

SESO THO COMPOUNDING

A thesis submitted in fulfilment of the
requirements for the degree of
MASTER IN LINGUISTICS

at

THE UNIVERSITY OF THE FREE STATE
Department of Linguistics and Language Practice

by

'MASEANKOENA AMINA MOKOALELI

Supervisor: Dr Kristina Riedel

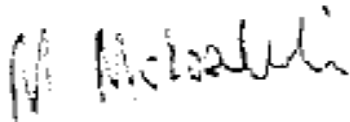
Co-supervisor: Prof Nancy Kula

November 2020

DECLARATION

I, 'MASEANKOENA AMINA MOKOALELI, hereby declare that the work on this thesis is my original work (all sources have been acknowledged) and that neither the whole work nor any part of it has been submitted at the University of the Free State nor any other university for the award of any degree.

Signature:



Date: 30/11/2020

ABSTRACT

This study explores the formation of Sesotho compounds from descriptive and theoretical perspectives. Attention is specifically given to the compounds' combinatorial patterns, morphological structure, number, agreement, headedness and properties which distinguish compounds from similar constructions. Compounding has featured quite extensively in the linguistic literature and it is a highly controversial phenomena in terms of whether compounds are morphological or syntactic constructions. Descriptively, this study reveals that Sesotho compounds are morphological constructions which morphosyntactically behave like simple nouns. This, therefore leads to the conclusion that Sesotho compounds are different from similar syntactic constructions, being phrases. The difference between the two is explained within the Lexical Integrity Principle.

Theoretically, Sesotho compound data is accounted for within the Optimality Theoretic framework in which specific constraints set forth necessary conditions for constructions which qualify as compound nouns in Sesotho. This happens following ranking of constraints in strict hierarchy of domination. The shortcoming of standard Optimality Theory to account for Sesotho compound data following the ranking of constraints in strict hierarchy of domination is complemented by manipulation of Cophonology which allows re-ranking of constraints to account for variations in the data.

The data for this study was collected from participants at the National University of Lesotho and traditional doctors in Maseru through one on one interviews in which Sesotho nouns were elicited in order to collect compounds. The primary data from the participants was complemented with secondary data. A variety of Sesotho materials were consulted to collect Sesotho compounds. Such materials included books, theses, dictionaries amongst others.

ACKNOWLEDGEMENTS

I would like to extend my gratitude to both my supervisor and co-supervisor: Dr Kristina Riedel and Prof Nancy Kula for their support and guidance during my study. Words fail me in terms of explaining their patience and understanding throughout this journey. I am indebted to Dr Kristina Riedel for the role she played in my academic life. She went beyond and above the call of duty. She believed in me and saw potential no one has ever seen and for that I will always be grateful.

I am also grateful to a number of academics who gave of their time to have discussions with me regarding my study and were always ready to review my chapters and give feedback. These were: Prof Andrew Nevins, Prof Lee Bickmore and Dr Seunghun Lee. They always found time to discuss my research regardless of their busy schedules and for that I cannot thank them enough. I would also like to highlight that, with all the feedback I got, I am solely responsible for any errors found in this study.

I would also like to acknowledge and extend my gratitude to Ms Pulane Moliko, who is the senior assistant officer at our Department of Linguistics and Language Practice. She has been there for me in all respects and I am deeply grateful.

My appreciation also goes to family and friends for the support they gave me whenever I needed it. I would also like to thank all my informants from the National University of Lesotho in the Department of African Languages.

LIST OF GLOSSES

AGT	Agentive
APP	Applicative
ASSC	Associative marker
COP	Copula
DEM	Demonstrative
DIM	Diminutive
INF	Infinitive
NC	Noun class
NMLZ	Nominaliser
OC	Object concord
PASS	Passive
PFV	Perfective
QP	Qualificative particle
REL	Relative
SM	Subject marker
STAT	Stative

LIST OF ABBREVIATIONS AND SYMBOLS

ABBREVIATIONS:

CON	Constraints set
C1	Constraint 1
C2	Constraint 2
EVAL	Evaluator
GEN	Generator
OT	Optimality Theory

SYMBOLS:

*	Constraint violation
*!	Fatal constraint violation
→	Optimal or winning candidate
∅	Zero prefix

TABLE OF CONTENTS

DECLARATION	iii
ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
LIST OF GLOSSES.....	iv
LIST OF ABBREVIATIONS AND SYMBOLS	v
TABLE OF CONTENTS	vi
CHAPTER 1: INTRODUCTION	1
1.1 Background	1
1.2 Research Questions	2
1.3 Aim and Objectives of the Study	2
1.4 Background Notes on Sesotho.....	2
1.4.1 Nominal morphology.....	4
1.4.2 Phonology of Sesotho nouns.....	10
1.5 Semantics of Sesotho Compounds	16
1.6 Theoretical Framework.....	17
1.6.1 Basic components of OT.....	17
1.6.2 Cophonology	21
1.7 Organisation of the Study	21
CHAPTER 2: LITERATURE REVIEW	22
2.1 Introduction	22
2.2 What is Compounding?.....	22
2.3 Compound Nouns and Headedness	24
2.4 Compounding in Bantu Languages	26
2.5 Compounding in Other Languages	34
2.6 Differences Between Compounds and Phrases.....	35

2.7	Compounding in Optimality Theory	37
2.8	Conclusion.....	42
CHAPTER 3: METHODOLOGY		43
3.1	Introduction	43
3.2	Participants	43
3.4	Data Collection Method.....	44
3.5	Data Processing and Analysis.....	45
3.6	Ethical Considerations.....	46
CHAPTER 4: MORPHOSYNTAX OF SESOTHO COMPOUNDS		47
4.1	Introduction	47
4.2	Types of Sesotho Compounds	47
4.2.1	Root compounds	47
4.2.2	Synthetic compounds.....	58
4.3	Compound Nouns and Headedness	66
4.3.1	Endocentric compounds.....	67
4.3.2	Exocentric compounds	68
4.3.3	Coordinate compounds.....	71
4.4	Differences Between Compounds and Phrases	73
4.5	Conclusion.....	79
CHAPTER 5: OPTIMALITY THEORY ANALYSIS.....		80
5.1	Introduction	80
5.2	Descriptive Generalisations about Sesotho Compounds	80
5.3	Proposed Constraints.....	81
5.3.1	Constraints definitions	81
5.4	Optimality Theoretic Analyses of Sesotho Compound Data	83
5.5	Conclusion.....	99
CHAPTER 6: CONCLUDING REMARKS		100
6.1	Introduction	100

6.2	Concluding remarks	100
6.3	Recommendations for Future Research	102
	REFERENCES	103
	APPENDIX 1: Sesotho Noun + Noun Compound Examples	108
	APPENDIX 2: Sesotho Noun + Adjective Compound Examples	116
	APPENDIX 3: Sesotho Verb + Noun Compound Examples	117
	APPENDIX 4: Sesotho Noun + Adverb Compound Examples.....	118
	APPENDIX 5: Sesotho Noun +Associative marker + Noun Compound Examples.....	119
	<u>APPENDIX 6: Questionnaire for National University of Lesotho Students</u>	119
	APPENDIX 7: Questionnaire for Traditional Doctors	121
	APPENDIX 8: Ethical clearance.....	122

CHAPTER 1: INTRODUCTION

1.1 Background

This study explores the formation of Sesotho compounds from descriptive and theoretical perspectives. Attention is specifically given to the compounds combinatorial patterns, morphological structure, number, agreement, headedness and properties which distinguish compounds from similar constructions, those being their corresponding phrases.

From a descriptive perspective, this study confirms that Sesotho compounds are morphological constructions which morphosyntactically behave like simple nouns. As a result, they are concluded to be different from their corresponding phrases which are syntactic constructions. From a theoretical perspective, the Optimality Theoretic analyses reveal that when forming a compound noun, all constraints that are needed to set requirements for compounds work together to select the optimal form which qualifies as a compound in Sesotho.

Compounding has featured quite extensively in the linguistic literature. It has been studied cross-linguistically for example, in English (Alonso, 2011; Bauer & Renouf, 2001; Fabb, 1998; Katamba, 1993; Matthew, 1974; Spencer, 1991), English and German (Berg, 2012), Greek and Turkish (Ralli, 2013), and Spanish (Rao, 2015). In connection to Bantu languages, compounding has been studied in languages such as: Sepedi (Mphasha, 2006); Setswana (Cole, 1955; Pretorius et al., 2008); Tshivenda (Musehane, 1995, 2004); Cibemba (Kula, 2012) and Mwera (Harries, 1950). Basciano et al. (2011) and Schadeberg (2003) discuss Bantu noun + noun compounds. Downing (2003) discusses compounding, reduplication and tone in Bantu languages.

For Sesotho, compounding has been described in Doke and Mofokeng (1957); Guma (1971); Kalane (2005); Kunene (1961); Lekhotla la Sesotho (1984); Sebotsa (2016) among others. Although, these studies on Sesotho compounding have been undertaken, collectively they concentrate primarily on combinatorial patterns of Sesotho compounds. A detailed study is by Kalane (2005) in an unpublished paper in which the morphological

and semantic properties of Sesotho compounds are discussed. These studies provide an important basis which the present study elaborates on. The present study differs from the abovementioned because of its prime focus on morphosyntactic properties of Sesotho compounds, and the Optimality Theoretic approach it takes to account for such properties.

1.2 Research Questions

This study intends to address the overarching question in (1) in an attempt to determine the morphosyntactic properties of Sesotho compounds. Question (1) is followed by sub-questions (1.1, 1.2, 1.3).

1. What are the morphosyntactic properties of Sesotho compounds?
 - 1.1 What are the combinatorial patterns, morphological structure, number and agreement properties of Sesotho compound nouns?
 - 1.2 Where is the head position of Sesotho compound nouns?
 - 1.3 Are Sesotho compounds morphological or syntactic constructions?

1.3 Aim and Objectives of the Study

The aim of this study is to provide a descriptive and theoretical account of the morphosyntactic properties of Sesotho compounds, and it sets out to achieve the following objectives:

1. To provide the descriptive analyses of the morphosyntactic properties of Sesotho compounds, and
2. To present an Optimality Theoretic account of Sesotho compounds.

1.4 Background Notes on Sesotho

This section presents background information about Sesotho which is relevant for understanding how Sesotho nouns behave. Sesotho (Guthrie's S33) is a Bantu language of the Sotho-Tswana (S30) group, which is spoken in Lesotho and in South Africa, with 13,525,500 speakers (5,625,500 as their first language and 7,900,000 as their second

language) (Eberhard et al., 2020). This study concerns itself with standard Sesotho as it is spoken and written officially in Lesotho. It should be noted that Lesotho orthography is different from the South African orthography. Table 1 is provided below with important examples of representations of sounds used in this study following Lesotho orthography in comparison with the South African one and the IPA. It can be seen from Table 1 that Basotho from Lesotho use the alveolar stop [t] together with the symbol š to represent the aspirated alveolar affricate [tʰ]. In word initial, a syllabic nasal followed by a syllable starting with the same nasal resulting in a geminate, for example *nn* or *mm* is written as ‘*n* or ‘*m* in Sesotho replacing the other nasal with an apostrophe.

Table 1: Difference between Lesotho and South African orthographies

Sesotho (Lesotho)	Sesotho (South Africa)	IPA
kh	kg	[x]
tš	tsh	[tʰ]
ch	tjh	[tʃ]
l	l / d (before high vowels i/u)	[l], [d]
‘n	nn	[n]
‘m	mm	[m]

1.4.1 Nominal morphology

A distinctive system of noun classes is one of the major characteristics of Bantu nouns. As in other Bantu languages, each noun in Sesotho belongs to a particular noun class. The class to which a noun belongs is indicated by a class prefix. There is generally a regular association of singular and plural pairs. This pairing is shown in Table 2 below.

Table 2: Sesotho noun class prefixes (Machobane et al., 2007:2)

Singular				Plural			
Classes	Prefixes	Examples		Classes	Prefixes	Examples	
1	mo-	monna	'man'	2	ba-	banna	'men'
1a	∅	malome	'uncle'	2a	bo-	bo-malome	'uncles'
3	mo-	mose	'dress'	4	me-	mese	'dresses'
5	le-	lerako	'wall'	6	ma-	marako	'walls'
7	se-	selelu	'chin'	8	li-	lilelu	'chins'
9	N-	Nkho	'bucket'	10	li-	liNkho	'buckets'
14	bo-	bolepo	'water moss'	(no plural or class 6 plural)			
15	ho-	ho tsamaea	'to walk'	(no plural)			

Sesotho has 14 noun classes as seen from Table 2 above. It is observed from the table that the class prefixes corresponding to the Proto-Bantu classes 11, 12 and 13 do not occur in Sesotho noun classes. Class 12 comprises diminutives while class 13 consists of augmentatives in the Bantu languages which have retained them. Instead of these classes, Sesotho uses the nominal suffixes: *-nyana*, *-ana* for diminutives and *-hali* to indicate augmentation and femininity. The singular-plural pairs are: 1/2, 1a/2a, 3/4, 5/6, 7/8, 9/10. Nouns in class 14 do not take singular-plural pairing pattern for referring to one and many entities. They can take plural in class 6 indicating collectiveness. An example is *bohobe* 'bread' of class 14 which takes collective plural *mahobe* 'kinds of bread' in class 6. Class 15 comprises infinitives which also do not take plural forms.

In non-derived nouns such as the examples in Table 2 above, one can easily predict the corresponding plural of a singular noun where for example, class 1 takes its plural in class 2. There are exceptions, however, as can be seen with nouns from class 14 which can take their collective plural in class 6, instead of following the singular-plural pairing. Other examples can take two forms of plural depending on whether the meaning is collective or plain plural, for instance *nku* 'sheep' of class 9 can take class 10 plural *linku* 'sheep' which follows the singular-plural pairing pattern and class 6 collective plural *manku* 'kinds of sheep'. This pattern is found in animate nouns only.

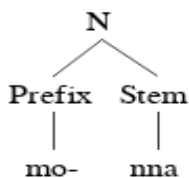
As seen from non-derived noun examples in Table 2, the prefix in singular form is replaced by the prefix in plural form as a differentiation in number, where such a prefix is attached to the noun stem. Another instance is where a plural prefix is attached to the singular noun instead of a noun stem. For example, a class 2a plural prefix *bo-* is attached to a singular noun in some nouns derived from a verb referring to humans. This plural is contextual as it refers to 'one in company of others'. For instance, the noun *molisa* 'cattle herder' derived from the verb stem *lisa* 'herd' of class 1 can take its plural in class 2 as *balisa* 'cattle herders' following the singular-plural pairing where the singular class prefix *mo-* is replaced by the plural class prefix *ba-* like it has been seen with the non-derived nouns in Table 2. The same noun *molisa* depending on context may take class 2a plural and become *bo-molisa*, referring to the 'cattle herder in company of others'. So, class 2a can either be the plural class for class 1a, following the singular-plural pairing or contextual plural for certain nouns derived from verbs. The noun classes which allow this contextual plural are mostly class 1 and class 7 referring to humans.

This contextual plural is also seen when pluralising synthetic compound nouns referring to humans. The compound *moletsa-phala* (one who whistles + whistle) 'referee' of class 1 follows its expected singular-plural pairing pattern and takes the plural form *baletsa-phala* 'referees' in class 2 and another form as *bo-moletsa-phala* 'referee in company of others'. In compounds, this also happens with compounds belonging to classes 1 and 7. Other nouns are non-count examples which identify masses, characteristics, states, among other things. These nouns are referred to as one-class nouns in this study. For example, the noun *mafi* 'sour milk' belongs to class 6 only.

The above discussion gives rise to the conclusion of types of plurality identified. The most widespread or predominantly used is singular-plural pairing pattern of count nouns in which one class takes its plural from the next class. This type of plural is specifically for number distinction indicating one or more entities. Another type of plural is contextual. This kind of plural is found in class 2a as has been seen earlier.

It was seen from the above discussions that a noun is made up of a prefix which is attached to a noun stem and this is its morphological structure. Katamba (2003:103) supports this by saying that a simple noun in Bantu languages can be analysed as consisting of two morphemes being a prefix and a stem. An example of the morphological structure of the noun *monna* 'man' given in Table 2 above is illustrated in (1) below.

(1)



The prefix is an important element in the analysis of Bantu nouns. Katamba (2003:103) perceives the system of noun class prefixes as the hallmark of Bantu nominal morphology. It has been seen that the prefixes in Sesotho reveal the class to which a particular noun belongs and designate number. The prefix determines agreement as well. There is noun class agreement relating to singular-plural distinction, as well as where all pronouns, qualificatives and predicates modifying a noun agree with it in noun class. Examples showing these agreement patterns are given in (2) below.

- (2) a. *Se-fate* *se-se-holo* *sa* *li-perekisi* *se-robehile.*
 NC7-tree QP7-NC7-big ASSC7 NC10-peach SM7-break.STAT.PFV
 'Big tree of peaches is broken.'
- b. *Li-fate* *tse-kholo* *tse* *li-perekisi* *li-robehile.*
 NC8-tree QP8-NC8.big ASSC8 NC10-peach SM8-break.STAT.PFV
 'Big trees of peaches are broken.'

While the class prefixes have been mentioned as important elements in Bantu languages, it has been realised that some nouns can appear without them. This is referred to as class prefix drop and it is common in Sesotho nouns as it is noticeable not only in spoken but also in written Sesotho. It is further identified in the formation of compound nouns. Machobane et al. (2007:6) provide examples of noun class prefixes that are able to shed the class prefix and those that cannot in Table 3 below. Classes 1a and 9 are left out because most nouns in these classes take covert class prefixes whereas, class 15 are infinitives.

Table 3: Sesotho class prefix drop (Machobane et al., 2007:6)

Classes which allow class prefix drop			Classes which do not allow class prefix drop		
Classes	Prefixes	Examples	Classes	Prefixes	Examples
5	le-	lemati 'door'	1	mo-	monna 'man'
7	se-	sehoete 'carrot'	2	ba-	banna 'men'
8	li-	lihoete 'carrots'	2a	bo-	bo-malome 'uncles'
10	li-	liNkho 'buckets'	3	mo-	mose 'dress'
14	bo-	bolepo 'water moss'	4	me-	mese 'dresses'
			6	ma-	mamati 'doors'

Machobane et al. (2007), provide a phonological and syntactic account of class prefix drop in Sesotho nouns. They point out that, phonologically the prefixes starting with the consonants: [l], [s], and [d] with the [+coronal] feature can be lost, while the ones which cannot be lost begin with the consonants: [m] and [b] with the [-coronal] feature. Class 14 is an exception as seen from the table because the prefix starts with [b] yet it allows the class prefix drop, therefore this contradicts the phonological explanation provided since [b] has the feature [-coronal]. There is no answer provided as to why this is the case. Syntactically, Machobane et al. (2007) indicate that agreement, whether nominal or verbal creates an environment for the noun class prefix drop. An example is shown in (3) below, indicating that the sentence is grammatical whether the noun *sefate* 'tree' is used with or without the prefix, arguably because of the verbal agreement.

- (3) (Se)fate se-robehile.
NC7.tree SM7-break.STAT.PFV
'The tree is broken.'

The nouns which have been discussed above were mostly non-derived, and it has been seen that they comprise the prefix and the noun stem as per their morphological structure. Another group of nouns comprises those that are derived from verbs. An example of such nouns, *molisa* 'cattle herder' was discussed earlier with regards to pluralisation. These nouns are agentive and non-agentive denoting nouns. Non-agentive nominalisation involves those deverbatives which do not denote an agent of an action and are mostly non-human nouns which take the non-human suffix *-o*. These deverbatives are abstract and their examples are provided in (4) below.

- (4) a. *thut-o* (*<<-ruta* 'educate')
NC9.educate-NMLZ
'education'
- b. *tseb-o* (*<<-tseba* 'know')
NC9.know-NMLZ
'knowledge'

Deverbatives denoting agentive meaning are made up of action verbs conveying the meaning 'one which verbs' (Comrie and Thompson, 2007). Schadeberg (2003:80) states that primary kinds of agent nouns are derived from verbs by adding the suffix *-i* to the base and that stems are generally assigned to classes 1/2 since common agents are humans. In Sesotho, kinds of agentive nouns are formed by affixing different class prefixes and the personal suffix *-i* to verbal radicals. This is a very productive process in Sesotho as in other Bantu languages (Schadeberg, 2003). The prefix and the suffix are attached to the verbal radical to yield examples such as in (5) below and these nouns belong to their respective noun classes determined by their class prefixes.

