoretic Syntax*. MIT Press, 29-60.

Ackema, P. and Neeleman, A. D. (2004*). Beyond Morphology*. Interface Conditions on Word Formation: Studies in Theoretical Linguistics.

Acquaviva, P. (2014). The Roots of Nominality, the Nominality of Roots.*The Syntax of Roots and the Roots of Syntax*, (ed.). Artemis Alexiadou, Hagit Borer, and Florian Schaefer.Oxford: Oxford University Press, 33-56.

Altakhaineh, A.R.M. (2016a). *Compounding in Modern Standard Arabic, Jordanian Arabic and English*: PhD Thesis. New Castle University.

Altakhaineh, A.R.M. (2016b). What is a Compound? The Main Criteria for Compoundhood. *ExELL (Explorations in English Language and Linguistics*) 4 (1): 58-86.

Anderson, S.R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Baker, M.C. (1988). *Incorporation*. Chicago and London: Chicago University Press.

Baker, M. (2003). *Lexical categories: Verbs, Nouns and Adjectives*. Cambridge: Cambridge University Press.

Baker, M. (2008). *The Sytax of Agreement and Concord*. Cambridge: Cambridge University Press.

Baker, M.C. and Fasola, C.A (2009). Araucanian: Mapudungun: In Lieber, R. and Štekauer, P. (eds) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 594-608.

Basciano, B., Kula, N. and Melloni, C. (2011). Modes of Compounding in Bantu, Romance and Chinese. *Rivista di Linguistica*, 23(2): 203-249.

Bauer (2009). Typology of compounds. In R. Lieber and P. Štekauer (eds.). *The Oxford Handbook of Compounding.* Oxford: Oxford University Press, 343-356.

Bauer, L. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.

Bauer, L. (1998a). When is a Sequence of two Nouns a Compound in English? English *Language and Linguistics* 2: 65–86.

Bauer, L. (2001). 'Compounding'. In Haspelmath, M., Ekkehard, K.., Wulf, Ö. and Wolfgang, R. (eds) *Language Typology and Language Universals*. Berlin: Mouton de Gruyter, 695-707.

Bauer, L. (2004). Adjectives, Compounds, and Words. *Nordic Journal of English Studies*, 3(1), 7-22.

Bauer, L. (2008). Exocentric Compounds. *Morphology* 18:51-74.

Bauer, L. (2009). 'Typology of Compounds'. In Lieber, R. and Štekauer, P. (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 343-356.

Bauer, L., Lieber, R. and Plag, I. (2013). *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.

Bentley, M. (1998). The Marking of Grammatical Relations in Swahili'. *Studies in African Linguistics,* (27) 177-197.

Bisetto, A. and Scalise, S. (2005). 'The Classification of Compounds'. *Lingue e linguaggio,* 4(2): 319-332.

Booij, G. (1992). Morphology, Semantics and. *Thematic structures: Its role in Grammar*, 47

Booij, G. (2005). Compounding and Derivation: Evidence for Construction Morphology. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, *264*, 109.

Booij, G. (2009). Compounding and Construction Morphology. In Lieber, R. and. Štekauer, P. (eds). *The Oxford Handbook of Compounding.* Oxford: Oxford University Press, 201-216.

Borer, H. (2005). *Structuring Sense*. Oxford: Oxford University Press.

Borer, H. (2009). Afro-Asiatic, Semitic: Hebrew. In Lieber, R., Štekauer, P. (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 386–399.

Bosch, S. E. (1997). Possible Origins of the Possessive Particle-ka-in Zulu. *South African Journal of African Languages*, *17*(1), 1-5.

Bošković, Z. (2014). Now I am a Phase, now I am a Phase: on the Variability of Phases with Extraction and Ellipsis: *Linguistic Inquiry*, 45: 27-89.

Botha, R.P. (1980). Word Based Morphology and Synthetic Compounds. *Stellenbosch Papers in Linguistics, 5: 1-166.*

Bresnan, J. (1991). Locative Case vs. Locative Gender. In *Annual Meeting of the Berkeley Linguistics Society* 17, 1, 53-68).