- (5) a. *mo-hlahlob-i* (*<<-hlahloba* 'inspect')
 NC1-inspect-AGT
 'inspector'
- b. *se-bapal-i* (*<<-bapala* 'play')
 NC7.play-AGT
 'player'

Guma (1971:44) also mentions another type of agentive nouns, where the class prefix is affixed to the verbal radical to create agentive nouns with the suffix *-a*. Examples of this are illustrated in (6) below.

- (6) a. *mo-lis-a* (*<< -lisa* 'herd')
 NC1-herd-AGT
 'herdsman'
- b. *mo-ren-a* (*<< -rena* 'reign')
 NC1-reign-AGT
 'chief'

Some agentive nouns are created from a radical with a verbal extension before the final suffix in the examples in (7) below, consequently resulting in the following morphological structure; prefix+root+passive+suffix.

- (7) a. *mo-rat-u-a*
 NC1-love-PASS-AGT
 'loved one'
- b. *mo-nyal-u-a*
 NC1-marry-PASS-AGT
 'bride'

There are a lot of nouns derived from verbs in Sesotho (Doke and Mofokeng, (1957); Kunene, 1961). The few discussed here are to act as a necessary basis for the discussion of Sesotho synthetic compounds in Chapter 4, because most of them are made up of nouns of this nature.

1.4.2 Phonology of Sesotho nouns

This section discusses the phonology of Sesotho nouns with special attention given to some phonological sound changes which take place when particular deverbatives are derived and nominal tone. Sub-section 1.4.2.1 discusses the phonological sound changes and nominal tone is discussed in sub-section 1.4.2.2.

1.4.2.1 Phonological sound changes in the formation of deverbatives

As has been mentioned, deverbatives are derived from verbs and during the derivation process some sounds incur phonological changes. The following environments: deletion, assimilation and strengthening have been identified to allow such sound changes. In examples in (8) below, the processes of deletion and assimilation take place. The vowel of class 1 prefix *mo-* or the object concord *mo-* of class 1 is deleted and then the consonant [b] is fused, to make it a geminate, together with affixation of the agentive suffix *-i* in example (8a) and the *-e* for imperative form in (8b).

Example (8c) is a compound noun which has the first constituent as a deverbative which is derived the same as in example (8a) where the class 1 prefix *mo-* has its vowel deleted and then the consonant [b] is fused, resulting in the geminate. The difference between examples (8a and 8c) is that example (8c) is affixed with the agentive suffix *-a* not *-i*. It was seen in earlier discussions that some deverbatives can take the suffix *-a* and this is also the case with the first agent-denoting constituent of synthetic compounds.

- (8) a. [b] > [m]
mo-bus-a >> mbusi >> mmusi >> 'musi'¹
 NC1-govern-AGT 'governor' (agentive)

¹ Refer to section 1.4 on how orthography is handled.

- b. *mo-bits-a* >>mbitse >> mmitse >> 'mitse!
 OC1-call-AGT 'call him!' (imperative)
- c. 'mak-a *li-taba* >>mo-bak-a >>mbaka >> mmaka >>'maka
 NC1a.cause-AGT NC10-affair
 'troublemaker'

There is also an instance of a phonological influence of an underlying N- nasal class 9 prefix on certain verbal stems in derivation of class 9 deverbatives which results in the phonological process of strengthening. The sound changes are shown in the first constituents of compound examples in (9) below and they are: s> ts^h, t > t^h, and j> tj.

- (9) a. *tšal-o*² *morao* (<<-sala 'follow')
 NC9.follow-NMLZ after
 'follow up'
- b. *chebel-o* *pele* (<<-sheba 'see')
 NC9.see-NMLZ forward
 'vision'
- c. *tłhoph-o* *bo-cha* (<<-hłlopha 'structure')
 NC9.structure-NMLZ NC14-new
 'restructuring'

More compound examples with sound changes which take place under the influence of the underlying N- nasal class 9 prefix are provided in (10) below, where it can be seen that: voiced stops are devoiced (b>p) as shown in example (10a), voiced continuants become stops and devoice (l>t) as illustrated in (10c) and voiceless continuants change to homorganic aspirated affricates (s> ts^h) as seen in (9a), and a vowel takes the initial velar stop /k/ as in (10b). It should be noted that /k/ is the epenthetic consonant in

² Again refer to section 1.4 on how orthography is handled

Sesotho. According to Rose and Demuth (2006), this epenthetic consonant is used in vowel-initial stem contexts where there is an underlying nasal. For example, in nominalisation process where *otla* 'beat' becomes *kotlo* 'beating' in (10b).

- (10) a. *pul-a* *ma-liboho* (<<-*bula* 'open')
- NC9.open-AGT NC6-border
- 'pioneer'
- b. *kotl-o* *qobell-o* (<<-*otla* 'beat')
- NC9.beat-NMLZ NC9-force-NMLZ
- 'sanction'
- c. *tik-a* *mo-tse* (<<-*lika* 'surround')
- NC9.surround-AGT NC9-village
- 'plant used as Sesotho medicine'

1.4.2.2 Nominal tone

Sesotho is a tonal language in which tone is used to convey lexical and grammatical meaning (Raborife, 2011). The nouns are no exception to this since they are marked tonally. Sesotho has two basic tones: high (H) and low (L). Low tone has variations such as mid (M) (Kunene, 1961). In this discussion, a high tone is indicated with an acute accent, low tone with a grave accent and mid tone with two dots. Tone can be used to differentiate meaning between two identical words. For example, *lèhátà* 'skull' has a LHL pattern while *lèhàtà* 'liar' has a LLL.

It is realised from these examples and others from the literature that the noun class prefixes in Sesotho bear a low tone and that the tone spreads from left to right. It is, however, important to note that not all nouns composed of identical phonemes can be distinguished by tone as some are also identical in tone. For such words, only the context in which they are used clarifies the meaning. The following in (11) below, from Doke and Mofokeng (1957, 38) are examples of such words in which both words in (11a) have LLL pattern and those in (11b) have LL.

- (11) a. *li-kòtò*
 NC10-knobkerrie/piece
 'knobkerries or pieces'
- b. *hlòlò*
 NC9.hare/creation
 'hare or creation'

Tone may also be used to show grammatical relationships. For instance, the distinction between a copula without a subject and first person singular subject marker can be judged by means of tone. The following examples in (12) below from Doke and Mofokeng (1957:39) illustrates this grammatical relationship.

- (12) a. *ké tichere*
 COP NC9.teacher
 'He/she is a teacher'
- b. *kè tichere*
 SM1sg NC9.teacher
 'I am a teacher'

Kunene (1961) provides examples of Sesotho nouns and marks their tone in a study on the sound system of Sesotho. An example of a derived noun which is agentive denoting is given in (13) below, bearing LHHM tone pattern.

- (13) *mò-sébéts-ì*
 NC1.work-AGT
 'worker' [Kunene, 1961:232]

Kunene (1961) points out that, when the suffix or the final vowel of the derived noun stem is *-i*, the tone pattern of the stem is the same as that of the verb stem from which it is derived. That is, the verb stem *sebetsa* 'work' takes the tonal pattern; HHM. Some words which are syntactically related to the noun, and which appear immediately before or after

it affect such a noun tonally. Kunene (1961: xiii) refers to such words as influencing elements and they can be seen in example (14) below. It should be noted that this data is not cited from sources. The tone patterns were marked by the researcher.

- (14) a. *lè-tónà*
 NC5-minister
 ‘minister’
- b. *lè-tóná* *lènà*
 NC5-minister DEM5
 ‘minister this’
- c. *léná* *lè-tónà*
 DEM5 NC5-minister
 ‘this minister’
- d. *lè-tóná* ⁺*lá* ⁺*thútò*
 NC5-minister ASSC5 NC5.education
 ‘minister of education’

From the examples in (14), it can be seen that there is a high tone spread where the H spreads from the penult to final syllable in the noun *letona* in example (14b), in which LHL changes to LHH when a demonstrative *lena* follows the same noun *letona*. In the case of example (14c), when the demonstrative *lena* appears before the same noun *letona*, there is no tone spread. In (14d), there is a high tone spread, where the H spreads from the penult to final in the noun, then from the final to the associative *la* causing two downsteps followed by L.

With respect to tonal patterns found in Sesotho compounds, Kunene (1961:267) provides the examples cited in (15a, b and c) and claims that the constituent elements of some Sesotho compounds have the same tone patterns in combination and in isolation.

Although this study was conducted many years ago, the examples (15a, b and c) are of interest as it is observed that there is indeed a case where elements of a compound have the same tonal patterns in combination and in isolation, in comparison with tonal patterns marked recently by the researcher in (15d, e and f).

- (15) a. *lè-tè-à* *tšoènè*
 NC5-strike-AGT NC9.baboon
 ‘warrior’
- b. *tè-à*
 strike-FV
 ‘strike’
- c. *tšoènè*
 NC9.baboon
 ‘baboon’
- d. *mò-làlà* *mò-tšò*
 NC3-neck NC3-black
 ‘turtle dove’
- e. *mòlàlà*
 NC3-neck
 ‘neck’
- f. *mò-tšò*
 NC3-black
 ‘black’

The compound *letea-tšoene* ‘warrior’ in (15a) has the same tonal patterns with its individual elements *tea* in (b) and *tšoene* in (c). This can also be seen from an example

of a noun + adjective compound in (15d) in which both elements in combination and in isolation in (15e and f) hold the same tonal patterns.

1.5 Semantics of Sesotho Compounds

Semantically, nouns can be categorised based on their meaning, for example nouns belonging to semantic areas such as: humans, animals and so on. Most examples of Sesotho compounds in the data collected are found in the following semantic areas in (16) below, nevertheless, others are found in other areas but are not as prevalent as those mentioned below.

(16). a. Types of humans (mostly synthetic compounds)

<i>'maka-litaba</i>	1.cause.AGT -10.affair	'troublemaker'
<i>seila-tsatsi</i>	7.avoid.AGT -5.sun	'a beautiful woman'
<i>leapara-nkoe</i>	5.wear.AGT -9.leopard	'warrior'
<i>sebala-makhulo</i>	7.count.AGT -6.pasture	'vagabond'

b. Animals

<i>kolobe-moru</i>	9.pig-3.forest	'wild pig'
<i>sethoena-moru</i>	7.hop.AGT-3.forest	'cape robin-chat (bird)'
<i>molisa-lipela</i>	3.herd.AGT-10.rock jumper	'Drakensberg rock jumper'
<i>khoho-noka</i>	9.chicken-9.river	'common moorhen'
<i>tsie-balimo</i>	9.locust-2.ancestor	'type of locust'

c. Plants

<i>molika-letšana</i>	3.surround.AGT-5.lake.DIM	'Scleria Dieterleni'
<i>mohlaba-lerumo</i>	3.pierce.AGT-5.spear	'bedding grass'
<i>sehoete-moru</i>	7.carrot-3.forest	'type of edible wild roots'
<i>sehamela-poli</i>	7.milk.APP.AGT-9.goat	'kind of leaves; traditional medicine'

d.	Place names		
	<i>thupa-kubu</i>	9.stick-9.hippopotamus	'name of a place'
	<i>ngope-tšoeu</i>	5.donga-5.white	'name of a place'
	<i>khohlo-ntšo</i>	9.gorge-9.black	'name of a place'
	<i>motse-mocha</i>	3.village-3.new	'name of a place'
e.	Academic terminology		
	<i>sebopeo-puo</i>	7.structure.NMLZ-9.language	'grammar'
	<i>kutu-ketso</i>	9.root-9.action.NMLZ	'verb'
	<i>sebapali-moelelo</i>	7.play.AGT-3.meaning.NMLZ	'thematic role'

1.6 Theoretical Framework

This section provides the theoretic basis of which this study undertakes. This study adopts a unified account of standard Optimality Theory and Cophonology which is a development within Optimality Theory to account for morphosyntactic properties of Sesotho compounds. Original or standard Optimality Theory (OT) was introduced by Prince and Smolensky (1993/2004) and further advanced by McCarthy (1993, 2002, 2003a, 2003b); McCarthy & Prince (1993a, 1993b). This theory was primarily used in phonology and later its usage spread to other domains such as morphology, syntax, semantics, amongst others. OT is a constraint based representational model of grammar.

The central thesis of OT is that the interaction between conflicting constraints results in observed forms of a language. OT, sees grammars as systems providing mappings from underlying representations to surface representations.

1.6.1 Basic components of OT

The first component is CON (the constraint set). Universal Grammar consists of a set of principles called constraints which demonstrate universal linguistic tendencies in languages. These constraints are described according to the following characteristics.

i. Universality

In OT, constraints are described as universal, but different languages are accounted for by differences in constraint rankings. That is, the basic insight of OT is that even if grammars are driven by the same principles, these principles are ranked in different ways in different languages. For example, this is seen from the theoretic analyses of the present study which reveals that compound nouns comprise different patterns which cannot be explained purely within standard OT. Which then leads to manipulation of cophonologies discussed in section 1.6.2 which results in re-ranking of constraints to account for variation as opposed to ranking of constraints in a strict hierarchy of domination which is the basic principle of OT.

ii. Ranking of constraints

For individual grammars to resolve conflicts there must be ranking of the constraints. A particular constraint A may be ranked very high in some languages, so that the surface forms always or nearly always satisfy it.

The same constraint A may be ranked lower in another language. In this case it would be of less importance in this language. In this language, the constraint A can be crucially dominated by for instance, constraint B which conflicts with it so that surface forms satisfy B and violate A. Meaning, constraints A and B conflict with each other and the conflict is resolved differently in one language and in another. This kind of conflict is normally represented by means of the tableaux which will be revealed in later discussions (Kager, 1999).

iii. Violability

A central property of constraints is that they are violable. In principle, all constraints are violable even when they are ranked highest in a hierarchy. A candidate which comes out as optimal may violate at least some constraints but it must incur the least violations of the ranked constraints. This brings rise to the fundamental claim in OT that no linguistic form manages to satisfy all constraints' requirements, meaning that constraints do not

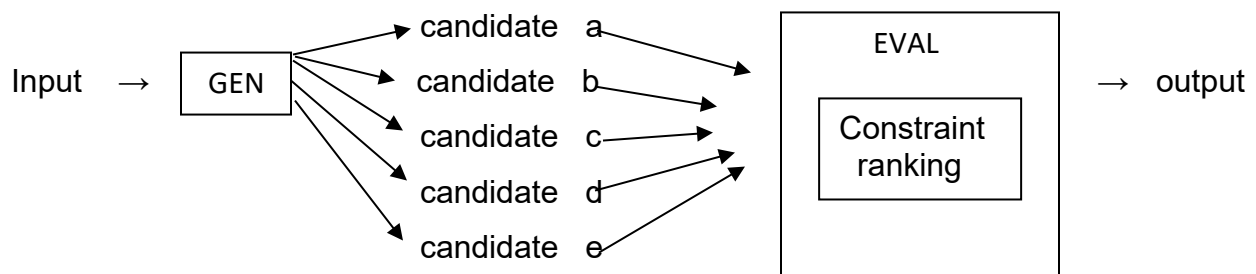
need to be satisfied in all contexts. They are also violable in relation to the needs of the grammar in production of surface forms (McCarthy, 2008).

iv. In conflict

Constraints are essentially in conflict. That is, a possible surface form in any language will necessarily violate at least a certain constraint. As a result, grammars should be able to regulate conflicts between universal constraints in identifying the most acceptable surface form (Kager, 1999).

The above-mentioned constraints are represented by two types, being the markedness and faithfulness constraints. The markedness constraints on one hand enforce well-formedness of the output, meaning that they do not allow input that is marked and the faithfulness constraints on the other hand, enforce similarity between input and output. There is also an addition of the third type; namely the alignment constraints which are relevant to mention as the analyses of this study involves them. These alignment constraints are responsible for proper concatenation of morphemes.

The second component is GEN (generator). GEN is responsible for generating a list of output forms from the input. The list of these output forms is referred to as a candidate set. OT claims that all the output forms produced are in conflict with each other in order to win the competition and as a result, be evaluated as optimal (Fanselow & Féry, 2002). The third component is the EVAL (evaluator). After GEN has generated a set of candidates, EVAL then applies the ranked constraints to the set of candidates provided and evaluates which is the winning candidate. The above discussion can be represented by an OT architecture adopted from Nishimura (2013:16) below.



In summary, it is observed from the OT architecture that when GEN is provided with an input, it generates a candidate set from that input and then EVAL using the constraint hierarchy of the language chooses the winner from the candidate set. An interpretation of the above schema through a tableau is provided in Tableau (1) taken from McCarthy (2002:4).

Tableau 1: A ranking argument

	C1	C2
a. → Cand _{opt}		*
b. Cand _{comp}	*	

As can be seen, candidate comparison is shown in a *tableau*, where the winner is compared with other candidates with regards to how they behave towards the constraint hierarchy. In summary, a tableau provides information on the significance of constraints as well as their rankings in selecting an optimal form. The constraints as can be seen from Tableau 1 are listed according to their importance where their domination order is from the left to the right. The rows comprise competing candidates which the winner is chosen from.

The violation marks incurred by any candidate violating a particular constraint are shown by an asterisk (*). Less serious violations are shown by the asterisk alone as seen from the tableau, while fatal or serious violations are indicated by the asterisk followed by an exclamation mark, this case will be seen in chapter 4 as the above tableau does not demonstrate it. The optimal form or the winner is shown by the arrow appearing before it. From the tableau, the constraints C1 and C2 are in conflict relative to their evaluation of two candidates provided. C1 favours Cand_{opt}, while C2 prefers the competitor Cand_{comp}. Considering Cand_{opt} as the target form, C1 has to be ranked above C2. Meaning, C1 dominates C2 in the hierarchy, which is written as [C1 » C2].

1.6.2 Cophonology

Cophonology (constraint re-ranking) is a development within standard OT which is used to account for internal language variation and exceptionality which OT following its emphasis on ranking of constraints in a strict hierarchy of domination does not sufficiently account for. According to Cophonology, there are two Cophonologies corresponding to each input. For instance, there is a number of phonological patterns related to certain morphological constructions which provide an environment where constraints may show a particular ranking for a certain morphological construction and a reverse ranking in another (Inkelas and Zoll, 2007; Orgun, 1996). For example, this is seen in accounting for particular counterexamples of compound nouns in Sesotho.

1.7 Organisation of the Study

This study comprises six chapters. Chapter 1 is this introductory chapter. Chapter 2 examines compounding phenomena. Chapter 3 outlines the methodology of this study. Chapter 4 provides a discussion on morphosyntactic properties of Sesotho compounds. Chapter 5 presents an Optimality Theoretic account of Sesotho compounds. Chapter 6 outlines the concluding remarks.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter presents a review of literature in order to provide needed background to this study. The core issues given attention are compounding as well as theoretic aspects which are relevant to this study. Another important aspect of the literature is background on Sesotho which is discussed in chapter 1, section 1.4. The chapter comprises eight sections. Section 2.1 is this introductory section. Section 2.2 explains what compounding is. Section 2.3 deals with the headedness phenomenon in compounding. Section 2.4 discusses compounding in Bantu languages. Section 2.5 is on compounding in other languages. Section 2.6 provides a distinction between compounds and phrases. Section 2.7 discusses compounding in Optimality Theory followed by section 2.8 on conclusions.

2.2 What is Compounding?

Compounding is a very productive way of adding new words to different languages and it has proven to be a topic of interest cross-linguistically. Plag (2003:169) claims that compounding is the most controversial phenomena in relation to its linguistic analysis. The subject that arouses debates in many studies is what a compound is. Lieber and Štekauer (2009:2) mention that extensive discussions from the literature are relative to what a compound is and further point out that there is difficulty for scholars to come up with universally accepted definitions of a compound. Consequently, different scholars offer different definitions and give attention to various aspects by making use of different terms in order to highlight for instance, certain crucial characteristics shared by most compounds (Nakov, 2013).

According to Plag (2003:170), compounding involves combination of two words to form a new word. This definition points to two conclusions; first, that compounds are specifically made up of only two words while the second points to the fact that the elements which make up a compound are words. The first conclusion therefore leaves out compounds with more than two words. An example of such a compound in English is provided as; *power source requirement*. As a result, this definition which indicates that compounding

involves only two words is found to be confining. So, Plag (2003:173) modifies the previously mentioned definition and says, “a compound is 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.”