Bresnan, J. (1994). Locative Inversion and the Architecture of Universal Grammar. *Language,* 70(1): 72–131.

Bresnan, J. and Kanerva, J.M. (1989). Locative Inversion in Chichewa. *Linguistic Inquiry* 20, 1-50.

Bresnan, J. and Mchombo, S. (1987). Topic, Pronoun, and Agreement in Chichewa. *Language*, 63, 741-782.

Bresnan, J. and Mchombo, S. (1995). The Lexical Integrity Principle: Evidence from Bantu languages. *Natural Languages & Linguistic Theory*. Springer, 13 (2): 181-254.

Bresnan, J. and Mugane, J. (2006). Agentive Nominalizations in Gıkuyu and the Theory of Mixed Categories. *Intelligent Linguistic Architectures: Variation on Themes by Ronald M. Kaplan*, 201-234.

Buell, L.C., Riedel, K. and van der Wal, J. (2011). What the Bantu Languages can Tell us about Word Order and Movement. *Lingua*, *121*(5): 689-701.

Caha, P. and Pantcheva, M. (2015). Locatives in Shona and Luganda. Unpublished manuscript, Lingbuzz/002220.

Carstairs-Mc Carthy, A. (2002). *An introduction to English Morphology: Words and their Structure*. Edinburgh: Edinburgh University Press.

Carstens, V. (1991). *The Morphology and Syntax of Determiner Phrases in Kiswahili*, PhD Thesis. University of California, Los Angeles.

Carstens, V. (1993). On Nominal Morphology and DP Structure. In Sam A. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*, 151−180. Stanford: CSLI.

Carstens, V. (1997). Empty Nouns in Bantu Locatives. *The Linguistic Review,* 14 (4): 361-410.

Carstens, V. (2000). Concord in Minimalist Theory*. Linguistic Inquiry.* The MIT Press*,* 31(2): 319–355.

Carstens, V. (2008). DP in Bantu and Romance. In K. Demuth and C. De Cat (eds.). *The Bantu-Romance connection.* Amsterdam: John Benjamins, 131–166.

Carstens, V. (2005). Agree and EPP in Bantu. *Natural Language & Linguistic Theory,* 23(2), 219-279.

Carstens, V. (2011). Hyperactivity and Hyperagreement in Bantu. *Lingua,* 121: 721−741.

Carstens, V. (2018). *Concord and Labeling*. Southern Illinois University.

Carstens, V. and Zeller, J. (2020). 'Only'in Nguni: A phrase-final particle meets antisymmetry theory. *Linguistic Inquiry*, 51(2), 199-235.

Centnarowska, B. (2016). Headedness of Coordinate Compounds in Polish and English. *Studies in Lexicogrammar. Theory and Application,* 54, 234.

Chomsky, N. (1970). Remarks on Nominalization. In *Readings in English Transformational Grammar.*

Chomsky, N. (1986). *Knowledge of Language: its Nature, Origins, and Use*. New York: Praeger.

Chomsky, N. (1995). *Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist Inquiries: The Framework. In R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by step: Essays on Minimalist Syntax in Honour of Howard Lasnik* (pp. 89 - 155). Cambridge, Massachusetts: MIT Press.

Chomsky, N. (2001). Derivation by Phase. In *Ken Hale: A Life in Language*, M. Kenstowicz (ed.), 1-52. Cambridge: MIT Press.

Chomsky, N. (2007). Approaching U.G. from below. *Interface +Recursion=Language? Chomsky's Minimalist and the New from Syntax-Semantic*s, 1-29.

Chomsky, N. (2008). On Phases. In *Foundational issues in linguistic theory*. R. Freidin, C.P. Otero, & M.L. Zubizarreta, Eds. Cambridge Mass: The MIT Press. 132–166.

Clark, E.V., Hecht, B.F. and Mulford, R.C. (1986). Coining Complex Compounds in English: Affixes and Word Order in Acquisition. *Linguistics*, 24:7-29.

Conti, S. (2006). *Compound Adjective in English. A descriptive approach to their morphology and functions*. PhD Thesis. University of Pisa, Italy.

Corbett, G.G. and Mtenje, A.D. (1987). Gender Agreement in Chichewa. *Studies in African Linguistics*, 18(1):1–38.

Creissels, D., Dimmendaal, G., Frajzyngier, Z. and König, C. (2008). Africa as a Morphosyntactic Area. In Heine, B. & Nurse, D. (eds), *A Linguistic Geography of Africa. 86-150*. Cambridge University Press.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler Syntax*. Oxford, UK: Oxford University Press.

De Belder, M. (2013). The root and Nothing but the Root: Primary Compounds in Dutch. *Syntax,* 20(2): 138-168.

De Dreu, M. (2008). *The Internal Structure of the Zulu DP*. MA. Thesis. Leiden University, Netherlands.

Delfitto, D, Fiorin, G. and Kula, N. (2011). Syntactic Gradients in Compounding: Bemba Associative Nominals vs. Italian Prepositional and Deverbal Compounds. *Rivista di Linguistica*, 23 (2): 301-325.

Den Dikken, M. (2007): Phase Extension. Contours of a Theory of the Role of Head Movement in Phrasal Extraction. *Theoretical Linguistics*, 33: 1-41.

Diercks, M.J.K. (2010). *Agreement with Subjects in Lubukusu*. PhD. Thesis. Georgetown University: Washington, DC.

Di Sciullo, A.M. (2009). Why are Compounds a Part of Human Language? A view from Asymmetry Theory. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 145-177.

Di Sciullo, A.M. and Williams, E. (1987). *On the Definition of Word*. Cambridge Mass: MIT Press.

Embick, D. and Noyer, R. (2007). Distributed Morphology and the Syntax/Morphology Interface. In G. Ramchand and C. Reiss (eds.). *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, 289-324.

Embick, D., and Marantz, A. (2008). Architecture and Blocking. *Linguistic inquiry*, *39*(1): 1-53.

Englund, H. (2001). Chinyanja and the Language rights. Nordic Journal of African studies,Uppsala, Sweeden, 10(3): 299-319.

Fabb N. (1998) Compounding. In: Spencer A., Zwicky A.M. (eds). *The Handbook of Morphology*. Oxford: Blackwell, 66–83.

Fabb, N.A.J. (1984). *Syntactic affixation*. PhD Thesis. Massachusetts Institute of Technology, Boston.

Ferrari, F. (2005). *A Syntactic Analysis of the Nominal System of Italian and Luganda – how nouns can be formed in the syntax*. NYU PhD Thesis.

Fiorentino, R.D. (2006). *Lexical Structure and the Nature of Linguistic Representations*. PhD Thesis. University of Maryland, College Park.

Fortune, G. (1985). *Shona Grammatical Constructions: Parts 1*, 3rd edition. Harare: Mercury Press.

Fuchs, Z. and Van de Wal, J. (2018). Nominal Syntax in Bantu languages: How Far can we Get with Gender on *n*? *Talk presented at Standford Smircle.*

Gallego, A. (2007). *Phase Theory, and Parametric Variation*. PhD Thesis. University of Barcelona: Barcelona.

Guevara, E. and Scalise, S. (2009). Searching for Universals in Compounding. *Universals of Language Today,* 101-128.

Güldemann, Tom (1999). 'Toward a Grammaticalization and Typological Account of the *ka* possessive in Southern Nguni'. *Journal of African Languages and Linguistics* 2(2): 157-184.

Guthrie, M. (1967). *The Classification of Bantu Languages*. London: Dawsons of Pall Mall.

Guthrie, M. (1971). *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages.* Farnborough: Gregg Press.

Haegeman, L. (1994). *Introduction to Government and Binding Theory*. UK Blackwell publishers.

Halle, M. and Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection. In: K. Halle and S.J. Keyser (eds.). *A View from Building 20: Essays in Linguistics in Honour of Sylvian Bromberger*. Cambridge, Mass. MIT Press, 111-176.