This definition is still restrictive regarding the number of elements which form a compound because it still mentions that a compound is made up of two elements but it accommodates the fact that words are not the only elements used in the formation of compound nouns. There are instances of compounds in which the first member is not made up of an independent word such as those compounds which have the first element as *bio* like an English compound - *biochemistry*. The element *bio* is referred to as a root, a bound root because it behaves more like a word not an affix. Additionally, other compounds involve the first element as a phrase like in *over-the-fence gossip*.

Katamba (1993:54) on one hand, explains compounds as comprising at least two bases which can be words or root morphemes. Katamba's (1993) use of the expression 'at least' shows consideration that some compounds may not consist of two words only and also acknowledges that not only words but also roots can be used in the formation of compounds. Fabb (1998:66) on the other hand, defines a compound as a word consisting of two or more words. So, Fabb (1998) is of the view that only words can be used to create compounds and also considers the fact that a compound may consist of two or more words.

Two types of compounds are distinguished in the study of compounding cross-linguistically. In a discussion by Spencer (1991) there are two types of compounds identified, namely root and synthetic. In distinguishing between the two compound types, Lieber (2005:375) states that synthetic compounds are those compounds which have the second constituent derived from a verb, while root compounds are those in which the second constituent is non-derived. An English synthetic compound example provided is *truck driver* which has the second noun *driver* derived from the verb 'drive'. The English root compound example is given as *file cabinet* in which the second noun *cabinet* is not derived.

Various compounding patterns are attested cross-linguistically in the formation of these root and synthetic compounds. Words from different word classes are combined to create these compound types. These compounding patterns are an instance of classification of compounds according to composition where examples such as: noun + noun which is the most attested pattern from the literature, noun + adjective, noun + adverb, verb + noun amongst others are found. Noun + noun compounds are by far the most productive in English (Lieber, 2005). This is different in Bantu languages as Basciano et al. (2011) state that Bantu languages generally have a limited number of noun + noun compounds which are mostly kinship terms and phytonyms. Additionally, some languages have more complex structures like *over-the-fence gossip* which was discussed earlier. Other complex structures in some languages comprise two nominals which are linked by a linking element or an associative marker producing the pattern; noun + associative marker + noun.

2.3 Compound Nouns and Headedness

The notion of headedness has also been extensively debated in the literature. This topic mainly centres on the criteria used in determining the head of a compound. The head is defined as a word in a syntactic construction for example, in a phrase or a morpheme in a morphological construction like in a compound noun which determines the meaning of the construction as a whole (Aronoff and Fudeman 2011).

Different languages identify different positions of the head in compound nouns. Compound nouns in Bantu are left-headed (Basciano et al., 2011). Languages like Dutch and German are shown to be strongly right-headed, while others such as Arabic and French are strongly left-headed, whereas English is perceived to cut the boundary between left-headedness and right-headedness (Altakhaineh, 2018).

Different types of heads are attested in the literature relating to whether the head can be semantic, syntactic, morphological or morphosyntactic (Bauer, 2009; Fabregas & Masini, 2015; Hoeksema, 1992; Lieber, 2010; Scalise & Fabregas, 2010). Attention in this study is drawn to the semantic head. The semantic analysis of the position of the head in

compounding produces three types of compound nouns being; endocentric, exocentric and coordinate compounds.

Endocentric compounds according to Nakov, (2013:9) are a type of compounds where one element is the head while the other acts as its modifier, that is, attributing a property to the head. This is interpreted as 'AB is (a) B' in English whereas it is the opposite; 'AB is (a) A' in other languages such as Bantu. Nakov further provides an English compound *steamboat* pointing out that, as compared with *boat*, *steamboat* is inherently found in the same semantic contexts as the noun *boat*. Steamboat also takes the category noun of *boat*, since they are both nouns. As a result, a *steamboat* is a type of *boat*, where the class of *steamboats* is a subclass of the class of *boats*.

Endocentric compounds in different languages from the literature appear to take the left-headed, and right-headed structures. In English for example it was mentioned that there exist left and right-headed compounds but right-headed ones are predominant. An example of an endocentric right-headed compound noun in English is *handbag* which has the head on the right and an example of a left-headed endocentric compound is *darkroom* which has the head on the left. This is established as well in this study to see if both structures are found in Sesotho compounds.

With regards to exocentric compounds, Nakov (2013:9) states that these compounds lack a semantic head. That is, the semantics of the compound does not derive from the combined elements but an inference of meaning outside the combined words. Lieber, (2005:378) provides an English exocentric compound; *red head* which is not a type of head, but a metonym for a type of a person. In other words, a person with red hair. There is also a debate which arises from the topic relating to the existence of exocentric compounds in the studies of compounding. Other researchers raise concerns about whether exocentric compounds exist. Among other issues discussed is the metaphorical inferences of the semantics of exocentric compounds which produce their endocentric interpretation.

As for coordinate compounds, these are endocentric compounds with two equal heads. Dressler (2005:275) provides a coordinate compound example in English as *speaker-*

hearer and states that both constituents of this compound have equal status. The relation between the constituents of a compound of this type is that of coordination. Meaning, a referent is characterized by both elements of the compound. The constituents of a coordinate compound are marked with dual inflection as a further indication of their equal status (Khan, 2013).

2.4 Compounding in Bantu Languages

Compounding in Bantu languages, including Sesotho has been studied as has been mentioned in chapter 1. The studies on Sesotho compounding except for Kalane (2005) and Kunene (1961) only describe compound types according to their composition, where their possible combinations in terms of the lexical categories in which their constituents belong to are discussed. Kalane's (2005) work is discussed later. Kunene discusses tonal patterns of Sesotho compounds and this work has been discussed in chapter 1, section 1.4.2.2. Doke and Mofokeng, (1957:104) state that the most commonly found compounds in Sesotho may be classified according to composition as follows: noun + noun, noun + adjective, noun + possessive, verb + subject, verb + object, verb + adverb among others. The compound examples in (1) have been identified from Doke and Mofokeng and they all take the category noun. Sesotho has a large number of these compounds.

(1)	Noun + noun	<i>kolobe-moru</i>	(bird + forest)	'wild pig'
	Noun + adjective	<i>ntate-moholo</i>	(father + old)	'grandfather'
	N + possessive	<i>boea-ba-ntja</i>	(dog's hair)	'plant'
	Verb + logical subject	<i>boea-batho</i>	(lit. where people go)	'death'
	Verb + object	<i>bohlaba-phieo</i>	(state of piercing + kidney)	'traitorship'
	Verb + Adverb	<i>moeta-pele</i>	(leader + front)	'leader'

In Sebotsa (2016:62-64), the following combinatorial patterns in (4) are identified: noun + noun, noun + adjective, noun + adverb, noun + verb, verb + noun, verb + adjective, verb + noun as well as verb + adverb and they all take the category noun except for the verb + noun and verb + adverb examples which take the category verb. This still emphasises the fact that most compounding patterns in Sesotho produce nouns.

(2)	[N + N] N	<i>thaba-lesoba</i>	(mountain + hole through a thing)	'place name'
	[N + A] N	<i>nko-phara</i>	(nose + flat)	'type of person'
	[N + Adv] N	<i>motho-feela</i>	(person + nothing)	'ordinary person'
	[N + V] N	<i>mosika-phalla</i>	(that which goes parallel + run to)	'centipede'
	[V + N] N	<i>thala-boliba</i>	(frolic + deep hole in the river)	'whirligig insect'
	[V + A] N	<i>khofa-ntšo</i>	(husk of a grain + black)	'referent not known'
	[V + N] V	<i>hata-koekoe</i>	(step + koekoeken)	'walk carefully and slowly'
	[V + Adv] V	<i>bolela-teng</i>	(mention + there)	'not known'

Other compound examples are provided in (3a, b and c); in which two nouns are combined (Guma, 1971). Guma (1971:39) further gives other examples, in which a class prefix is added to a radical to form a derived noun that is combined with another noun, which produce the compound examples in (3 d, i and ii). These are synthetic compounds.

- | | | | | |
|-----|----|-----------------------|---|-----------------------|
| (3) | a. | <i>mokopu-ntja</i> | (pumpkin + dog) | 'warthog' |
| | b. | <i>fito-khomo</i> | (knot + cow) | 'well-made knot' |
| | c. | <i>khoho-moru</i> | (chicken + forest) | 'a kind of wild fowl' |
| | d. | i) | Where second noun is the logical subject of the verbal idea: | |
| | | <i>boea-batho</i> | (ea 'go' + batho 'people') | 'death' |
| | | <i>bochaba-tsatsi</i> | (chaba 'rise' + letsatsi 'sun') | 'east' |
| | | ii) | Where the second noun is the logical object of the verbal idea: | |
| | | <i>sethiba-pitsa</i> | (thiba 'stop' + pitsa 'pot') | 'pot-mender' |
| | | <i>maapara-nkoe</i> | (apara 'wear' + nkoe 'leopard') | 'warriors' |

In addition to the above compound examples, compounds with the complex structure: noun + associative marker + noun are identified in a discussion by Guma (1971) and examples are given in (4) below.

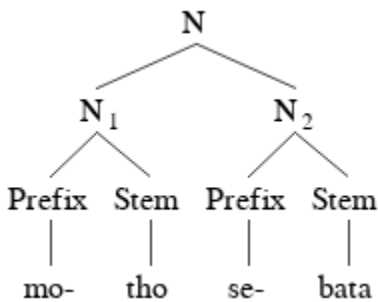
- | | | | |
|-----|----------------------------|--------------------------|----------------------|
| (4) | <i>koloi-ea-mollo</i> | (vehicle + of + fire) | 'train' |
| | <i>mala-a-libe</i> | (intestines + of + sins) | 'gonorrhoea' |
| | <i>setsoalle-sa-lintja</i> | (friendship + of + dogs) | 'useless friendship' |

Kalane (2005) focuses on morphological and semantic analyses of Sesotho compounds. The present study elaborates on this work. While Kalane (2005) discusses the morphological properties of Sesotho compounds, there are gaps with regards to discussions on the role played by class prefixes, which are crucial to Bantu nouns. This relates to class prefix drop as well as agreement patterns. Additionally, the issue of headedness is not addressed in detail. Relating to the morphological properties of compounds, Kalane (2005) discusses compounds' morphological structure where constituents of compounds are analysed as made up of free and bound morphemes, as in the examples in (5) below.

- (5) a. Free + free → compound (lit.) → 'other meaning'
thaba + phatšoa (black and white mountain) 'name of a mountain'
phatla + nchocho (a sharp forehead) 'kind of forehead'
- b. Free + bound + bound + bound + bound
thab+ana+n+tle+nyana (least beautiful mountain) 'name of a mountain'
 mountain-diminutive-adj concord-adj stem-diminutive
- c. Bound + bound + bound + free
ma+sis+a+pelo (heartbreak) 'tragedy'
 nominal prefix-verb stem-verbal ending-heart

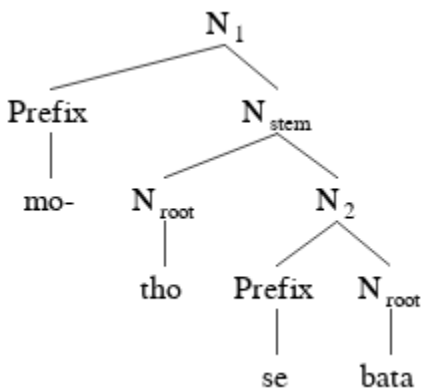
Apart from discussing the morphological structure of compound nouns on the basis of free and bound morphemes, Kalane (2005) provides a tree diagram to account for the morphological structure of a compound noun. The morphological structure that is followed by Kalane (2005:5) shows that, for example noun + noun compounds are made up of two nouns with each comprising a class prefix and a stem. This is illustrated in (6) below.

(6)



Mphasha (2006) when discussing Sepedi compound nouns provides a different illustration of the morphological structure of noun + noun compounds, adopting Musehane's (1995:113) morphological structure of Tshivenda compounds. Mphasha mentions that the prefix of the first member is the head of the whole compound which is contrary to Kalane's (2005) structure which treats the first noun as the head. The following in (7) below is the structure of a Sepedi compound noun.

(7)



When interpreting the above structure, Mphasha (2006) states that, the compound noun *motho-sebata* (person + beast) 'carnivorous man' has two class prefixes as seen in (7) and further indicates that all noun + noun compounds have the class prefix as the head which, therefore, carries properties of the whole compound noun.

Mphasha's (2006) morphological structure is not accepted in this study in that, if the compound is said to have two class prefixes then it is made up of two nouns with each

comprising a class prefix and a stem, which contradicts Mphasha's (2006) morphological structure which suggests that the compound noun comprises the prefix and the noun stem, and that this noun stem is made up of the noun root and another noun which consists of the prefix and the stem. Somehow, the prefix of the second noun as shown by this structure is absorbed within the noun stem of the first noun.

Mphasha's (2006) morphological structure is therefore found to raise questions as to the formation of a compound noun, which clearly indicates that a compound may be formed from combining two words or more, which in this case should be the combination of two nouns, with each comprising the class prefix and the stem.

As a result, the prefix is not treated as the head of the compound in this study, but as a part of the head following Kalane's morphological structure. So, the interpretation should be that the head noun is important as it is the determinant of morphosyntactic features.

Semantically, Kalane (2005) identifies the head of a compound as important in the interpretation of compound nouns and discusses compound examples with and without heads. Kalane (2005:6) provides the headed Sesotho compound, *kolobe-moru* 'wild pig' which is analysed as a kind of pig and the non-headed compound provided is *poho-tšehla* which literally means 'yellow bull' which is not a kind of bull colour but Sesotho medicine. Kalane (2005) also points out that the characteristic of the head of a compound is that it can carry inflectional affixes and inflect Sesotho compounds for plurality as in example (8) below, and additionally, points out that the prefixes are added to headwords to show the plural forms of compounds by replacing the singular prefix with the plural one.

- (8) *kolobe-moru* > *likolobe-moru* 'wild pigs'
tsie-balimo > *litsie-balimo* 'type of locust'
ntja-me > *bo-ntja-me* 'brothers or sisters'

Mphasha (2006) discusses compound nouns in Sepedi as mentioned earlier. This study explores among other issues, the morphology and semantic properties of Sepedi compound nouns. It reveals many properties which are the same as in Sesotho, and this is expected since the two languages are very closely related. Focusing on noun + noun

- (12) *seno-tagi sa gago*.
'the alcoholic drink of yours.'

Mphasha (2006) also categorises compound examples as belonging to different semantic areas such as: natural phenomena, humans, animals, plants, natural objects, body parts, cultures, cognition, direction, food, artefacts and states.

Another study on compounding in Bantu languages is by Pretorius et al. (2008) on Setswana. The formation of noun + noun compounds in Setswana is investigated in this study. In a discussion about endocentric groups, Setswana examples are identified as left-headed. An example provided is that of a compound noun *modula-setulo* (sitter + chair) 'chairperson' where *modula* refers to 'sitter', therefore is regarded as the head element.

Pretorius et al. (2008) point out that, some Setswana compounds originate from the underlying word group structure. For instance: a compound is formed from a word + a word group where the processes of reduction, replacement and elision of morphemes take place. An example of a possessive group is discussed and it is stated that in this case, the first noun is an antecedent and it is qualified by a possessive group as in *ditsetela tsa nko* 'roads of the nose' and then the compound *litsetela-nko* 'air passages' is created. Another example is from the infinitive group as in *go dula setulo* 'to sit on the chair' where the infinitive verb is nominalised to the preferred noun class and the compound *modula-setulo* 'chairperson' is formed.

Other compounds discussed in this study are those which have developed from a simple juxtaposition. These are compounds indicating place names or locations. Such examples are given in (13) below.

- | | | | |
|------|-------------------------|------------------|----------------|
| (13) | <i>Aferika-borwa</i> | (Africa + South) | 'South Africa' |
| | <i>Bokone-bophirima</i> | (North + West) | 'North-West' |
| | <i>Bokone-botlhaba</i> | (North + East) | 'North-East' |

Tshivenda is another Bantu language in which compound nouns are discussed by Musehane (2004). This study investigates the structure of root compound nouns. It is

pointed out that there are nominal compounds in Tshivenda, but no verbal compounds like in Sotho-Tswana group. Compounds discussed here are of two types; those created from combining two nouns and those in which any constituent is made up of a root. These are given in (14) below. The compound in (14a) is made up of two nouns. In (14b) the root of the first noun adopts a new prefix, while the second noun retains its prefix. The new class prefix is *tshi-* which replaces *mi-*. The first constituent is however not seen as a root as Musehane (2004) states because it still appears as a noun comprising the prefix and the stem before and after it has adopted the new prefix. This is therefore still an example of a noun + noun compound, not root + noun.

In the compound in (14c), the first constituent keeps its class prefix and then the second member loses the prefix, thus appearing as a root. This is an instance of a class prefix drop which was discussed in Chapter 1, Section 1.4.1. The same goes for the compound in (14d) where the first member adopts a new class prefix and therefore is not a root but a noun and then the second element sheds its class prefix, thus appearing as a root.

(14) a. Noun + Noun

ḷiivha *thavha*
 NC5.pigeon NC9.mountain
 'rock pigeon'

b. Root + noun

tshiḷodzimare >*miḷodzi*, 'tears' of class 4 + *mare* 'saliva' of class 6.
 'weeping'

c. Noun + root

nungupfa >*nungu*, 'porcupine' of class 9 + *-pfa* <*mupfa* 'thorn' of class 3
 'porcupine'

d. Root + root

musimbiri >*tsimbi* 'iron' of class 9 + *muri* 'tree' of class 3.
 'iron wood tree'

Lusekelo (2019) discusses Swahili compounds and points out that compound nouns in Swahili are made up of class prefixes and roots. The Swahili compound nouns like in other Bantu languages bear the grammatical feature number, which is hosted by the head noun where the singular prefix is replaced by the plural prefix. This was seen with Sesotho examples discussed earlier as well. The number is expressed by the morphosyntactic process of prefixation. Lusekelo (2019:25) provides the Swahili singular compound in (15) and its plural form in (16).

(15) *∅-garimoshi ∅-kuu li-me-pinduka*
 NC5-train NC5-old SM5-PFV-overturn
 ‘an old train has overturned.’

(16) *ma-garimoshi ma-kuu ma-wili ya-me-pinduka*
 NC6-train NC6-old NC6-two SM6-PFV-overturn
 ‘two old trains have overturned.’

In example (15) above the noun + noun compound *garimoshi* ‘train’ takes a zero class marker and the adjective *kuu* ‘old’ too. The noun class 5 is determined by verbal agreement. The plural form *magarimoshi* ‘trains’ surfaces with an overt class prefix *ma-* (class 6). In the same vein, the head noun controls agreement as this is indicated by the modifier *makuu* ‘old’ in (16). So, in both examples (15 and 16), the first noun controls agreement pattern for adjectives and verbs, therefore, the leftmost member carries the morphosyntactic features.

2.5 Compounding in Other Languages

Compounding in languages from different language families has been extensively studied. Basciano et al. (2011) amongst other languages discuss Italian, Mandarin Chinese as well as German noun + noun and verb + noun root compounds.

In discussing noun + noun root compounds, they mention that Italian exhibit similarities with Bantu in relation to root compounding, which is different from Mandarin. They further indicate that Italian compounds are left-headed. Germanic languages are describes as

right-headed, and it is pointed out that root compounding is highly productive. Mandarin resembles Germanic languages in terms of being right-headed. Root compounding in Italian is attested even though it is interpretively constrained, whereas Chinese is different in this respect because the productivity of the root compounds is similar to that of Germanic languages, where noun + noun compounding is a very productive word formation process. So, in comparison, the occurrence of noun + noun compounds in Italian is lower than in Germanic languages.

The verb + noun compounds are pointed out to present challenging issues cross-linguistically. For example, Basciano et al. (2011) state that the form of the verb is not clearly defined and it is difficult to analyse. The unusual properties of verb + noun compounds seem to be shared in most languages including Bantu. For example it is not easy to decide if the form of the verb should be seen as an imperative, a noun or even a verb stem. These compounds in Sesotho are discussed as noun + noun synthetic compounds where the head noun which is derived has lost the class prefix, as opposed to them being verb + noun compounds. Instances of class prefix drop have been discussed in the previous chapter and are discussed more in relation to compounds in chapter 4.