Halpert, C. (2012). *Argument Licensing and Agreement in Zulu*. PhD Thesis. Massachusetts Institute of Technology: Cambridge.

Hamungole, M. (2015). *Television and the Cultivation of Personal Values among Catholics in Zambia*. PhD Thesis. Leuven, Belgium.

Harley, H. (2009). Compounding in Distributed Morphology. In R. Lieber and P. Štekauer (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 129-144.

Harley, H. (2014). On the identity of roots. *Theoretical Linguistics*, 40: 225-276.

Harley, H., & Noyer, R. (1999). Distributed Morphology. *Glot International, 4*(4), 3 - 9.

Haspelmath, M. (2002). *Understanding* Morphology. Arnold: London.

Haspelmath, M. and Sims, A.D. (2010). *Understanding Morphology*. Oxford: Oxford University Press.

Haugen,E.(1966). Dialect, Language, Nation 1. *American Anthropologist*, 68(4): 922-935.

Henderson, B. (2006). Multiple agreement, concord and case checking in Bantu. In Arasanyin, Olaoba F. & Michael A. Pemberton (eds.), *Selected proceedings of the 36th Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 60-65.

Hsieh, F.F. and Sybesma, R. (2011). On the Linearization of Chinese Sentence-final Particles: Max spell out and why CP moves. *Korea Journal of Chinese Language and Literature*, 1 (1), 53-90.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, Massachusetts: MIT Press.

Katamba, F. and Stonham J.T. (2006). Morphology. Basingstoke: Palgrave Macmillan.

Katamba, F. (2003). Bantu nominal morphology. In Derek Nurse & Gerard Philippson (eds.) *Bantu Languages*. London: Routledge, 103-120.

Kashoki, M.E. (1978). A General Sketch of the Bantu Languages of Zambia*. Language in Zambia*. (S Ohanessian and M.E Kashoki, eds). London, international African institute, 47-100.

Kavka, S. (2009). 'Compounds and Idiomatology'. In Lieber, R. and Štekauer, P. (eds) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 19-33.

Kayne, R. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kiefer, F. (2009). Uralic, Finno-Ugric: Hungarian. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 527-541.

Kinyalolo. K. (1991). *Syntactic Dependencies and the SPEC-head Agreement Hypothesis in Kilega*.

Ph.D. Thesis. University of California. Los Angeles.

Kiso, A. (2012). *Tense and Aspect in Chichewa, Chitumbuka and Cisena. A Description and Comparison of the Tense-aspect system in three southern Bantu languages.* PhD Thesis. Stockholm University.

Kornfeld, L.M. (2009). IE, Romance: Spanish. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 436-452.

Kosch, I. (2011) Descriptive Issues of Derivation and Inflection in Sesotho sa Leboa (Northern Sotho) with particular reference to 'number'. *Southern African Linguistics and Applied Language Studies*, 29(1): 89-96.

Kramer, R. (2015). *The Morphosyntax of Gender*. Oxford: Oxford University Press.

Kratzer, A. and Selkirk, E. (2007). Phase Theory and Prosodic Spellout: The case of verbs. *The Linguistic Review,* 24(2-3): 93-135.

Kremers, J. (2003). The Noun Phrase in Arabic: A Minimalist Approach. PhD Thesis. University of Nijmegen, Netherlands.

Kula, N. C. (2012). Nominal Compounding and Associative Phrases in Bemba. In M. Brenzinger and A. Fehn (eds.). *Proceedings of the 6th world congress of African Linguistics*. (Cologne 2009). Cologne: Koppe 431-441.

Langaendoen, D.T. (1970). The Accessibility of Deep Structure. *Readings in English transformational Grammar***,** ed. By Roderick A. Jacobs and Peter S. Rosenbaum, 99-104, Waltham, MA: Ginn and Company.

Lapointe, S. (1980), A Theory of Grammatical Agreement. PhD. Thesis. Amherst, University of Massachusetts.

Letsholo, R. (2004). Clausal and DP-internal Agreement in Ikalanga. *Studies in African Linguistics*, 33(1):91–127.