2.6 Differences Between Compounds and Phrases

This section focuses on distinguishing between the sequences of morphemes which yield compound nouns and those that produce phrases. The topic of distinguishing compounds from their corresponding phrases has proven to be another area of interest among different studies of compounding as well. The literature review shows that most scholars suggest that compounds straddle the boundaries between morphology and syntax, whereas some scholars view compounding as a strict morphological process.

As a result, attempts have been made in the literature to find distinguishing criteria between compound nouns and phrases. Various authors (Bauer, 1998; Bauer, 2003; Booij, 2007; Fabb, 1998; Katamba, 1993; Lieber & Štekauer, 2009; Spencer, 1991) distinguish between compounds and phrases in their studies. Various criteria such as semantic, phonological, orthographic, and syntactic have been suggested in an attempt

to differentiate compounds from phrases (Nakov, 2013; Altakhaineh, 2016). However, there are debates regarding their reliability as most of them are found to be unreliable in accounting for the difference. Some are language specific as they apply to some extent in some languages and do not in others.

Bauer (2003) explains that compounding is similar to phrase formation because of the fact that compounds are sequences of lexemes which are constructed through rules of syntax. When defining a phrase, Katamba (1993:332) says it is a syntactic construction whose head belongs to different word classes such as an adjective, noun, verb, an adverb or a preposition. The phrase can comprise one word, two words or more. In the same way, a compound consists of two words or more. This therefore shows that the number of words in a construction does not indicate whether a construction is the compound or the phrase, except for when the phrase is made up of only one word.

Spencer (1991) is of the opinion that compounds constitute the interface between morphology and syntax and that this is specially the case with synthetic compounds. In distinguishing between compounds and phrases, Spencer mentions that concatenation of words to form phrases is an instance of syntax, while compounding is the concatenation of words to form other words.

Di Sciullo and Williams (1987) cited by Khan (2013:113) propose three ways in which words can be distinguished from phrases. Amongst those is the Lexical Integrity which is central to this study for the analyses of compounds and phrases. The Lexical Integrity proposes that sequences of morphemes constitute a word and not a phrase. According to the Lexical Integrity Principle (Booij, 2009:3), no parts of a word can be separated, moved, or deleted by rules of syntax. As a result, the principle acts against insertion (by modification) of elements which are not part of a word and disallows movement of the constituents of a word. The structure of a word is represented by principles which are in part different from syntactic ones (Anderson,1992). In other words, syntactic principles are not responsible for the formation of inflected or derived words.

Lexical Integrity therefore accounts for a clear-cut division between syntax and morphology. Lieber and Štekauer (2009:12) discuss modification criterion in

distinguishing compounds and phrases. According to the Lexical Integrity, as mentioned earlier, no elements which are not part of a compound can be inserted between its two constituents, whereas other elements such as the modifier word *ugly* can be inserted into the phrase such as *a black bird*, which would become *black ugly bird*. Nonetheless, no insertion is allowed with the compound *blackbird*. The modifier word *ugly* can only modify the compound as a whole and become *ugly blackbird* (Lieber & Štekauer, 2009).

In relation to movement of constituents of a compound, Booij (2009:3) points out that the syntactic movement of a word constituent, and the splitting up of words by intermediate constituents is impossible and as a result, perceives this impossibility as a necessary requirement for a particular linguistic unit to be a word.

2.7 Compounding in Optimality Theory

The study of compounding has been investigated in the light of various theoretic approaches; for example, in groups of theories such as transformationalist theories, lexicalist theories and syntactic theories. The discussion in this section however centres on the Optimality Theoretic approach to compounding. This is a constraint-based approach which was initially introduced to deal with phonological aspects but is now used in other linguistic fields such as semantics, syntax, morphology, pragmatics, among others.

A study on compounding within the Optimality Theoretic approach in Swahili which is a Bantu language has been conducted by Mashauri (2018). This study demonstrates that the constraint-based approach best accounts for the interface of grammatical modules in the formation of Swahili compounds. The interaction is between three modules namely; morphology, phonology and semantics.

According to Mashauri (2018:1), this constraint-based approach accounts for how words, which are compound nouns in this case, are accepted for use in Swahili when they meet all the grammatical requirements. An example is shown in the Swahili compound *kukumaji* (chicken + water) 'moorhen/water hen' analysis that is presented in Tableau 1 below.

The ranking schema for the analysis in Tableau 1 is provided below and the constraints clear definitions can be referred to in (Mashauri, 2018:5-8).

{ $\omega \Sigma \omega \Sigma n$ } & δ Erase) φ ({comp} & δ COMP1(M1)) >> Stress11(Comp) >> Stress1 {W# σ #1} comp & δ Vlength(W# σ #1) comp >> *Vlength(W1 σ #1 Σ Wn# σ #1) comp >> W1(SemHead) comp & δ NTE >> Align(W,R; ω ,R) comp & δ Align(W,L; ω ,L) comp >> FAITH {Mor} >> FAITH [Pho] >> Express {AffixNum/Ncl} >> W1 {Ncl.} comp & δ W1 {AffixNum} Comp

Tableau 1

/kukumaji/

Inputs ({[ku:ku]} ω), ({[ma:ji]} ω)	{ $\omega \Sigma \omega \Sigma n$ } & δ Erase) φ ({comp} & δ COMP1(M1))	Stress11(Comp)	Stress1 {W# σ #1} comp & δ Vlength(W# σ #1)	*Vlength(W1 σ #1 Σ Wn# σ #1) comp	w1(SemHead) comp & δ NTE EEXPR ESS NCL. (WORD 1)	Align (W,R; ω ,R) comp & δ Align (W,L; ω ,L) comp	FAITH {Mor}	FAITH [Pho]	Express {AffixNum/Ncl}	W1 {Ncl.} comp & δ W1 {AffixNum}
a. ({[ku:ku'ma:ji]} ω)				**				*	*	*
b. ({[ku:ku]('ma:ji]} ω)	* ³									
c. ({[kukuma:ji]} ω)		**						*	*	*
d. → ({[kuku'ma:ji]} ω)								*	*	*
e. ({[ku:ku'ma:ji]} ω)		**	*	*					*	*

³ Note that in OT analysis, the fatal violation rules out other violations of constraints that appear after it, that is why other violations for example, the last three constraints which are violated by candidate (b) are not indicated by the asterisk in this analysis. It is because their violation does not matter anymore since candidate (b) is ruled out by the first strong constraint. However in this study, all violations of constraints following the fatal one are shown.

Tableau 1 shows that candidate (a) is not optimal because it violates the vowel lengthening constraint $*Vlength(W1\sigma\#1\Sigma Wn\#\sigma\#1)comp$ by having long vowel on both words of the compound. Constraint FAITH[Phon] is violated by candidates (a, c, and d) by dropping stress and vowel length on the first word.

Express{AffixNum/Ncl}, $W1\{Ncl.\}comp$ and $\delta W1\{AffixNum\}Comp$ are violated by all candidates because they do not show noun class and number affixes. However, these constraints are weak therefore, tolerable in evaluating compound nouns of this structure. Candidate (b) violates the strong $\{\omega\Sigma\omega\Sigma n\}\&\delta Erase)\varphi(\{comp\}\&\delta COMP1(M1)$ constraint which requires all compound nouns to form a single unit of meaning by combining together all its component words (morphologically and phonologically). Candidate (c) violates the stress placement Stress11(Comp) constraint by having no stress on all words forming it and it also violates the last three constraints; so it loses competition.

Candidate (d) surfaces as the winner because it does not violate any strong constraints except for the weak ones. It violates FAITH(Phon) by dropping stress and vowel length on the first word of the compound as well as Express{AffixNum/Ncl}, $W1\{Ncl.\}comp$ and $W1\{AffixNum\}Comp$. The last candidate (e) violates many constraints including strong ones. It violates the constraints which require a compound to have one primary stress on the final word of the compound by having two primary stresses (on both words forming it). It also violate a constraint that forbids the compound to have long vowel on the penultimate syllable of the first and the second word of the compound and further violates Express{AffixNum/Ncl}, $W1\{Ncl.\}comp$ and $W1\{AffixNum\}Comp$.

In conclusion, Mashauri's (2018) Optimality Theoretic analysis of Swahili compounds reveals that for a compound noun to be considered acceptable in Swahili, it has to satisfy requirements set forth by the constraints of all the three grammatical modules, being morphological, phonological and semantic, and as a result, violating such constraints makes a construction uninterpretable and therefore not acceptable as has been seen from Tableau 1.

Another research in which compounding is analysed within an Optimality approach is that of Onwukwe, (2018) on Ibeme Igbo. The language analysed in this case is not Bantu but the study is discussed because it examines the morphophonological processes that underlie nominal compound formation in the same vein as Mashauri's (2018) which adopts Optimality Theory to account for the interaction between grammatical modules. This is relevant to the present study as it also investigates morphosyntactic properties of Sesotho compounds hence analysing the interaction between morphology and syntax modules.

This study reveals that there are two forms of nominal compounds in Ibeme Igbo namely; the phonologically sensitive and insensitive compounds where the sensitive class shows a high ranking of phonological constraints over morphological constraints and then the reverse is the case for the insensitive class. These two forms of compounds show variation in the data, as a result, two forms of ranking order in analysing morphophonological processes in the derivation of phonologically sensitive and insensitive noun compounds are identified. This study therefore accounts for this variation within OT alone by re-ranking constraints which OT does not support due to emphasis on ranking of constraints in strict hierarchy of domination to account for any data. It was shown that variation encountered in the present study in accounting for Sesotho compound data is accounted for by OT and Cophonology.

The Tableaux 2 and 3 illustrate the two forms of ranking order for Ibeme Igbo phonologically sensitive and insensitive compound nouns.

Tableau 2 corresponds to the compound *teetekwu* 'grandfather'

N1 + N2 =NC	NO V-HIATUS	OCP	ALIGN (N₁. L, N₂. R)	FAITH
a. teete+ukwu=teeteukwu	*!	*		
b. teete+ukwu =ukwoteete			*!	* *
c. teete+ukwu =ukwuteete			*!	*

d. → teete + ukwu =teetekwu				*
-----------------------------	--	--	--	---

The ranking schema for Tableau 2 is as follows: NO V-HIATUS>>OCP>>ALIGN>>FAITH. Candidate (a) shows no assimilatory process and tone changes in the formation of the noun compound while candidates (b) and (c) show wrong morphological form whereby N2 becomes N1. These do not capture the morphophonological processes in the derivation of the compound noun (teetekwu ‘grandfather’). No assimilation process rule enforced by NO V-HIATUS constraint which forbids the occurrence of adjacent vowels. This is a situation which triggers assimilation. No tone change rule is enforced by OCP which does not allow identical tones across word boundary or associative constructions. The morphological form of these nouns is enforced by the ALIGN constraint which generates the right morphological form of the nouns in the derivation of the compound noun. That is, ALIGN is violated by compound which have their constituents reversed. Candidate (d) is the winner because it incurs minimal violation of the least ranked FAITH constraint.

Tableau 3 corresponding to the compound *okpobekee* ‘main road’

N₁ + N₂ = NC	ALIGN (N₁. L, N₂. R)	FAITH	OCP	NO V- HIATUS
a. okpo+bekee =bekeeokpo	*!			*
b. →okpo+bekee=okpobekee				
c. okpo+bekee =ukorubekee		*!		
d. okpo+bekee =bekeeokporu	*!	*	*	*

The ranking schema for this type of compounds is ALIGN>>FAITH>> OCP>>NO V-HAITUS. The above tableau shows that, candidates (a) and (d) are not regarded as the morphological forms as N1 and N2 are juxtaposed. Candidate (c) does not depict the morphological form of N1 for the realization of a compound noun. The constraints NO V-

HAITUS and OCP in this case are ranked lower than ALIGN and FAITH. Candidate (b) is the optimal candidate. This is because in (b), the morphological form of the constituent nouns are retained which forecloses the phonological aspects of assimilation and tone changes leading to the derivation of *okpobekee*. This case shows that ALIGN AND FAITH constraints must not be violated for the optimal candidate to be realized in the derivation of the noun compound *okpobekee*.

The studies discussed above adopt an Optimality Theoretic approach in analysing compound nouns within more than one grammatical module. As a result, they set forth an important background for the present study.

2.8 Conclusion

This chapter discussed the relevant aspects of compounding to the present study. Highlighting issues that surround compounding studies. It was acknowledged that compounding has been studied cross-linguistically and that there are topics of interest to different scholars ranging from defining what compounding is, compounds' headedness as well as whether compounds should be analysed as morphological or syntactic constructions. Another important aspect is the theoretical basis of this study which was discussed in relation to studies of compounding which were conducted within it.

CHAPTER 3: METHODOLOGY

3.1 Introduction

The purpose of this chapter is to introduce the research methodology of this qualitative study which describes the Sesotho compounding phenomena within the morphosyntactic domain. The participants, data collection method, data processing and analysis as well as ethical considerations are the primary components of this chapter. Initially, this study sought to investigate the phonological, morphological, syntactic and semantic properties of Sesotho compounds in answering the following research questions:

1. What are the phonological and morphological properties of Sesotho compounds?
2. What are the syntactic and semantic properties of Sesotho compounds?

The above research questions were narrowed down to one question in (1) below, followed by sub-questions (1.1, 1.2, 1.3).

1. What are the morphosyntactic properties of Sesotho compounds?
 - 1.1 What are the combinatorial patterns, morphological structure, number and agreement properties of Sesotho compound nouns?
 - 1.2 Where is the head position of Sesotho compound nouns?
 - 1.3 Are Sesotho compounds morphological or syntactic constructs?

Modification of the initial research questions arose due to the fact that during the analysis process, the phonological and semantic properties of Sesotho compounds presented little information worth writing independent chapters on. Hence the discussion of such properties of Sesotho compounds in Chapter 1, Section 1.4 which outlines general background notes on Sesotho.

3.2 Participants

This study adopted a purposive sampling for the selection of participants. In a discussion by Tongco (2007), the purposive sampling technique is concerned with the deliberate choice of participants, based on the qualities they display. The sample for this study was

drawn from the student population of the National University of Lesotho and traditional doctors practising in Lesotho. The sample of twelve participants comprised: ten students majoring in African languages and two traditional doctors practicing in Maseru city. All participants had to be native speakers of Sesotho. The ages for the National University of Lesotho students ranged from 18-37 as there were out of high school and mature participants. The two traditional doctors were aged 41 and 60. They were both from the district of Maseru, whereas the students came from all over the country representing most of the districts of Lesotho. A request was sent to the head of department in African languages and literature at the National University of Lesotho to reach out to the students, of whom ten of them volunteered to take part in this study.

As for the traditional doctors, the requests for participation were hand-delivered and two traditional doctors also volunteered to participate. The inclusion of traditional doctors was based on their knowledge of names of traditional medicines, which was an area believed to have many Sesotho compounds.

3.4 Data Collection Method

During the data collection phase, primary and secondary data was collected. The methods used to collect the primary data involved interviews and questionnaires. The interviews were conducted one on one between the researcher and the participants. The participants were informed beforehand that there would be a recording taking place for the convenience of saving all the data provided, and they were comfortable with the recording process. Each interview lasted between ten and fifteen minutes. The interviews were recorded using a cell phone. After each session, the recording was saved to a laptop.

The interviews were conducted on campus for the National University of Lesotho participants. Each participant had indicated the time when they would be free to be interviewed. With regards to the traditional doctors, the interviews were conducted at their workplace.

The interview and questionnaire methods of data collection were used in combination in that the interviews were guided by the questionnaire. The questionnaire for the students comprised questions which required participants to give a list of names of things, for example, body parts, mountains, places, traditional food, among others; in order to elicit Sesotho nouns of which compound nouns were expected to be found. Such questions included for instance, “Can you give any place names that you know?” or “Do you know of any Sesotho traditional food?” The questionnaire also comprised English prompts wherein students were asked to provide Sesotho translations. Examples of the English prompts were nouns such as: an aeroplane, train, lightning, head office, library, among others.

A different questionnaire was administered to the traditional doctors. The questions required each traditional doctor to give a list of names of traditional medicines from different domains such as in animal names, plant names among others. The questions asked in this case for example were: “Do you give names to the medicines you produce?” “Can you tell me such names?”.

As for the secondary data collection, a variety of Sesotho materials were consulted to collect Sesotho compounds. These included books, dictionaries, articles such as: Doke and Mofokeng (1957); Guma (1971); Hlalele, (2005); Kalane (2005); Lekhotla la Sesotho, (1984); Mabile and Dieterlen (2011); Sebotsa, (2016).

3.5 Data Processing and Analysis

After the completion of data collection, the recordings from the interviews were listened to and a list of nouns was compiled from each group being: the ten National University of Lesotho students and the two traditional doctors. There were two hundred and eighty one Sesotho nouns collected from students and ninety of those collected from the traditional doctors. In total there were three hundred and seventy one Sesotho nouns. Target data which were Sesotho compound nouns were extracted from the list of nouns collected from both groups and a list of one hundred and thirty two compound nouns was compiled. Another word list was compiled from the compounds collected from the literature which were two hundred and fifty and then both word lists were compared, and duplicates were

deleted. At the end, Sesotho compound nouns collected for this study in total were three hundred and sixty excluding deleted duplicates. A list of Sesotho compounds is provided as an appendix but it does not include all the compounds collected for this study due to the great amount of space the tables occupied.

Following the data processing, a database was made from the compiled word lists and annotated using fieldworks language explorer (Flex), SIL International (2018). Flex enabled morphosyntactic analyses of Sesotho compound data, where the compounds were classified according to their shared properties such as: noun classes, components of each compound, headedness among others.

3.6 Ethical Considerations

Prior to the process of data collection considerations were made regarding important principles guiding conduct in collecting data from human participants. Before commencing with the interviews, the information leaflets and consent forms were given to the participants to read and understand what the study entailed. Then each participant was required to sign the consent form agreeing to all the requirements stated in the information leaflet. It was made clear to the participants that they were not obliged to participate in this study and that they could withdraw at any time, without giving reasons. Additionally, since recordings were made during the interviews, the participants were assured that, the information gathered for this study was confidential, as their names would not be mentioned and that nothing in the content would be traced back to an individual participant.

CHAPTER 4: MORPHOSYNTAX OF SESOTHO COMPOUNDS

4.1 Introduction

This chapter presents the types of Sesotho compounds identified from the data collected for this study and further discussed their formation within the morphosyntactic domain. Attention is specifically given to the compounds' combinatorial patterns, morphological structure, number assignment, agreement patterns, headedness and properties which distinguish compounds from similar constructions, being phrases.

The chapter is organised as follows: Section 4.2 outlines the types of Sesotho compounds, including discussions relating to their combinatorial patterns, morphological structure, number assignment and agreement patterns. Section 4.3 discusses the headedness phenomenon in compounding while section 4.4 distinguishes between compounds and phrases, then followed by section 4.5 which is on concluding remarks.

4.2 Types of Sesotho Compounds

The compound types include root and synthetic compounds. The following sub-sections show that the types of compounds are formed through combining different lexical categories which produce the following compound patterns: noun + noun, noun + adjective, noun + adverb, and noun + associative marker + noun. These compound patterns are discussed with regards to morphological structure, number and agreement.

4.2.1 Root compounds

As defined in Chapter 2, section 2.2, root compounds are compounds in which two non-derived words are combined. The compound patterns of this nature are identified from the data as noun + noun and noun + adjective. The noun + noun compounds are combined in two ways; first, two non-derived nouns are combined directly without a linking element to form a noun + noun pattern and second, two non-derived nouns are combined by the linking element which produces the pattern; noun + associative marker + noun. The following sub-sections discuss these compounds in detail.

4.2.1.1 Noun + noun compounds

Dressler (2007:33) points out that in most languages, the largest sub-class is made up of noun + noun compounds. Sesotho is not an exception, as there are many compounds belonging to the noun + noun pattern and these include both root and synthetic compounds. The noun + noun root compounds within the noun + noun sub-class are productive to a lesser extent and this was already mentioned by Basciano (2011) to be the case in Bantu languages. The noun + noun root pattern is less productive in Sesotho; however, the compounds are not restricted to kinship terms nor phytonyms like Basciano states. They exist in various semantic areas such as: humans, animals, plant names and place names among others. This can be seen from the compound examples listed in (2).