Lieber, R. (1983). Argument Linking and Compounds in English. *Linguistic Inquiry*. The MIT Press, 14 (2): 251-285.

Lieber, R. (1992). *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago, Illinois: University of Chicago Press.

Lieber, R. (2009a*). Introducing Morphology*. New York: Cambridge University Press.

Lieber, R. (2009b). A Lexical Semantic Approach to Compounding. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 78-104.

Lieber, R. (2009c). IE, Germanic: English. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 357-369.

Lieber, R. and Stekauer, P. (2009). Introduction: Status and Definition of compounding. In R. Lieber and P. Štekauer (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 3-18.

Lipka, L. (1992). *An outline of English Lexicology: Lexical Structure, Word Semantics, and Word-Formation*. Tubingen: Niemeyer.

Lusekelo, A. (2013). DP-internal and V-external Agreement Patterns in Eastern Bantu: Re-statement of the Facts in Eastern Bantu. *Journal of Linguistics and Language in Education*, 7(1):19–47.

Lusekelo,A. (2009). The structure of Nyakyusu Noun Phrase. Nordic Journal of African Studies, 18(4): 305-331.

Marantz, A. (1997). No Escape from Syntax: don't Try Morphological Analysis in the Privacy of your own Lexicon. *UPenn Working Papers in Linguistics*, 4: 201-225.

Marantz, A. (2001). Words. Ms., Massachusetts Institute of Technology.

Marantz, A. (2007). Restitutive re and the First Phase Syntax /Semantics of VP. *University of Maryland, college park, 20.*

Marchand, H. (1960) *The Categories and Types of Present-Day English Word-formation: A Synchronic-Diachronic Approach*. Wiesbaden: Otto Harrassowitz.

Marten, L. (2012). Agreement in Locative Phrases in Luganda. In M. Brenzinger and A. Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics (Cologne 2009)*, 433−443. Cologne: Köppe.

Marten, L., & Kula, N. C. (2007). Zambia 'One Zambia, one nation, many languages'. *Language and National Identity in Africa*, 291-313.

Matambirofa, F. (2000). Some Aspects of the Architecture of the Possessive Noun Phrase in Bantu. *Zambezia*, 27 (1):71-92.

Mchombo, S. A. (1978*). A Critical Appraisal of the Place of Derivational Morphology within Transformational Grammar, Considered with Primary Reference to Chichewa and Swahili*, PhD Thesis. School of Oriental and African Studies, University of London.

Mchombo, S.A. (2004). *The Syntax of Chichewa*. Cambridge: Cambridge University Press.

Miti, L. (2001*). A linguistic Analysis of Cinsenga. A Bantu language spoken in Zambia and Malawi*. Cape Town: The Centre for Advanced Studies of African Society (CASAS).

Miti, L. M. (2006). *Comparative Bantu Phonology and Morphology: a Study of the Sound Systems and Word Structure of the Indigenous Languages of Southern Africa*. Centre for Advanced Studies of African Society (CASAS).

Mletshe, L. (2017). Deverbal Nominals Derived from Intransitive State Verbs in isiXhosa: A Generative Lexicon approach. *South African Journal of African Languages*, 37:1. 29-39.

Mletshe, L.K. (2010). *Deverbal nominals in Xhosa*. PhD Thesis. University of Stellenbosch, South Africa.

Mtenje, Al. (1987): Tone Shift Principles in the Chichewa Verb. *Lingua* 72, (72):169-209.

Mufwene, S. (1980). Bantu Class Prefixes: Inflectional or Derivational? In *Papers from the 16th Regional Meeting Chicago Linguistics Society*. J. Kreiman & A.E. Ojeda, Eds. Chicago: CLS. 246–258.

Mugane, J. M. (1997). *Bantu Nominalization Structures*. PhD Thesis. The University of Arizona.

Mukai, M. (2006). *A comparative Study of Compound Words in English, Japanese, and Mainland Scandinavian*. PhD Thesis. University of Newcastle, England.

Msaka, P.K. (2019). *Nomnal Classification in Bantu Revisited: The Perspective from Chichewa*. PhD Thesis. University of Stellenbosch, South Africa.

Myers, P.S. (1987). *Tone and the Structure of Words in Shona*. PhD Thesis. University of Massachusetts, Boston.

Nakov, P (20:13). On the Interpretation of Noun Compounds: Syntax, Semantics, Entailment. *Natural Language Engineering.* United Kingdom: Cambridge University Press, 1(1):1-40.

Neef, M. (2009). 'IE, Germanic: German'. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 386-399.

Nunberg, G, Sag, I and Wasow, T. (1994). Idioms. *Language* 70, 49 1 -538.

Nurse, D. (2008). *Tense and Aspect in Bantu*. Oxford: Oxford University Press.

Okubo, T. (2013). Two Types of Compounds in Distributed morphology. *Tsukuba English Studies*, 32:147- 158.

Olsen S. (2017), Synthetic Compounds from a Lexicalist Perspective. In.Vogel, P.M,.Donalies, E,. Eichinger, L.M, Habermann, M, Maibauer J., Schluker, Solms, B., H. J, Stekauer, P. and Hernandez, S.V. (eds.). *Journal of Word Formation*, an Open Access Journal, 17-45.

Paas, S. (2016). *Oxford Chichewa-English/ English-Chichewa Dictionary.* Cape Town: Oxford University Press ORBIS (Pty) Limited.

Paas, S. (2005). *English Chichewa-Chinyanja Dictionary*. Malawi: A.G.L.C Press.

Paas, S. (2013). *Chichewa/Chinyanja English Dictionary*. Blantyre: Christian Literature Action in Malawi.

Padrosa-Trias, S. (2009). Are there Coordinate Compounds? In Ralli, A., Booij, G., Scalise, S. and Karasimos, A. (eds.). *Mediterranean Morphology Meeting*, 7:98-110.

Padrosa-Trias, S. (2010). *Complex Word-Formation and the Morphology-Syntax Interface*. PhD Thesis. Universidad Autónoma de Barcelona, Barcelona.

Pafel, J. (2017). Phrasal Compounds and the Morphology-Syntax Relation. In Trips, C & Kornfilt, J. (eds.). *Further Investigation into the Nature of Phrasal Compoundin*g. Berlin: Language Science Press, 233-259.

Pesetsky, D. and Torrego, E. (2001). 'T to C Movement: Causes and Consequences'. In Michael Kenstowicz (ed.).*Ken Hale: a Life in Language*, Cambridge, MA: MIT Press, 355–426.

Plag, I. (2003). *Word-Formation in English*. Cambridge: Cambridge University Press.

Poulos, G. (1990). *A Linguistic Analysis of Venda***.** Via Afrika Limited.

Radford, A. (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Ralli, A. (2010). 'Compounding Versus Derivation'. In Scalise, S and Vogel, I (eds). *Cross Disciplinary Issues in Compounding*. Amsterdam: John Benjamins Publishing. 57-76.

Ritter, E. (1992). Cross-Linguistic Evidence for Number Phrase. *The Canadian Journal of Linguistics,* 37 (2): 197-218.

Roeper, T. and Siegel, M.E.A. (1978). A Lexical Transformation for Verbal compounds. *Linguistic inquiry*, The MIT Press, 9(2):199-260.

Rosenberg, M. (2007). Classification, Headedness and Pluralization Corpus Evidence from French Compounds. *Acta Linguistica Hungarica*, 54 (3):341-360.

Rugemalira, M.J. (2007). *The Structure of Bantu Noun Phrase*. SOAS Working Papers in Linguistics 15: 135–148.

Saeed, J. (2003). *Semantics.* 2nd ed. Oxford: Blackwell.

Solano-Flores, G. (2006). Language, Dialect and Register: Sociolinguistic and the estimation of Mesurement Error in Testing of English Learners. *Teachers College Record*, 108(11): 2354-2479

Savini, M. (1984). Phrasal Compounds in Africans: A Generative Analysis. *Stellenbosch Papers in Linguistics*, 12: 34-114.