In comparison to noun + noun root compounds, the noun + associative marker + noun root compounds are highly productive, existing in various semantic areas but largely identified as plant names (Doke and Mofokeng, 1971). The noun + noun synthetic compounds in which the head noun is derived from a verb are by far the most productive within the noun + noun sub-class in Sesotho, and they are discussed in sub-section 4.2.2, which presents an overview of synthetic compounds.

Mphasha (2006) in a study on the compound noun in Sepedi (S32) which is a Bantu language of the Sotho-Tswana group identifies the same types of noun + noun compounds as in Sesotho. Mphasha provides examples of compounds without the linking element and those with the linking element. Cibemba (M42), which is a Bantu language spoken in Zambia also has these types of compounds although it belongs to a different language group from Sesotho and Sepedi. Kula (2012) identifies the same constructions as those in Sesotho and Sepedi in Cibemba and mentions that there are compounds which assume two root nominals, directly combined without the linking element and constructions which are referred to as associative constructions; in which two nominals are connected by a linking element; being the associative marker. Examples of root nominals without a linking element in Sepedi and Cibemba are provided in (1) below.

- (1) a. *mo-du* *∅-kala*
 NC3-root NC5-branch
 'branched root'

[Sepedi, Mphasha (2006:129)]

- b. *na-kulu* *ba-ntu*
 NC1a-elderly woman NC2-people
 'elderly (distinguished) woman'

[Cibemba, Kula (2012:431)]

The Sepedi compound in (1a) comprises the noun *modu* 'root' of class 3 and *lekala* 'branch' of class 5 which appears as *kala* without the class prefix *le-*. This is talked about in the later discussions as it happens in Sesotho compounds as well, where some nouns appear without their class prefixes in the formation of compounds. In (1b), the Cibemba compound is made up of *na-kulu* 'elderly woman' of class 1a and *a-ba-ntu* 'people' of class 2. It can be seen from these examples that there is a direct combination of two non-derived nouns, without the associative marker. Sesotho, Sepedi and Cibemba all have this type of noun + noun root compounds in which two nouns which are not derived are combined directly without a linking element. Sesotho examples are provided in (2) below.

The compound in (2a) for instance, is made up of two nouns: *sehoete* 'carrot' of class 7 and *moru* 'forest' of class 3. The two nouns also are not derived. The same analysis applies to the compound examples in (2b, c and d) as their constituents are non-derived nouns.

- (2) a. *se-hoete* *mo-ru*
 NC7-carrot NC3-forest
 'wild carrot'

- b. *∅-tsie* *ba-limo*
 NC9-locust NC2-ancestor
 'kind of locust'

c. *ø-khoho* *ø-noka*
 NC9-chicken NC9-river
 'common moorhen'

d. *ø-buka* *ø-ntsoe*
 NC9-book NC5-word
 'dictionary'

As mentioned earlier, another type comprises those compounds which are made up of two non-derived nouns connected by the linking element. Sesotho, Sepedi and Cibemba were earlier indicated to have these types of compounds. Examples regarding nominals connected by the associative marker in Sesotho and Sepedi are provided in (3a, b, and c) and (3d) for Cibemba. In (3a), the Sesotho compound is made up of the noun *boea* 'hair', the associative marker *ba* 'of' and the noun *ntja* 'dog'. The Sepedi example in (3c) consists of the two nouns *pudi* 'goat' in class 9 and *tsela* 'path' from class 9 which are connected by the associative *ya* 'of'. The Cibemba example in (3d) is made up of two nouns *ulukasu* 'hoe' and *mfumu* 'chief', which are linked by the associative maker *lwa* 'of'.

(3) a. *ø-koena* *ea* *ø-thaba*
 NC9-crocodile ASSC9 NC9-mountain
 'species of mint' [Sesotho, Doke & Mofokeng (1971:103)]

b. *bo-hobe* *ba* *li-nonyana*⁴
 NC14-bread ASSC14 NC10-bird
 'Digitaria monodactyla' [Sesotho, Doke & Mofokeng (1971:103)]

c. *ø-pudi* *ya* *ø-tsela*
 NC9-goat ASSC9 NC9-path
 'rumour' [Sepedi, Mphasha (2006:333)]

⁴ It should be noted here that the noun *linonyana* 'birds' is not in diminutive form. This is the result of frozen morphology.

d.	<i>ukuulu</i>	<i>kwa</i>	<i>nkoko</i>
	NC5.leg	ASSC5	NC9.chicken
	'chicken's leg'		[Cibemba, Kula (2012:436)]

Although in Sesotho, Sepedi and Cibemba the same constructions are established, in Sesotho and Sepedi these constructions are considered as compounds and are therefore analysed as words, whereas in Cibemba they are regarded as associative constructions which cut across the boundary between morphology and syntax and as a result, are seen as not being 'true compounds'. Mphasha (2006) in Sepedi labels these constructions as compounds although there is no account as to why they are considered compound words rather than phrases. This issue is further discussed in section 4.4 which discusses the difference between Sesotho compounds and phrases.

Referring back to examples (2a-d), the noun + noun root compounds without a linking element are morphologically formed by combining two nouns, each comprising the prefix and the noun stem. Basciano et al. (2011:207) in a discussion about noun + noun compounds in Bantu, presents the general schema of noun + noun Bantu root compounding as follows: [CM + N] + [(CM) + N]. The interpretation of this schema is that the root compound is made up of two nominal stems, each following its class marker (CM). Sesotho noun + noun root compounds conform to this schema.

This is evident from example (2a) where the first noun *sehoete* 'carrot' of the compound *sehoete-moru* 'wild carrot' consists of the prefix *se-* and the noun stem *hoete*. The second noun *moru* 'forest' is also made up of the prefix *mo-* and the noun stem *ru*.

The prefix of the second noun in the general schema above is put in brackets as an indication that sometimes the second noun may appear without the prefix. An example in Sesotho in which the second noun in a compound appears without the class prefix is *buka-ntsoe* 'dictionary' in (2d). The compound has the first noun *buka* 'book' from class 9 with the null class prefix and the second is *lentsoe* 'word' of class 5, which appears as *ntsoe* 'word' in the compound after shedding its class prefix *le-*. This is the case in Sepedi too, as this issue was raised in example (1a) in which the second noun *lekala* 'branch'

has lost its class 5 prefix *le-* and appears as *kala* 'branch' in the compound. In Tshivenda as well, this phenomenon is attested as it was seen in Chapter 2, section 2.4.

It is perceived that in Sesotho noun + noun root compounds, the first constituent may also shed the noun class prefix. An example is shown in (4) below, where the first noun *lefito* 'knot' of class 5 has lost its class prefix *le-* thereby appearing as *fito* 'knot' in the compound *fito-ntja* 'badly made knot'.

- (4) *∅-fito* *∅-ntja*
 NC5-knot NC9-dog
 'badly made knot'

From the root compound examples discussed, it is concluded therefore that in Sesotho, the first or second constituent of a compound may drop the class prefix. There is no case from the data in which both constituents appear without their class prefixes in a compound. Machobane et al. (2007) as discussed in Chapter 1, section 1.4.1 provides an account of Sesotho noun class prefix drop and provides examples of noun classes that are able to lose the class prefix and those that cannot and the circumstances in which nouns can lose their class prefixes. Phonologically, the class prefixes which can be dropped must belong to noun classes which have the prefixes which start with coronal consonants. Syntactically, agreement is a condition for class prefix drop in Sesotho simple nouns, this was indicated with an example in chapter 1, section 1.4.1.

Relative to compound nouns, the class prefix drop can be accounted for in that a compound has two elements in which the second element modifies or acts as a complement to the first one, as a result create an environment for the class prefix drop.

This issue of class prefix drop raises a question of what happens in terms of agreement in the case where the prefix of the first constituent of the compound is lost. The answer to this question is that the noun even though it has lost the prefix is still the determinant of agreement in which the modifiers adopt the noun class to which the noun belongs to. This is illustrated in the sentence in (5) below.

- (5) *Fito-ntja* *le le-holo le-chele.*
 NC5.knot-NC9.dog QP5 NC5-big SM5-burn.PST
 'Badly made knot is burned.'

It can be seen that the compound *fito-ntja* 'badly made knot' in (5) above still has its modifiers agreeing with the noun class which the compound belongs to, which is class 5 although the prefix does not appear in the compound. Regarding number, the compound examples in (2a-d) are listed in their singular forms. It was mentioned in Chapter 1, section 1.4.1 that Sesotho nouns generally have a particular singular prefix replaced by a corresponding plural prefix for number distinction. For instance, the plural form of the compound *sehoete-moru* 'wild carrot' in (2a) is *lihoete-moru* 'wild carrots' shown in (6) below. The singular prefix of class 7 *se-* is replaced by the plural prefix *li-* of class 8. It is observed with the noun + noun root compounds discussed here that the first constituent which is the head is the one which incurs change in relation to pluralisation. There is no change to the second constituent and if it were to be inflected for plural, the compound would lose its intended meaning.

- (6) *li-hoete* *mo-ru*
 NC8-carrot NC3-forest
 'wild carrots'

The same is the case with the compound examples in (2c-d) as their first constituents incur change where the singular prefix is replaced by the plural one, following the singular-plural pairing which was discussed in detail in Chapter 1, section 1.4.1. The observation with the singular-plural pairing as is seen with the examples discussed, is that one can predict the plural form of a given singular compound. This is however not always the case as there are exceptions. For instance, there are one class compound nouns. Examples of such compounds are given in (7) below.

- (7) a. *mo-hloa* *ma-fika*
 NC3-grass NC6-rock
 'kind of grass'

- b. *∅-thupa* *∅-kubu*
 NC9-stick NC9-hippopotamus
 ‘place name’

The compound in (7a) belongs to class 3 and it has the first constituent *mohloa* ‘grass’ as a mass noun. Thus, it does not take the plural form following the singular-plural pairing discussed earlier. The compound *thupa-kubu* ‘place name’ in (7b) of class 9 cannot take the plural form **lithupa-kubu*, where the class 9 null prefix of *thupa* ‘stick’ would take its plural as class 10 prefix *li*⁵- following the singular-plural pairing. Both the compounds do not distinguish the corresponding plural. These are therefore examples of one class compound nouns. In terms of noun class agreement, both compounds in (7) above still have the first noun as the determinant of agreement, where it controls which modifying elements to take.

Relating to Sesotho noun + noun root compounds in which the two nouns are combined by a linking element, the following morphological structure is produced: N1 + AM + N2. The first noun comprises the prefix which is followed by a noun stem, and then there is the associative marker which is followed by the second noun that is made up of the prefix and the noun stem. For instance, the compound in (3b); *bohobe-ba-linonyana* ‘digitaria monodactyla plant’ is made up of the noun *bohobe* ‘bread’ which has the prefix *bo-* of class 14 and the noun stem *hobe*, then followed by the associative marker *ba* ‘of’ and the noun *linonyana* ‘birds’, which consists of the prefix *li-* of class 10 and the noun stem *nonyana*. This compound noun is also a one class compound as it belongs to class 14 and does not take the corresponding plural form.

If the compound *bohobe-ba-linonyana* was to take the plural form, it would be in class 6 producing **mahobe-a-linonyana* which is regarded incorrect as a compound plural form of the singular compound *bohobe-ba-linonyana*. This would be contrary to the singular-plural pairing where one singular class was mentioned to take its plural from the

⁵ The plural class prefixes in noun classes 8 and 10 are morphologically identical, so the compound in (7b) should not be mistaken to would have taken its plural in 8. The compound would take the plural class prefix in 10 because of the influence of the underlying nasal prefix *N-* of class 9 as opposed to the prefix *se-* of class 7 where corresponding plural nouns would be in class 8.

corresponding or following plural class, so this would be another exception of the singular-plural pairing, where class 14 would take plural in class 6 instead of 15. Then it would be a collective plural. So, *mahobe a-linonyana* is not a compound plural but a phrase in plural form referring to the kinds of bread belonging to the birds. A compound example which follows the singular-plural pairing is provided in (8a) below, together with its plural form in (8b).

- (8) a. *ø-koloi* *ea* *mo-llo*
 NC9-car ASSC9 NC3-fire
 'train'
- b. *li-koloi* *tša* *mo-llo*
 NC10-car ASSC10 NC3-fire
 'trains'

The noun class agreement is seen clearly in plural forms of noun + associative marker + noun examples. It can be seen that agreement is driven by class of the first noun, which is carried on to the associative marker. Hence, the associative marker agrees in number with the noun class of the first noun in plural form. This is a further indication of the importance of the the head noun, in relation to agreement.

4.2.1.2 Noun + adjective compounds

These compounds as seen from the examples in (9) comprise a noun and an adjective. The constituents of the noun + adjective compounds are non-derived just as noun + noun compounds discussed in the previous section are. The adjectives in compound nouns are marked with the noun class of the noun they modify, and they also distinguish between singular and plural. As a result, these compound patterns exhibit noun-adjective internal agreement.

- (9) a. *mo-lala* *mo-tšo*
 NC3-neck NC3-black
 'turtle dove'

- b. *ma-ru* *ma-tšo*
 NC6-cloud NC6-black
 'name of a dog'
- c. *∅-saka* *le-holo*
 NC5-kraal NC5-big
 'collective kraal for the community'
- d. *∅-ntlo* *∅-kholo*⁶
 NC9-house NC9-big
 'head office'

The compound example in (9a) *molala-motšo* 'turtle dove' is morphologically composed of: the noun *molala* 'neck' with the prefix *mo-* and the noun stem *lala* and the adjective *motšo* 'black' with the prefix *mo-* and the adjectival root *tšo* 'black'. The compound *maru-matšo* in (9b), shows that adjectives distinguish singular and plural in that, the noun *maru* 'clouds' of class 6 is in plural form; and therefore, the adjectival root takes the plural prefix *ma*. Whereas if the noun was *leru* 'cloud' of class 5 in singular, then the adjective would take the singular prefix *le-* and become *leru-letšo*⁷.

This means therefore, that all adjectival roots can take all 14 Sesotho noun class prefixes. This is proven by examples in (9a-d), where the adjectival roots take prefixes in classes 3, 5, 6 and 9, depending on the nouns they modify.

It should be observed once more that the compound in (9c) has its first noun *lesaka* 'kraal' appearing without the class prefix *le-* of class 5, just as it was seen with the noun + noun compounds in section 4.2.1.1. It should be observed here as well that the adjectival root adopts the noun class of the noun it modifies even though the class prefix is lost. This is still an indication of noun class agreement. The head in this noun + adjective pattern is

⁶ In this compound, there is a nasal influence of class 9 prefix *N-* on the following adjectival root. As a result, the sound */h/* in *-holo* 'big' goes through the process of strengthening and becomes */x/* in *-kholo* in the compound *ntlo-kholo* 'head office'.

⁷This is a non-existent example but a possible one in Sesotho.

still the determiner of agreement relative to which modifying elements such a compound should take.

It was mentioned that adjectives in noun + adjective compound pattern must agree with the nouns that they modify. As a result, they have a morphological structure that is similar to that of nouns in the compound formation. For instance, the second word *motšo* in the compound in example (9a) is actually an adjective not a noun. This is simply the case in which the adjectival root is preceded by the class prefix of the noun it modifies.

The word *mo-tšo* is concluded to be an adjective based on the following arguments. It cannot function in a sentence the same way a noun can. If the word *motšo* was a noun, it would syntactically function as a subject or object in a sentence. For example, * *motšo o fihlile* with the intended meaning 'motšo has arrived', but because it is an adjective it cannot assume the subject position, hence the example is not semantically acceptable. The example is grammatically correct but semantically unacceptable because there is no such noun as *motšo* in Sesotho. The correct sentence would be: *monna e-motšo o fihlile* 'black man has arrived' where the adjective syntactically functions as a modifier of the noun. This is the same as supposing we had a compound *sefate-setšo* (tree + black). *Setšo* would not be a noun as the adjectival root *tšo* is just preceded by the class prefix *se-* of *sefate*. Rather, examples of nouns derived from the adjectival stem *tšo* are such as those given in (10) below.

- (10) a. *bo-tšo*
 NC14-black
 'state of being black'
- b. *ra-bo-tšo*
 NC1a-NC14-black
 'name of a person'

The noun *botšo* 'the state of being black' is derived from the adjectival root *tšo* which is preceded by the class prefix *bo-* of class 14. This noun appears morphologically the same as the word *motšo*, which was earlier argued to be an adjective based on the fact that it

could not assume the same syntactic positions as nouns. Note that supposing we had a compound *bohobe-botšo* (bread + black) from class 14, *botšo* would be regarded as an adjective therefore different from the noun *botšo* of class 14, even though they are morphologically identical.

The difference is that in a compound the word *botšo* would be modifying the noun preceding it, while as a noun *botšo* refers to the state of being black. As for the noun *rabotšo*, the prefix *ra-* meaning ‘father of’ which is found in class 1a is affixed to the noun *botšo* to form a proper noun which is the name of a person. It is concluded that since a noun is a word which identifies for instance, a person, a place, an object or an idea, *motšo* in the compound *molala-motšo* as well as other examples in noun + adjective compounds do not do that. Instead they modify nouns. Relating to number, an example of the plural form of a singular compound *molala-motšo* in (9a) is provided in (11) below.

- (11) *me-lala* *me-tšo*
 NC4-neck NC4-black
 ‘turtle doves’

From example (11) above, the noun and the adjective as the constituents of the noun + adjective pattern take plural at the same time, due to noun-adjective internal agreement. The noun-adjective noun class internal agreement is obligatory as the compound formed by pluralising the noun only, as it happened with the noun + noun root compounds would be ungrammatical. For example, **melala-motšo* with the intended meaning ‘turtle doves’ would be unacceptable. The singular compound *molala-motšo* belongs to class 3 and it takes its plural as *melala-metšo* in class 4, following the singular-plural pairing.

4.2.2 Synthetic compounds

The term synthetic compounding has long been established to indicate that this kind of word-formation involves the concurrence of compounding and derivation (Booij, 2005). In synthetic compounds as opposed to root compounds, the head of a compound is derived from a verb. These were discussed in Chapter 2, section 2.2. The derivation process results in nouns derived from verbs which are referred to as deverbatives. The word

classes combined to form these compounds found in the data are noun + noun, noun + adjective and noun + adverb. Their examples are given in the following sub-sections in which it is shown that constituents of a compound may be made of deverbatives which convey agentive or non-agentive meaning.

Additionally, it is shown that the synthetic compounds are of two types. First, the compounds whose first constituent (a derived noun) appears with the class prefix. Second, there are those compounds which lack the class prefix just like the examples seen in the previous sections about root compounds, in which some of the members lost the class prefixes. Sub-section 4.2.2.1 discusses synthetic compounds with a class prefix and sub-section 4.2.2.2 concerns itself with synthetic compounds without a class prefix.

4.2.2.1 Synthetic compounds with the class prefix

These compounds comprise deverbative nouns which have the morphological structure; class prefix + verbal root + suffix. It was mentioned earlier that the derived constituents of these compounds denote agentive and non-agentive meaning. Non-agentive nominalisation involves those deverbatives which do not denote an agent of an action and are mostly non-human nouns. Such compound examples are listed in (12) below.

- (12) a. *∅-tseb-o* *∅-ntsoe* (verb stem; *tseb-a* 'know')
 NC9-know-NMLZ NC5-word
 'lexicon'
- b. *∅-kutu* *∅-kets-o* (verb stem; *ets-a* 'do')
 NC9-root NC9-act-NMLZ
 'verb'
- c. *mo-la-o* *∅-phetol-o* (verb stems; *la-a* 'advise' and *fetol-a* 'change')
 NC3-advise-NMLZ NC9-transform-NMLZ
 'transformational rule'

In the compound noun *tsebo-ntsoe* ‘lexicon’ in (12a), the first deverbative noun *tsebo* ‘knowledge’ is derived from attaching a class 9 null prefix and the suffix *-o* for non-humans to the verbal stem *tseb-a* ‘know’, which is made up of the verbal root *tseb* and the final vowel *-a*. The compound therefore has the first noun as the deverbative. The deverbative in this case does not have the function of an agent doing something.

Kutu-ketso in (12b) involves the second noun as the deverbative. *Ketso* ‘action’ is formed from the verbal stem *ets-a* (verbal root *ets* + final vowel *-a*) ‘do’ which is affixed with the null class 9 prefix and the suffix *-o*. The compound noun *molao-phetolo* ‘transformational rule’ in (12c) is made up of two deverbatives: *molao* which is derived through attaching the prefix *mo-* and the suffix *-o* to the verbal stem *la-a* (verbal root *la* + final vowel *-a*) ‘advise’ and the second noun *phetolo* ‘transformation’ is derived by attaching a class 9 null prefix and the suffix *-o* to the verbal stem *fetol-a* (verbal root *fetol* + final vowel *-a*) ‘change’. It can be seen that the non-agentive deverbatives in the compound examples carry the non-human suffix *-o*.

These compound nouns in Sesotho comprise a non-agent denoting deverbative as the first constituent, second or both, as indicated from the examples. As a result, the conclusion is that there are three possible structures of compounds made up of non-agent denoting nouns. It should be noted that there are some phonological changes taking place when some of these deverbatives are derived, for example in the case of deverbatives; *ketso* and *phetolo*. This was discussed in Chapter 1, section 1.4.2.1 on phonological sound changes in deverbatives. In relation to pluralisation, the compound *kutu-ketso* of class 9 in (12b) takes its plural in class 10, as indicated in (13) following the singular-plural pairing.

- (13) *li-kutu* *∅-ketso*
 NC10-root NC9-action
 ‘verbs’

In this compound and similarly, as with noun + noun root compounds, the plural prefix is attached to the first constituent of the compound and the second constituent is left as is. The compound *tsebo-ntsoe* of class 9 in (12a) is a one class compound noun, it does not

take the plural form in class 10 as would be expected. Example (12c), *molao-phetolo* of class 3 takes its plural in class 4 as *melao-phetolo*. The head noun in the compound examples in (12) and (13), whether in singular or plural form, is still in control of the accompanying modifiers and, as a result, emphasises the fact that the first noun of a compound determines agreement.

Concerning deverbatives denoting agentive meaning, these are made up of action verbs which express something that a person, an animal or an object for example, is doing. These were also discussed with examples in Chapter 1, section 1.4.1. The agentive nouns in Sesotho are used in formation of compound nouns as well. Schadeberg (2003:79) says that a productive type of compounds in Bantu has the morphological structure; [NPx=B-a] + N. What this structure means is that a compound can be made up of the agentive noun comprising: [noun class prefix + the base or verb root + the final vowel -a] and then followed by a noun.

Sesotho synthetic compounds adhere to this morphological structure. However, the second constituent of a compound which Schadeberg (2003) mentions is a noun in the morphological structure may be represented for example by: a noun, an adjective or an adverb in Sesotho. This is proven in examples in (14) where the deverbative *sebitsa* in (a) is followed by the noun and then the deverbative *moja* in (b) is followed by an adverb, while the deverbative *maoa* in (c) is complemented by the adjective.

- (14) a. *se-bits-a* *ba-nna*
 NC5-call.AGT NC2-man
 ‘muti used to attract men’
- b. *mo-j-a* *morao*
 NC1-eat.AGT last
 ‘one who eats last’

- c. *ma-o-a* *ma-fubelu*
 NC6-fall.AGT NC6-red
 'traditional medicine'

Relating to pluralisation, the compound noun in (14a) is a one class example as it does not take the plural form. The compound in (b) takes the singular-plural pairing pattern and becomes *baja morao* 'those who eat last' of class 2. The compound *maoa-mafubelu* is a one class which is already morphologically plural. The first nouns in this case also determine agreement patterns.

Other compounds with the deverbative denoting agentive meaning are provided in (15). The first noun in these examples is derived from an extended verbal base. For example, the compound in (15a) has the agentive *lithibela* made up of the class prefix + the verbal root + the applicative + the final vowel. In the compound in (15b), the agentive noun *mahlabisa* comprises the prefix + the verb root + the causative + the final vowel. The non-human agentive *lithibela* in the applicative form designates an action done for the object, which is the complement *pelehi*; while in the non-human agent *mahlabisa* in the causative form, an action is caused to happen by an agent.

- (15) a. *li-thib-el-a* *∅-pelehi*
 NC8-prevent-APP-AGT NC9-childbirth
 'contraceptives'
- b. *ma-hlab-is-a* *li-hlong*
 NC6-pierce-CAUS-AGT NC10-shame
 'things that cause shame'

The compound in (15a) on the one hand is already in plural form and has its singular form as *sethibela-pelehi* 'contraceptive'. On the other hand, the compound in (15b) is a one class compound which is morphologically plural. There is no singular form corresponding to it. Another type of plural which synthetic compounds exhibit is contextual plural which was discussed in chapter 1, section 1.4.1. There are synthetic compounds which belong to classes 1 and 7 referring to humans which take two plural forms depending on their

meaning, where the plural could be plain plural following singular-plural pairing and contextual. This contextual plural is found in class 2a. An example of such a compound is provided in (16) below.

- (16) a. *se-il-a* \emptyset -*tsatsi*
 NC7-avoid-AGT NC5-sun
 ‘beautiful woman’
- b. *li-il-a* \emptyset -*tsatsi*
 NC8-avoid-AGT NC5-sun
 ‘beautiful women’
- c. *bo-se-il-a* \emptyset -*tsatsi*
 NC2a-NC7-avoid-AGT NC5-sun
 ‘beautiful woman in company of others’

It can be seen from the examples that the singular compound *seila-tsatsi* in (16a) which is also another instance of the class prefix drop where the second noun *letsatsi* appears without the class prefix takes two plural forms. The plural form in (16b) follows the singular-plural pairing while the plural in (16c) is contextual. It does not refer to a beautiful woman only but the beautiful woman in company of others.

The observation made from the synthetic compounds discussed above is that agentive nouns appearing as the first constituent in these compounds can never exist as words on their own, even though they are referred to as nouns in the compound examples. For example, *sebitsa* in (12a) is not an independent word. They assume the noun category when followed by their complements. As a result, a word which is a compound only arises when these agentive nouns appear with complements like in *sebitsa-banna*.

Bloomfield (1933) is of the view that forms found in complex constructions may have special features in comparison with their independently occurring counterparts. This is why there is the special noun *sebitsa* ‘caller’ in a compound *sebitsa-banna* instead of *sebitsi* ‘one who calls’ which is an independent agentive noun.

Another form of a compound in example (17) below was found, which is formed through combining an independent agentive noun with a non-agentive denoting noun. This means therefore that this compound is different from those in which their agentive nouns end with the suffix *-a*, such as examples in (14), (15) and (16) which are deemed special.

- (17) *se-bapal-i* *mo-elel-o*
 NC7-play-AGT NC3-mean.NMLZ
 'thematic role'

The agentive noun *sebapali* 'player' of the compound *sebapali-moelelo* above is derived from attaching the prefix *se-* and the agentive personal suffix *-i* to the verbal root *-bapal-* 'play' and the second constituent is a non-agentive denoting deverbative *moelelo*, which is made up of the prefix *mo-*, the verbal root *elel* 'mean' and the non-human suffix *-o*. The agent *sebapali* 'player' does the action of playing. The compound *sebapali-moelelo* 'thematic role' is from class 7 and takes its plural form *libapali-moelelo* 'thematic roles' from class 8.

4.2.2.2 Synthetic compounds without the class prefix

These compounds consist of two constituents being a verb stem as the first constituent and a noun as the second constituent. These types of compounds are mostly referred to as verb + noun compounds and there are debates as to what verb form should be given to the first constituent, for example whether it is a verb stem or an imperative. In this study, they are argued to be synthetic compounds which are different from the ones discussed in sub-section 4.2.2.1; because their first constituent which is the deverbative does not appear with its class prefix, which was referred to as the class prefix drop in root compound discussions, as a result appears like a verb stem. The plural forms support this argument in that they carry overt noun class prefixes, proving that the singular form has just undergone the class prefix drop.

The first constituent in singular and plural forms is still considered an agentive noun just as in the synthetic compound, with the first constituent appearing with the class prefix. This is argued further in the following discussions. Basciano et al. (2011:220) states that

these compounds are attested in Bantu languages and indicate that they exhibit the structure: (CM/prefix) + [\sqrt{V} + fv] + [(CM) + \sqrt{N} Noun]. This structure suggests that the first constituent is a derived noun because of the indication that the first morpheme could be a class marker or a prefix. The interpretation of the structure is therefore that; the first constituent of a compound which is the derived noun, may appear with or without the class prefix (CM) and that the second constituent which is always a noun may sometimes shed the class prefix as it was seen with the root compounds or the synthetic compound *seila-tsatsi* discussed from the previous section.

The Sesotho compounds given in (18) below follow the same structure provided by Basciano et al. (2011:220). The compound in (a) has the first constituent with a zero prefix, thereby appearing as the verb stem *phutha* ‘collect’. The second constituent is a noun which is made up of the prefix *li-* of class 8 and the noun stem *chaba* ‘multitude’.

- (18) a. \emptyset -*phuth-a* *li-chaba*
 NC9-collect-AGT NC8-multitude
 ‘place name’
- b. \emptyset -*thul-a* *ma-bota*
 NC9-hit-AGT NC6-wall
 ‘type of dance’
- c. \emptyset -*thets-a* *ba-lisa*
 NC9-deceive-AGT NC2-herdsman
 ‘red capped lark’
- d. \emptyset -*thal-a* *bo-liba*
 NC9-frolic-AGT NC14-deep hole in a river
 ‘whirligig’

It was mentioned earlier that the plural forms of this compounds take the same structure as the synthetic compounds discussed in the previous section, as a result, this led to the

conclusion that these compounds also are synthetic compounds with the difference being that the singular forms appear without the class prefix. Examples of the plural forms of the compounds *thetsa-balisa* and *thala-boliba* in (18c and d) are provided in (19) below. The compounds in (18 a and b) are one class examples which do not take plural forms.

- (19) a. *li-thets-a* *ba-lisa*
 NC10-deceive-AGT NC2-herdsman
 ‘red capped larks’
- b. *li-thal-a* *boliba*
 NC10-frolic-AGT NC14-deep hole in a river
 ‘whirligigs’

In conclusion, the synthetic compounds from the literature are mostly defined as those compounds which involve the derived noun as the head of the compound. It is seen from the discussions above that the class of synthetics in Sesotho is actually larger than those with a derived head only, meaning that in Sesotho compounds, the term synthetic is considered broadly to refer to all the synthetic compound structures where the first constituent which is the head is always derived from a verb, or where the second constituent may also be derived and that synthetics can appear with or without the class prefix.

4.3 Compound Nouns and Headedness

This section focuses on identifying the position of the head in Sesotho compounds. Illonga (2016) is of the view that in Bantu languages the constituent on the left of a compound is most likely to be the head. Sesotho is no exception as compounds are predominantly left-headed. The head is the fundamental element of a compound (Dressler, 2007), as it carries the morphosyntactic features. The notion of headedness determined semantically, results in three categories of compounds being endocentric, exocentric and coordinate which are discussed in the following sections.

4.3.1 Endocentric compounds

An endocentric compound is one that has a semantic head as has been discussed in Chapter 2, section 2.3. This type of compound inherits most of its semantic and syntactic information from its head. Sesotho compounds presented in earlier sections are analysed in this section in terms of endocentricity. Starting with noun + noun compounds without a linking element, which were discussed in (2), these compounds are endocentric and they are semantically analysed on the basis that X is a kind of Y and syntactically, in that the head assigns the lexical category to the whole compound. For instance, the compound *tsie-balimo* ‘kind of locust’ in example (2b) is a kind of locust and it takes the noun as its lexical category, therefore *tsie* ‘locust’ is the head, respectively. The same goes for examples (a and c), in that the nouns *buka* ‘book’ in *buka-ntsoe* and *khoho* ‘chicken’ in *khoho-noka* are heads. So, the first noun in these noun + noun compounds is the head; thus, they are left-headed.

There is however, one noun + noun compound which was found in the data which has the second noun as the head. Thus, it is right-headed as opposed to examples discussed which reflect left-headedness. Such a compound *mochana-pula* ‘little rain’ is provided in (20) below. The head of this compound is the second element *pula* ‘rain’, therefore it is regarded as endocentric as well. This is so because *mochana-pula* is a kind of rain. This kind of compound then brings forth the question of whether there are more of these compounds. In any case, the data still indicates that Sesotho compounds are predominantly left-headed.

- (20) *mo-chana* *ø-pula*
NC1-nephew NC9-rain
‘little rain’

Noun + noun compounds with a linking element were also discussed and the endocentric ones were not found. All the compounds of this nature were found to be exocentric. Examples of these were discussed in examples (3) and (8). Another type of noun + noun compounds discussed earlier in section 4.2.2 is the synthetics, most of which denote endocentric compounds. Consider the examples in (21) below which are synthetic compounds with the class prefix. The compound *sebolaea-likokonyana* ‘insecticide’, in

(21a) has the head as the agentive *sebolaea* ‘killer’. The head in this case, still assigns the lexical category to the whole compound. In (21b), a noun and an adverb are combined, and the head is the first noun *moja* ‘eater’ and it shares the same properties as the whole compound. The synthetic compounds without the class prefix were all found to be exocentric and they are discussed in the following section.

- (21) a. *se-bolae-a* *li-kokonyana*
 NC7-kill-AGT NC10-insect
 ‘insecticide’
- b. *mo-j-a* *morao*
 NC7-eat-AGT last
 ‘one who eats last’

Noun + adjective compounds were also discussed and when analysed in terms of headedness, take the first element which is the noun as the head, just as the noun + noun compounds but they are not as productive as their exocentric counterparts. For example, *lijo-thollo* (food + grain) ‘grain food’ of class 10 is given in (22). This compound refers to the kind of food therefore deemed endocentric.

- (22) *li-jo* *ø-thollo*
 NC10-food NC9-grain
 ‘grain food’

4.3.2 Exocentric compounds

Exocentric compounds are compound nouns which lack a semantic head. These were discussed in Chapter 2, section 2.3. There are fewer exocentric compounds in noun + noun examples. Noun + adjective ones are productive. Noun + noun and noun + adjective exocentric compounds are provided in (23a-b).

The noun + adjective compound *poho-tšehla* in (23a) is not a type of a bull colour but Sesotho medicine. The noun + noun compound *ntja-‘me* in (23b) does not refer to a kind of dog but a brother or sister (Kalane, 2005).

The other noun + noun compounds were those mentioned to have the linking element. Examples are in (23c-d), where it is concluded that they are mostly exocentric because examples of endocentric ones were not found. Looking at example in (23c), *boea-ba-ntja* (hair + of + dog) is not dog's hair but a plant. In the same vein, example (23d) *koena-ea-thaba* (crocodile + of + mountain) is not a type of crocodile but a plant.

- (23) a. *ø-poho* *ø-tšehla*
 NC9-bull NC9-yellow
 'type of muti'
- b. *ø-ntja* 'me
 NC9-dog NC1a.mother
 'brother or sister'
- c. *bo-ea* *ba* *ø-ntja*
 NC14-hair ASSC14 NC9-dog
 'Bulbostylis striatella plant'
- d. *ø-koena* *ea* *ø-thaba*
 NC9-crocodile ASSC9 NC9-mountain
 'species of mint'

Noun + noun compounds also have exocentric metaphoric compounds, where it is observed that the two constituents do not contribute to the interpretation of the whole compound but do have descriptive features of how the referent looks like. The noun + noun compound in (24) below serves as an example of exocentric metaphoric compounds. *Molomo-khaba* 'African spoonbill' is a type of bird not a kind of mouth; however, the bird has a mouth which looks like a spoon hence the metaphoric exocentricity.

- (24) *mo-lomo* *ø-khaba*
 NC3-mouth NC9-spoon
 'African spoonbill'

Some exocentric noun + adjective compounds also have metaphoric character. An example is that of the compound *molala-motšo* ‘turtle dove’ in (25) below. This is actually a bird in which its features are used metaphorically to refer to it. It has a black line on its neck. As a result, the compound is not a kind of neck but a kind of bird.

- (25) *mo-lala* *mo-tšo*
 NC3-neck NC3-black
 ‘turtle dove’

There are examples of exocentric synthetic compounds as well. Considering the examples in (26) below none of the constituents from the examples given contribute to the interpretation of the meaning of the compounds and are hence categorised as exocentric compounds.

- (26) a. *se-ol-a* *mo-lora*
 NC7-collect-AGT NC3-ash
 ‘newly emerging breast’
- b. *le-hlab-a* *ø-phieo*
 NC7-pierce-AGT NC9-kidney
 ‘traitor’

In (26a), the compound refers to a breast, as a result the meaning of the whole compound is not deduced from its constituents being *seola* ‘collector’ and *molora* ‘ash’ The same goes for the compound in (26b). There are also exocentric metaphoric synthetic compounds and they are provided in (27) below. In (27a) the compound *sekoalla-bosulu* ‘jumpsuit’ is referred to by its characteristics. That is, it is said to keep gas that is emitted when farting inside it with nowhere to escape through. With regard to example (27b) *leja-hlapi* ‘European’, the belief is that Europeans come from overseas where there is plenty of fish, hence the given compound noun.

- (27) a. *se-koalla* *bo-sulu*
 NC7-keeper NC14-gas emitted when farting
 ‘jumpsuit’

- b. *le-ja* *ø-hlapi*
 NC5-eater NC9-fish
 ‘European’

All synthetic compounds without the class prefix in Sesotho fall under the category of exocentric compounds. It can be seen from the example given in (28) below that neither constituent of the compound contributes to the interpretation of the meaning of the whole compound.

- (28) *ø-tim-a* *le-bone*
 NC1a-extinguish-AGT NC5-lamp
 ‘name of a person’

4.3.3 Coordinate compounds

Coordinate compounds were defined in Chapter 2, section 2.3 as compounds in which there is a relation of coordination between their two constituents, and, as a result, they hold equal status with regards to headedness (Bisetto & Scalise, 2005). In these compounds, both constituents semantically contribute to the interpretation of the whole compound. Three compounds found in the data collected for this study are argued to be coordinate examples. Examples of such compounds in Sesotho are given in (29) below.

- (29) a. *mo-tho* *se-bata*
 NC1-person NC7-beast
 ‘carnibal’
- b. *ø-tšoene* *mo-tho*
 NC9-baboon NC1-person
 ‘ugly person’
- c. *le-jo-ana* *mo-tho*
 NC5-stone-DIM NC1-person
 ‘person-like stone’

The above are argued to be coordinate compounds because they are all made up of two equal components and this is evident in the following ways: the components of each compound in (29) above both require number marking, which is seen in (30) below.

- (30) a. *ba-tho* *li-bata*
 NC2-person NC8-beast
 ‘carnibals’
- b. *li-tšoene* *ba-tho*
 NC10-baboon NC2-person
 ‘ugly people’
- c. *ma-jo-ana* *ba-tho*
 NC6-stone-DIM NC2-person
 ‘person-like stones’

These compounds could be argued to be noun + noun endocentric compounds in which (29a) is a kind of man rather than a kind of beast, (29b) is a type of person, rather than a type of baboon and similarly (29c) as a kind of stone rather than a kind of person. However, this is not the case, based on examples in (30) above, in plural forms which show that the referents are described with equal characteristics and therefore are pluralised equally. For example, *motho-sebata* refers to a carnival with equal characteristics of a person and a beast, being a person who eats other people as the beast does. Pluralising the first noun only as in **batho-sebata* with the intended meaning ‘carnibals’ is not acceptable as the compound loses the equal interpretation of the characteristics defining the referent resulting in deviation of intended meaning.

As a further indication that the characteristics defining the referent are equal in these compounds, the linear order of the constituents can be changed without interfering with the meaning of these compounds. This is seen in (31) below.

- (31) a. *se-bata mo-tho*
 NC7-beast NC1-person
 ‘carnibal’
- b. *mo-tho ø-tšoene*
 NC1-person NC9-monkey
 ‘ugly person’
- c. *mo-tho le-joana*
 NC1-person NC5-stone
 ‘person-like stone’

The compound *lejoana-motho* in (31c) is a stone which looks exactly like a person. This, therefore, means that the compound has features of both being a stone and resembling a person. As a result, the compound can be reversed, producing *motho-lejoana*. The meaning being the ‘person-like stone’ is still retained because the reversed version still has both equal features of the stone and the person.

4.4 Differences Between Compounds and Phrases

This section looks at how compound nouns are distinguishable from their corresponding phrases. All languages have phrases. It has not been determined however, if all have compounds. With languages in which both compounds and phrases exist, these two constructions pose difficulty in trying to make a distinction between them. Many studies have been conducted to distinguish between compounds and phrases, as they share common properties (Bauer, 1998; Bauer, 2003; Booij, 2007; Fabb, 1998; Katamba, 1993; Lieber & Štekauer, 2009; Spencer, 1991). Hence, it is deemed necessary to make a distinction between the two in this study.