Scalise, S. and Guevara, E. (2005). The Lexicalist Approach to Word Formation. In Dordrecht: Springer, 147-187

Scalise, S. and Bisetto, A. (2009). The Classification of Compounds. In Lieber, R. and Štekauer, P. (eds.). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 34-53.

Scalise, S., Fabregas, A. and Forza, F. (2009). Exocentricity in Compounding. *言語研究* (*Gengo Kenkyu),* 135, 49-84.

Schadeberg, T.C. (2001). Number in Swahili Grammar. *Swahili Forum VIII. Afrikanistische Arbeitspapiere*. 68:7–16.

Selkirk, E.O. (1981). English Compounding and the Theory of Word Structure. In Moorigat et.al (eds.), 229-277

Selkirk, E.O. (1982). *The syntax of Words*. Cambridge, Mass: MIT Press.

Sideeg, A. (2016). Key Arguments for the Determiner Phrase Hypothesis. *IOSR Journal of Humanities And Social Science (IOSR-JHSS)*, 21 (2): 18-25.

Scotton, C.M. (1980). *Learning Chichewa Book 1*. African Studies Center. Michigan State University.

Spreda, K.W. (2000).*The Associative Noun Phrase in Metta*. Cameroon: SIL Cameroon.

Sproat. R. (1985) *On Deriving the Lexicon.* PhD Thesis. MIT, Cambridge.

Szymanek, B. (2009). IE, Slavonic Polish. In Lieber, R. and Štekauer, P. (eds). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 465-477.

Taraldsen, T. (2010). The Nanosyntax of Nguni Noun Class Prefixes and Concords. *Lingua,* 120:1522-1548.

Trips, C. (2012). Empirical and Theoretical Aspects of Phrasal Compounds: against the "syntax explains it all" attitude. *In Mediterranean Morphology Meetings*, 8: 322-346.

Uriagereka, J. (1999). Minimal Restrictions on Basque Movements. *Natural Language & Linguistic Theory*, 17(2), 403-444.

Van de Velde, M. L.O. (2006). Multifunctional Agreement Patterns in Bantu and the Possibility of Genderless Nouns. *Linguistic Typology*, 10 (2), 183-221.

Van de Velde, M. (2013). The Bantu Connective Construction. *The Genitive* (Case and grammatical relations across languages 5). A. Carlier and J.C. Verstraete (eds.). Amsterdam: John Benjamins, 217-252.

Van der Spuy A. (2010). Generating Zulu Noun Class Morphology. *Language Matters,* 41(2): 294–314.

Vermeullen, J. (1979). *Chinyanja (Chichewa)-English Dictionary*. Lusaka: Mission L. C.

Welmers, W.A. (1963). Associative *a* and *ka* in Niger-Congo. *Language*. Language society of America, 39 (3): 132-147.

Wiese, R. (1996). Phrasal Compounds and the Theory of Word Syntax. *Linguistic Inquiry.* The MIT Press, 27 (1):183-193.

Williams, E. (1981). On the Notions "Lexically Related" and "Head of a Word". *Linguistic Inquiry* 12(2): 245–274.

Zeller, J and Ngoboka, J.P. (2018). Agreement with Locatives in Kinyarwanda: a Comparative Analysis. *Journal of African Languages and Linguistics* 39 (1): 65-106.

Zeller, J. (2017a). A note on Bantu Extensions and Syntactic Word Formation. *Nordic Journal of African Studies*, *26*(4).

Zeller, J. (2017b). A Note on Root Projection and Labelling. *Stellenbosch Papers in Linguistics*, 48, 197-204

Zeller, J. (2018). Locatives in Bantu. Unpublished Manuscript, University of KwaZulu-Natal, Durban.

Ziering, P. and Van der Plas, L. (2014). What Good are 'Nominalkomposita' for 'Noun Compounds': Multilingual Extraction and Structure Analysis of Nominal Compositions Using Linguistic Restrictors. *In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 1047-1058.