Various criteria have been used in the study of compounding to differentiate between compounds and phrases. Sesotho compound nouns identified in this study are tested following the Lexical Integrity Principle (Booij, 2009). Booij’s (2009:3), Lexical Integrity

states that the syntactic rules do not apply to the internal structure of a word since the word is considered as a single morphological unit.

The implications of the Lexical Integrity Principle by Booij to be adopted in this discussion are that: firstly, syntactic rules cannot manipulate parts or internal structure of a word, which results in no movement of parts of a word. So, the linear order of the constituents of a compound is very important in most Sesotho compounds, secondly, syntactic rules do not have access to the internal structure of words, which results in no modification within parts of a word. These are expected to mostly hold true in Sesotho compounds as the exceptions to the first implication are coordinate compounds which allow movement of constituent without losing the intended meaning. The compounds introduced above are compared with their corresponding phrases in this section. In analysing noun + noun compounds, this study follows Kula (2012) and the Lexical Integrity. Kula (2012:441) rules noun + noun compounds without a linking element in Cibemba, as true compounds. Consider the compound *sehoete-moru* in example (32) below.

- (32) *se-hoete* *mo-ru*
 NC7-carrot NC3-forest
 'wild carrot'

In this case semantics can be one distinguishing factor between compounds and phrases, for instance, the compound *sehoete-moru* refers to an entity 'wild carrot' while the syntactic structure *sehoete-sa-moru* 'carrot of the forest' describes an entity. As a result, the compound on the one hand functions differently grammatically because it is treated as a single word, whereas the construction *sehoete-sa-moru* is not a single word, but a noun phrase.

Following the Lexical Integrity Principle with regards to modification, the compound *sehoete-moru* does not allow any elements between its parts. For example, (33a) is an instance of a compound noun which is internally modified. The adjective *se-seholo* 'big' cannot modify the noun *sehoete* 'carrot' independently since both the constituents *sehoete* and *moru* form one word which is the compound. Rather, the modifier element appears outside the whole compound as seen in example (33b). So, modifying elements

inserted between constituents of a compound produce ungrammatical structures. The noun phrase *sehoete-sa-moru* ‘carrot of the forest’ in contrast allows modification within its independent constituents. This is seen in (33c) in which two adjectives modify the noun *sehoete* ‘carrot’ and another adjective modifies the noun *moru* ‘forest’.

- (33) a. * *se-hoete se se-holo mo-ru*
 NC7-carrot 7QP NC7-big NC3-forest
 Int. ‘big wild carrot’
- b. *se-hoete mo-ru se se-holo*
 NC7-carrot NC3-forest 7QP NC7-big
 ‘big wild carrot’
- c. *se-hoete se se-holo se se-hla sa mo-ru o-mo-holo*
 NC7-carrot 7QP NC7-big 7QP NC7-yellow ASSC7 NC3-forest 3QP-NC3-big
 ‘big yellow carrot of the big forest’

In relation to movement of constituents of a compound, the Lexical Integrity was mentioned to posit that there should be no movement of the parts of a compound word. Considering the compound *sehoete-moru* ‘wild carrot’ again, if its constituents’ order is switched, it produces *moru-sehoete* which is acceptable in that it is a possible compound depending on the meaning it is intended for, which is nonetheless different from that of the ‘wild carrot’. In contrast, with phrases the constituents’ order can be switched without interfering with their complete meaning. This is shown in the sentences in (34a versus 34b) where in (34a) *sa-moru* ‘of forest’ has moved to the left, resulting in subtle change of meaning in (34b). The noun *sehoete* in both cases retains its position as the head, as it can be seen that it determines agreement patterns.

- (34) a. *Se-hoete sa mo-ru se se-holo.*
 NC7-carrot ASSC7 NC3-forest 7QP NC7-big
 ‘Big carrot of the forest’

- b. *Sa mo-ru se-hoete se se-holo.*
 ASSC7 NC3-forest NC7-carrot 7QP NC7-big
 'Of the forest, big carrot'

Creissels (2016:19) indicates that in Setswana noun phrases, the head-dependent order is not absolutely obligatory, but noun dependents preceding their head are extremely rare in spontaneous texts. This is true of Sesotho noun phrases, as it can be seen from the examples in (34a-b) that the noun dependents can precede or follow it, though the example in (34b) is not popular in speech. It can be seen that there is actually a difference in meaning between *sehoete-moru* 'wild carrot' and the examples in (34a and b), which refer to the 'carrot of the forest'.

Noun + noun compounds with a linking element are said to cut across the boundary between compounds and phrases and are referred to as associative nominals by (Kula, 2012). However, in this discussion it is argued that the noun + noun compounds with a linking element are true compounds with characteristics of words rather than phrases.

Delfitto et al. (2011) consider tests in assessing the possibility of manipulating the second noun by means of various types of modification, in order to show whether associative nominals are morphological constructions, or modifiable syntactic constructions in Cibemba. In one test, an adjective is used to modify the second noun in these kind of constructions and it is concluded that the second noun can be modified by an adjective. This is indicated in example (35) below where it is said that the adjective agrees in class with the second noun. It is pointed out that the agreement pattern shows that the adjective is modifying the second noun and not the whole associative nominal.

- (35) *ukuulu kwa nkoko iyi-kulu*
 NC5.leg ASSC5 NC9.chicken NC9-big
 'leg of a big chicken'

Another test is where a second noun is modified by a relative clause. In this case as well, the agreement in class between the second noun and the relative clause marker -yo

shows that the relative clause is modifying the second noun. This is indicated in example (36) below.

- (36) *ubwalwa bwa mataba ayo ba-shitile mailo na-bu-pwa*
 NC14.beer ASSC14 NC6.maize REL6 SM2-buy.PFV yesterday TNS-SM2-finish
 ‘the beer of the maize that they bought yesterday is finished’

Based on the above tests, Delfitto et al. (2011) conclude that Cibemba associative nominals are syntactic constructions whose internal constituents fully adhere to modification. In analysing the tests from Delfitto et al. (2011) for example, the test which produces example (35) above, it is shown that this is the case with Sesotho phrases but not with compound nouns where the second noun cannot be independently modified. It is also shown that even the first noun in Sesotho can be independently modified in phrase structures.

An example in Sesotho which is a translation of the Cibemba example in (35) above is provided in (37a), in which the adjective is modifying the second noun. In (37b) the adjective is modifying the first noun. These examples are treated as phrases in Sesotho.

- (37) a. *leoto la khoho e-kholo*
 NC5.leg ASSC5 NC9.chicken NC9-big
 ‘leg of a big chicken’
- b. *le-oto le le-holo la khoho*
 NC5-leg QP5 NC5.big ASSC5 NC9-chicken
 ‘big leg of a chicken’

Contrary to the above, the examples of noun + noun examples with a linking element as in (38) are compounds in Sesotho. The test of independent constituent modification does not work with these particular compounds.

- (38) *ø-koena ea ø-thaba*
 NC9-crocodile ASSC9 NC9-mountain
 ‘species of mint’

The compound in (38) above functions as a noun. Therefore, it is a single morphological unit which is the name of a plant. It neither allows internal modification of its parts nor their movement, in the same way as compounds without a linking element do, as shown in example (33a). The example in (39a) shows that the compound does not adhere to internal modification as its meaning changes to ‘big crocodile of the mountain’ which is a phrase rather than a plant name, while the compound in (39b) retains its meaning; the plant name when it is modified wholly, thus becomes ‘big species of mint’.

- (39) a. **ø-koena* *e-kholo ea ø-thaba*
 NC9-crocodile QP9-big ASSC9 NC9-mountain
 ‘big crocodile of the mountain’
- b. *ø-koena* *ea ø-thaba e-kholo*
 NC9-crocodile ASS9 NC9-mountain QP9-big
 ‘big species of mint’

It is seen that the compounds with a linking element have the same structure as their corresponding phrases which comprise two nouns connected by an associative marker. In distinguishing between the two, it is concluded that the compounds were initially syntactic structures which have undergone a process of lexicalisation and as a result, lost their syntactic properties. The modifier in these compounds appears outside the whole compound, just as has been seen with the noun + noun compounds without a linking element (Lieber & Štekauer, 2009). Changing the linear order of these compounds results in the formation of a phrase. The conclusion is therefore that the structure noun + associative marker + noun in Sesotho can produce both compounds and phrases. These compounds are morphological constructions with fairly fixed semantics, just like the noun + noun compounds discussed above and conform to the Lexical Integrity Principle, as a result should be categorised as true compounds as well.

Kula (2012) considers synthetic compounds as true compounds as well, indicating that they display consistent compound properties throughout their analyses. They are also regarded as true compounds in Sesotho, since they do not conform to internal modification nor change of the linear order of their constituents, and as a result, they act

in accordance with the Lexical Integrity Principle, just like the noun + noun root compounds do.

Coordinate compounds have also been discussed and they are regarded as compound nouns as well because they do not adhere to internal modification although in terms of movement of their constituents, they can be reversed and still retain the same meaning.

4.5 Conclusion

In conclusion, this section was dedicated to present Sesotho compound types and discuss their formation. The types which were discussed were root and synthetic compounds. It was observed from the discussions that the word classes: nouns, verbs (in synthetic compounds) and adjectives are those frequently or mostly used to form Sesotho compound nouns and are then followed by adverbs. All the compounds presented in the discussions in this chapter assume the noun, whether derived or not as the first word or constituent which more often than not, appear with the noun class prefix. The noun as the head of the whole compound, determines its morphological, syntactic and semantic properties. The compound patterns discussed above have been ruled out as compound nouns as compared to their corresponding phrases with regards to Lexical Integrity Principle.

CHAPTER 5: OPTIMALITY THEORY ANALYSIS

5.1 Introduction

This chapter presents an Optimality Theoretic analysis of the Sesotho compound data discussed in Chapter 4. Section 5.1 is this introductory section. Section 5.2 provides the descriptive generalisations about Sesotho compound data. Section 5.3 outlines proposed constraints to account for Sesotho compound data. Section 5.4 provides the Optimality Theoretic analyses, while section 5.5 presents the concluding remarks.

5.2 Descriptive Generalisations about Sesotho Compounds

The close examination of Sesotho compound data presented in Chapter 4 has led to the descriptive generalisations discussed in this section. In relation to compound formation, the following structures were identified in the data collected for this study:

1. Word1 + Word2 = Compound noun.
2. Word1 + Associative marker + Word2 = Compound noun.

The words combined to form compounds represent different lexical categories. For example, the patterns found were noun + noun, noun + adjective, noun + adverb and noun + associative marker + noun. The constructions following the patterns: word1 + word2 = compound noun and word1 + associative marker + word2 = compound noun are analysed as words, as opposed to phrases. That is, they are morphological constructions instead of syntactic structures.

As a result, they do not allow movement of their constituents (exceptions in this case are coordinate compounds) nor insertion by internal modification, as per Lexical Integrity, which was discussed in Chapter 4, section 4.4. The data revealed that the vast majority of Sesotho compounds are semantically left-headed. One example which was right-headed was found. The headedness phenomenon was determined on the semantic level in which three types of compounds were found; those with a semantic head referred to as endocentric compounds, those without the semantic head, referred to as exocentric compounds as well as coordinate ones which have two equal heads.

As mentioned in Chapter 1, section 1.4.1 the noun class system is central to the study of nouns in Sesotho as in other Bantu languages. Most nouns must have an overt class prefix, there are, however, some examples which take the zero prefix and their nominal class gender is determined from the agreement features.

Most compounds take plural forms and the plural affix is assigned to the leftmost element only and nothing changes on the second constituent, except in noun + adjective compounds, where the following adjective adopts the class prefix of the noun it modifies and coordinates compounds in which the two constituents are both marked for number.

The most widespread or predominant plural pattern in compounds is singular-plural pairing of count nouns in which one class takes its plural from the next class. Other compounds are non-count examples and are referred to as one class compound nouns. There are also instances of collective and contextual plural forms in compound nouns as has been seen from the previous discussions.

5.3 Proposed Constraints

This section presents constraints necessary for an account of Sesotho compound data namely: INTEGRITY which penalises structures which are not words or possible word forms, EXPRESS Ncl.(WORD1) which requires compound nouns to have overt noun class prefixes, ALIGN-LEFT (WORD, HEAD) which posits that the semantic head of a compound is the leftmost member, DEP-IO which does not allow insertion within parts of a compound, IDENT-IO (HEAD) and IDENT-IO which enforce identity between the input and the output. These constraints are defined in subsection 5.3.1.

5.3.1 Constraints definitions

1. Markedness constraint

- a) INTEGRITY (WORD): posits that the internal structure of a word cannot be interrupted by elements which are not part of that word [Anderson, 1999:22].
- b) EXPRESS Ncl. (WORD1): requires that the first constituent of a compound noun must have an overt noun class prefix [Mashauri, 2018:6].

2. Alignment constraint

- a) ALIGN-LEFT (WORD, HEAD): accounts for alignment of the semantic head to the left of the compound word. Teuber's constraint is ALIGN-LEFT (WORD, HEAD FOOT) which is modified by removing foot to account for compound word headedness

[Teuber, 2012:45] *modified*.

3. Faithfulness constraints

- a) DEP-IO: does not allow insertion.
- b) IDENT-IO: enforces input-output identity.
- c) IDENT-IO(HEAD): stipulates that the head feature of the input must correspond to the output

[McCarthy, 2008:228]

5.4 Optimality Theoretic Analyses of Sesotho Compound Data

The compounds analysed in this study were mentioned to have two types being root and synthetic compounds which take two structures: word + word = compound noun and word + associative marker + word = compound noun. The root and synthetic compounds were additionally semantically analysed, resulting in three categories being endocentric, exocentric and coordinate compounds. Firstly, endocentric root compounds following the structure word + word = compound are analysed.

Tableau 1 illustrates an analysis of an endocentric root compound *sehoete-moru* (carrot + forest) ‘wild carrot’.

/se-hoete-mo-ru/

se-hoete_H-mo-ru	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. →se-hoete-mo-ru						
b. se-hoete-sa-mo-ru	*!	*				*
c. mo-ru-se-hoete			*!	*		*
d. se-hoete-se-se-holo-mo-ru	*!	*				*
e. ∅-hoete-mo-ru					*	*

The candidate set in Tableau 1 above consists of a faithful candidate (a) *sehoete-moru* (carrot + forest) ‘wild carrot’. This candidate has incurred no change and therefore is faithful to the input, and, as a result, is analysed as optimal because it obeys all ranked constraints. The unfaithful candidate (b) *sehoete-sa-moru* (carrot + of + forest) violates INTEGRITY (WORD), DEP-IO and IDENT-IO. This candidate is actually a noun phrase which corresponds to the compound *sehoete-moru* and it introduces a new element being the associative marker, which in this case is not part of the compound.

Candidate (c) *moru-sehoete* (forest + carrot) is the reversed unfaithful candidate, which therefore violates ALIGN-LEFT (WORD, HEAD). It also violates IDENT-IO (HEAD) because the feature head is expressed on the rightmost member therefore does not correspond to the input in which the feature head is expressed on the leftmost member and there is no input-output correspondence which, as a result, violates IDENT-IO. The unfaithful candidate *sehoete se-seholo-moru* (carrot + big + forest) is a compound that is modified internally and therefore violates INTEGRITY (WORD), IDENT-IO and DEP-IO by inserting elements which are not part of a compound. The unfaithful candidate *hoete-moru* (noun stem; carrot + forest) is a compound appearing without the noun class prefix; consequently, violating EXPRESS Ncl. (WORD1) and IDENT-IO. This is a possible compound because it belongs to class 7 which is among the class prefixes which allow class prefix drop. In conclusion, the analysis in Tableau 1 does not justify any particular ranking schema since the winning candidate does not violate any of the ranked constraints.

It was mentioned in the descriptive generalisations that compound nouns have the feature number. The singular compound *sehoete-moru* ‘wild carrot’ of class 7 takes the plural form *lihoete-moru* ‘wild carrots’ in class 8, which was described as following the predominant singular-plural pairing pattern. The singular prefix is replaced by the plural prefix. Tableau 2 illustrates the formation of the compound *lihoete-moru* in plural form.

Tableau 2 corresponding to the compound *lihoete-moru* (carrots + forest) ‘wild carrots’

/li-hoete-mo-ru/

li-hoete_H-mo-ru	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS NCL. (WORD1)	IDENT-IO
a. →li-hoete-mo-ru						
b. li-hoete-tsa-mo-ru	*!	*				*
c. me-ru-se-hoete			*!	*		*
d. li-hoete-tse-kholo-mo-ru	*!	*				*

e. Ø-hoete-mo-ru					*	*
------------------	--	--	--	--	---	---

Tableau 2 above presents the same analysis as Tableau 1 in which the winning candidate does not incur any violations from the ranked constraints, thereby indicating that the analysis does not require a particular ranking. In conclusion, Tableaux 1 and 2 represent the formation of singular and plural endocentric compounds with the pattern word + word = compound, as has been indicated from the descriptive generalisations. The singular and plural forms discussed in Tableaux 1 and 2 are examples of compound nouns which were discussed to express the class prefixes which is the predominant property of nouns in general. Exceptions to this property were discussed and an example of such compounds is *fito-ntja* ‘badly made knot’ and its analysis is presented in Tableau 3 below.

Tableau 3 corresponding to the formation of the endocentric root compound *fito-ntja* (missing prefix-knot + dog) ‘badly made knot’.

/Ø-fito-Ø-ntja/

Ø-fito _H -Ø-ntja	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. → le-fito-Ø-ntja						*
b. le-fito-la-Ø-ntja	*!	*				*
c. Ø-ntja-le-fito			*!	*		*
d. le-fito-le-le-holo-Ø-ntja	*!	*				*
e. → Ø-fito-ntja					*	

Tableau 3 above illustrates the strict ranking schema to account for the endocentric root compound *fito-ntja* ‘badly made knot’: INTEGRITY (WORD)>> DEP-IO >> ALIGN-LEFT (WORD, HEAD)>> IDENT-IO (HEAD)>>EXPRESS Ncl. (WORD1)>> IDENT-IO. The ranking argument here

is that INTEGRITY (WORD) constraint is undominated since any form which violates it is not considered a compound noun or a possible one.

INTEGRITY (WORD) constraint distinguishes between word and phrase structures, by enforcing that there should not be insertion nor reversibility between members of a word or a compound noun. Non-reversibility and non-insertion points at the INTEGRITY (WORD) relationship with the constraints ALIGN-LEFT (WORD, HEAD) and DEP-IO. ALIGN-LEFT (WORD, HEAD) does not allow reversibility of members of a compound word, and, as a result, is responsible for maintaining the semantic and linear order of the constituents of a compound. DEP-IO enforces close correspondence between the input and output candidates in which it forbids the insertion of elements which are not part of the compound word.

Another constraint based on hierarchical importance which is ranked lower than ALIGN-LEFT (WORD, HEAD) is IDENT-IO (HEAD) which enforces head feature identity between the input and the output, then followed by EXPRESS Ncl. (WORD 1). Essentially this constraint enforces the importance of the noun class prefix of the left constituent of a compound word.

The last constraint that follows on the ranking hierarchy is IDENT-IO, which ensures identity between input and output. Tableau 3 above shows that both candidates (a and e) are optimal. Both candidates minimally violate the lower ranked constraints. It has been mentioned that Sesotho compounds mostly carry the class prefix like candidate (a) does. However, candidate (e) appears as an example of exceptions which appear without the class prefix, and thus are specifically from noun classes which were mentioned to allow the class prefix drop.

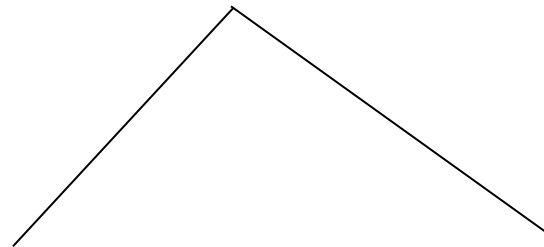
This means that a candidate without a class prefix belonging to the noun class which does not allow a class prefix drop fatally violates INTEGRITY (WORD). As a result, it is not considered a compound or a possible one. This is seen clearly in Tableau 7 in which such a candidate violates INTEGRITY (WORD). The above analysis does not sufficiently account for the formation of compounds without the class prefix as two winners' surface.

The existence of compounds without the class prefix bring forth the issue of variation, where it is observed that Sesotho on one hand prefers compounds with overt class prefixes, while on the other hand, it still allows compounds without overt class prefixes. This variation is accounted for by incorporating Cophonology within OT, which posits different rankings for the compound *fito-ntja*.

The Cophonology split adopted from Inkelas and Zoll (2007:137) relating to class prefixes inclusion in the formation of compounds is presented in example (1) below.

(1) Master ranking

INTEGRITY (WORD)>> DEP-IO >> ALIGN-LEFT (WORD, HEAD)>> IDENT-IO (HEAD) {EXPRESS NCl.
(WORD1), IDENT-IO}



Cophonology A (requires the presence of a class prefix)

INTEGRITY (WORD)>> DEP-IO >>
ALIGN-LEFT (WORD, HEAD)>>
IDENT IO (HEAD)>>

EXPRESS NCl. (WORD1)>> IDENT-IO

Cophonology B (does not require the presence of a class prefix)

INTEGRITY (WORD)>> DEP-IO >>
ALIGN-LEFT (WORD, HEAD) >>
IDENT IO (HEAD)>>

IDENT-IO >>EXPRESS NCl. (WORD1)

The interpretation of the above in example (1) is that there is a master ranking which is the partial ranking of constraints to which all Cophonologies conform to in the formation of Sesotho endocentric compound nouns, followed by a Cophonology split in relation to the class prefixes in which Cophonology A produces compound nouns with a class prefix whereas Cophonology B produces compounds without the class prefix. This is shown in Tableaux 4 and 5 below.

Tableau 4

/∅-fito-∅-ntja/

∅-fito_H-∅-ntja	EXPRESS NCl. (WORD1)	IDENT-IO
a. →le-fito-∅-ntja		*
b. ∅-fito-∅-ntja	*	

Tableau 5

/∅-fito-∅-ntja/

∅-fito_H-∅-ntja	IDENT-IO	EXPRESS NCl. (WORD1)
a. le-fito-∅-ntja	*	
b. → ∅-fito-ntja		*

With regards to pluralisation, the compound *fito-ntja* takes its plural form in class 6 as *mafito-ntja* 'badly made knots', following the singular-plural pairing pattern. Emphasis has been put on the importance of the first noun and its features in compound nouns. The pluralisation of the compound *fito-ntja* without the class prefix, serves as proof that the first noun is important as it carries a number feature. This is determined by agreement in which the post modifiers for example take class 6 prefix *ma-*. Tableau 6 presents the analysis of the plural form *mafito-ntja* which is still interpreted the same as Tableaux 1 and 2 in which candidate (a) wins and the ranking does not show any particular ranking for it.

Tableau 6 corresponding to the formation of the plural compound *mafito-ntja* (knots + dog) ‘badly made knots’

/ma-fito-∅-ntja/

ma-fito_{+H}-∅-ntja	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS NCL. (WORD1)	IDENT-IO
a. →ma-fito-∅-ntja						
b. ma-fito-a-∅-ntja	*!	*				*
c. ∅-ntja-ma-fito			*!	*		*
d. ma-fito-a-ma-holo-∅-ntja	*!	*				*
e. ∅-fito-ntja					*	*

Another example that was discussed was that of an endocentric, right-headed, noun + noun root compound. This compound is an exception to the proposition that Sesotho compound nouns are mostly left-headed. Such a compound is *mochana-pula* (nephew + rain) ‘little rain’ which has the second noun *pula* ‘rain’ as its semantic head which determines agreement. An example of agreement patterns confirming this are provided in example (2) below. This is a one class compound example; therefore, it is not easy to determine whether the first or second member would carry the plural prefix. The compound is analysed in Tableau 7.

- (2) *mo-chana* *∅-pula* *ena* *e-thib-el-a* *ba-na* *ho-bapala*
 NC1-nephew NC9-rain DEM9 SM9-hinder-APP-FV NC2-child NC15-play.INF
 ‘This little rain hinders children to play.’

Tableau 7 corresponding to the formation of the endocentric root compound *mochana-pula* (nephew + rain) ‘little rain’.

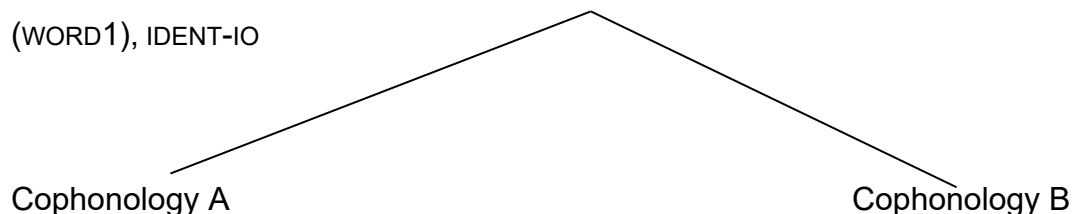
/mo-chana-∅-pula/

mo-chana.∅-pula _{+H}	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. mo-chana-∅-pula			*!			
b. mo-chana- <i>oa</i> -∅-pula	*!	*				*
c. →∅-pula-mo-chana				*		*
d. mo-chana- <i>e</i> -mo-holo-∅-pula	*!	*				*
e. ∅-chana-∅-pula	*!				*	*

The analysis in Tableau 7 provides an unwanted optimal candidate (c) in which the constituents of the compound have been reversed, resulting in the compound being left-headed. The ranking in this tableau favours the left-headed candidate as has been the case with the previous analyses. Candidate (c) obeys ALIGN-LEFT (WORD, HEAD) while the target candidate (a) fatally violates it. This kind of compound is the only one found in the data collected for this study and its existence shows that although Sesotho compounds are mostly left-headed, there might be some which are right-headed. This is yet another example of exceptionality which is best accounted for within the Cophonology approach. It is concluded that Sesotho compounds adhere to two Cophonologies. This is illustrated in example (3) below. The Cophonology A corresponds to Tableau 7 analysis in which constraints interaction favours the left-headed candidate. Cophonology B is illustrated in Tableau 8.

(3) Master ranking

INTEGRITY (WORD)>> DEP-IO >> {ALIGN-LEFT (WORD, HEAD), IDENT-IO (HEAD)}>> EXPRESS Ncl. (WORD1), IDENT-IO



(produces left-headed compounds)

INTEGRITY (WORD)>> DEP-IO >>

ALIGN-LEFT (WORD, HEAD)>>

IDENT-IO (HEAD)>>

EXPRESS Ncl. (WORD1), IDENT-IO

(produces right-headed compounds)

INTEGRITY (WORD)>> DEP-IO >>

IDENT-IO (HEAD) >>

ALIGN-LEFT (WORD, HEAD)>>

EXPRESS Ncl. (WORD1), IDENT-IO

Tableau 8

/mo-chana- \emptyset -pula/

mo-chana-\emptyset-pula	IDENT-IO (HEAD)	ALIGN-LEFT (WORD, HEAD)
a. \rightarrow mo-chana- \emptyset -pula		*
b. \emptyset -pula-mo-chana	*	

In Tableau 8, the interpretation is that candidate (a) *mochana-pula* ‘little rain’ which is right-headed, wins in this case because the head feature in the input corresponds to the head feature in the output. That is, the head noun *pula* ‘rain’ maintains its right-headed position since it has not incurred any movement whereas, the head noun *pula* in candidate (b) assumes the left position due to movement of the constituents of the input *mochana-pula*.

Other root compounds discussed were those with the pattern noun + adjective. An analysis of an endocentric noun + adjective compound *lijo-thollo* (lit. grain food) is provided in Tableau 9. The compound is morphologically plural; thus, it does not have a corresponding singular form. It can be seen from the tableau that this compound is analysed in the same way as with the compound *sehoete-moru* in Tableau 1, where candidate (a) surfaces as optimal without violating any of the constraints, hence there is no particular ranking required for the analysis of this compound.

Tableau 9 corresponding to the formation of the endocentric root compound *lijo-thollo* (food + grain) ‘grain food’

/li-jo-∅-thollo/

li-jo_{+H}-∅-thollo	INTEGRITY (WORD)	DEP-IO	ALIGN- LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS NCL. (WORD1)	IDENT-IO
a. →li-jo-∅-thollo						
b. li-jo-tse-thollo	*!	*				*
c. ∅-thollo-li-jo			*!	*	*	*
d. li-jo-tse-kholo-thollo	*!	*				*
e. ∅-jo-thollo					*	*

Another structure which was discussed is word + associative marker + word. There were no endocentric root compounds found following this structure. The examples identified from the data collected for this study are exocentric ones, an example of such is analysed in Tableau 15. Other compounds identified are synthetic compounds and a distinction was made where it was shown that there are synthetic compounds with and without the class prefix following the class prefix drop, which was found to be a property of some root compounds. The endocentric synthetic compounds with the class prefix follow the same compound formation analysis as the left-headed root compounds appearing with class prefixes discussed above. Tableau 10 illustrates the analysis of an endocentric synthetic compound noun *moletsa-phala* ‘referee’ which expresses the class prefix.

Tableau 10 corresponding to the formation of the synthetic compound *moletsa-phala* (one who whistles + whistle) ‘referee’

/mo-letsa-∅-phala/

mo-letsa _{+H} -∅-phala	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. →mo-letsa-∅-phala						
b. mo-letsa-∅-phala	*!	*				*
c. ∅-phala-mo-letsa			*!	*	*	*
d. mo-letsa-e-mo-holo-∅-phala	*!	*				*
e. ∅-letsa-phala	*!				*	*

The result shown from the analysis of the synthetic compound with a class prefix in Tableau 10 are the same as those shown in the analyses of the noun + noun compound *sehoete-moru*, noun + adjective *lijo-thollo*, in which candidate (a) is deemed optimal while candidates (b and d) both fatally violate INTEGRITY (WORD), as a result do not qualify as compounds. Candidate (c) also incurs a fatal violation from the constraint ALIGN-LEFT (WORD, HEAD) because the head has been moved from its leftmost position. Candidate (e) is ruled out by INTEGRITY(WORD) for losing its class prefix, because it cannot lose the class prefix since it belongs to class 1, which does not allow class prefix drop. This means, it is not considered a possible compound by INTEGRITY(WORD). When pluralised, the synthetic compounds in class 1 and 7 referring to humans discussed in Chapter 4, section 4.4.2.2 can both take two plural forms. Firstly, they follow the predominant singular-plural pairing pattern as shown in Tableau 11, which presents the plural form of the singular synthetic compound *moletsa-phala*. Secondly, they follow the contextual plural where the referent could be in company of others. This plural is shown in Tableau 12. Tableaux 11 and 12 present the same analysis as their corresponding singular form *moletsa-phala* analysed in Tableau 10.

Tableau 11, corresponding to the formation of the plural compound *baletsa-phala* (those who whistle + whistle) ‘referees’

/ba-letsa-∅-phala/

ba-letsa_{+H}-∅-phala	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS NCl. (WORD1)	IDENT-IO
a. →ba-letsa-∅-phala						
b. ba-letsa-ba-∅-phala	*!	*				*
c. ∅-phala-ba-letsa			*!	*	*	*
d. ba-letsa-ba-ba-holo-phala	*!	*				*
e. ∅-letsa-phala	*!				*	*

Tableau 12 corresponding to the formation of the plural compound *bo-moletsa-phala* (those who whistle + whistle) ‘referees in company of others’

/bo-mo-letsa-∅-phala/

bo-mo-letsa_{+H}-∅-phala	INTEGRITY(WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	DENT-IO (HEAD)	EXPRESS NCl. (WORD1)	IDENT-IO
a. →bo-mo-letsa-∅-phala						
b. bo-mo-letsa-ba-∅-phala	*!	*				*
c. ∅-phala-bo-mo-letsa			*!	*	*	*
d. bo-mo-letsa-ba-ba-holo-∅-phala	*!	*				*
e. ∅-letsa-phala	*!				*	*

The above analyses represent compound nouns which are semantically headed, which were mentioned to be left-headed because the leftmost element of these compounds determines the meaning of the whole compound. These compounds therefore conform to the constraint; ALIGN-LEFT (WORD, HEAD) which stipulates that the semantic head of a compound should be on the left. The semantic analyses of compound nouns does not result in headed compounds only but also those compounds which lack a semantic head as has been seen in discussions in Chapter 4, section 4.3. These are analysed next. The first exocentric compound analysed is the noun + noun compound *ntja-'me* in Tableau 13.

Tableau 13 corresponds to the formation of the compound *ntja-'me* (dog + mother) 'brother or sister'

/∅-ntja-∅-'me/

∅-ntja-∅-'me	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. →∅-ntja-∅-'me			*	*	*	
b. ∅-ntja-ea-∅-'me	*!	*	*	*	*	*
c. ∅-'me-∅-ntja			*	*	*	*
d. ∅-ntja-e-kholo-∅-'me	*!	*	*	*	*	*

The result shown from the analysis of the noun + noun exocentric is that candidate (a) wins because it incurs less violations, as compared to other candidates which are ruled out by high ranking constraints. The INTEGRITY (WORD) constraint recognises candidate (a) as a valid compound in Sesotho. The constraints ALIGN-LEFT (WORD, HEAD) and IDENT-IO (HEAD) are violated by compounds such as in (a) because they do not have the semantic head. This compound also violates the constraint EXPRESS Ncl. (WORD1) because it does not carry an overt noun class prefix.

Another exocentric compound noun analysed in tableau 14 is the noun + adjective *morena-motšo*.

Tableau 14 corresponds to the formation of the compound *morena-motšo* (chief + black) ‘Sesotho medicine’.

/mo-rena-mo-tšo/

mo-rena-mo-tšo	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	DENT-IO (HEAD)	EXPRESS NCL. (WORD1)	IDENT-IO
a. → mo-rena-mo-tšo			*	*		
b. mo-rena-e-mo-tšo	*!	*	*	*		*
c. mo-tšo- mo-rena	*!		*	*		*
d. mo-rena-e-mo-holo-mo-tšo	*!	*	*	*		*
e. ∅-rena-mo-tšo	*!		*	*	*	*

In Tableau 14, the winning candidate (a) minimally violates both ALIGN-LEFT (WORD, HEAD) and IDENT-IO (HEAD) as expected since it lacks a head and then all the remaining candidates are ruled out by the higher ranked constraint INTEGRITY (WORD). This is because candidate (b and d) have inserted elements which are not part of the input and then candidate (c) has its parts reversed resulting in an unacceptable structure in the formation of Sesotho compounds. This is because an adjective cannot appear as the head in the formation of Sesotho compound. The adjective rather follows the noun it modifies. Candidate (e) does not express the class prefix. It is not a possible compound structure because the nouns belonging to class 1 cannot lose the class prefix in the formation of compounds.

It was mentioned earlier that an exocentric compound with the structure: word1 + associative marker + word2 would be analysed in Tableau 15. This is shown below analysing the compound *bohobe-ba-linonyana* (bread + of + birds) ‘Digitaria monodactyla plant’

Tableau 15 corresponds to the formation of the compound *bohobe-ba-linonyana* (bread + of + birds) ‘Digitaria monodactyla plant’

/bo-hobe-ba-li-nonyana/

bo-hobe-ba-li-nonyana	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	IDENT-IO (HEAD)	EXPRESS Ncl. (WORD1)	IDENT-IO
a. →bo-hobe-ba-li-nonyana						
b. bo-hobe-ba-li-nonyana	*!					*
c. li-nonyana-tsa-bo-hobe			*!	*		*
d. bo-hobe-bo-bo-holo-ba-li-nonyana	*!	*				*
e. ∅-hobe-ba-li-nonyana					*	*

Tableau 15 has candidate (a) as the winner since it satisfies all the constraints. Candidate (b) is ruled out by the higher ranked INTEGRITY (WORD) constraint. It has been learned from the previous discussions that the structure word1 + associative marker + word2 produces both compounds and phrases hence the identical candidate (b). Candidate (b) is a phrase that is why it is ruled out by INTEGRITY (WORD) constraint which is violated by structures which are not compounds or possible compounds. Candidate (c) fatally violates ALIGN-LEFT (WORD, HEAD) constraint because of the reversed constituents. Candidate (d) is also ruled out by the higher ranked INTEGRITY (WORD) constraint as it is not a compound and it also inserts elements which are not part of the compound as a result violates DEP-IO. The last candidate (e) is a possible compound because class 14 nouns lose their class prefixes but it loses to candidate (a) in this case because it does not violate any constraints.

Other compounds which were discussed are coordinate compounds. These were mentioned to have two equal heads. Tableau 16 presents an analysis of an example of a coordinate compound *sebata-motho* (beast + person ‘carnibal’).

Tableau 16 corresponds to the formation of the compound *sebata-motho* (beast + person ‘carnibal’).

/se-bata-mo-tho/

se-bata_{+H}-motho_{+H}	INTEGRITY (WORD)	DEP-IO	ALIGN-LEFT (WORD, HEAD)	DENT-IO (HEAD)	EXPRESS NCL. (WORD1)	IDENT-IO
a. → se-bata-mo-tho						
b. se-bata-sa-mo-tho	*!	*				*
c. mo-tho-se-bata						*
d. se-bata-se-se-holo-mo-tho	*!	*				*
e. ∅-bata-mo-tho					*	*

Tableau 16 shows candidate (a) as the winner, it satisfies all the constraints, thus does not incur any violations. Candidates (b and d) are ruled out by INTEGRITY(WORD) as they are not acceptable as compounds. Candidate (e) is a possible compound which can appear without a class prefix but it is ruled out by EXPRESS NCL. (WORD1) and IDENT-IO. Candidate (c) which incurs the minimal violation is a possible coordinate compound as it still has the same meaning as the winning candidate because the winning candidate allows for reversibility of its constituent. This was discussed in detail in chapter 4, section 4.3. Candidate (c) is however ruled out of the competition by the lower ranked constraint IDENT-IO. Pluralisation of this compound *motho-sebata* is still analysed the same way as in Tableau 16.

5.5 Conclusion

This chapter has provided insights into constraint-based account of Sesotho compounds. This constraint-based examination of compounds revealed that there is a difference between compound nouns and similar structures being phrases emphasising the claim that the compounds discussed in Chapter 4 and in this chapter are morphological constructions.

The analysed compound nouns had to satisfy the constraints set forth, for them to qualify as word structures, based on their formational properties. Optimality Theory was initially developed in the phonology domain, and later expanded to other domains such as morphology, syntax, morphosyntax among others; and this study is an addition to the application of Optimality Theory in morphosyntax.

CHAPTER 6: CONCLUDING REMARKS

6.1 Introduction

This chapter outlines the concluding remarks drawn from the insights which the research questions brought. The recommendations for future research are also presented here. This study investigated the morphosyntactic properties of Sesotho compounds within an Optimality Theory framework.

6.2 Concluding remarks

A description of morphosyntactic properties of Sesotho compounds focusing on their combinatorial patterns, morphological structure, number, agreement, headedness and properties which distinguish compounds from similar constructions; those being their corresponding phrases was given. In addition, Optimality Theory was applied to Sesotho compound data in which the constraints which set forth the requirements for acceptable compound nouns in Sesotho were: INTEGRITY (WORD), DEP-IO, ALIGN-LEFT (WORD, HEAD), IDENT-IO (HEAD), EXPRESS Ncl. (WORD1) and IDENT-IO. The abovementioned then led to the below discussed conclusions on Sesotho compounds.

It was found that different word classes such as nouns, verbs, adjectives and adverbs were used in combining words to create compound nouns. The combination of noun + noun compounds was found to be predominant in Sesotho compounds analysed. This pattern of compounds has a large number of synthetic compound examples. This is by far the productive compounding type in Sesotho. The noun + noun root compounds were found not to be prevalent in Sesotho. This was already discussed in the literature that the noun + noun root compounds in Bantu languages are not productive.

The other combinations were found to be noun + adjective which were also root compounds examples. The synthetic type has a lot of these noun + adjective compounds as well. The last combination was the noun + adverb which also comprised synthetic compounds. It was revealed that the most compounds analysed in this study were noun headed and that they all belonged to the noun category. The morphological structures identified comprised the following: word1 + word2 = compound and word1 + associative

marker + word2. The majority of the compounds have a non-zero noun class prefix which has morphosyntactic features. Within Optimality analysis, the constraint EXPRESS Ncl. (WORD1) enforced that the compound nouns carry class prefixes, but that contention was violated by some compounds which appeared without class prefixes but were acceptable in the language. For simple nouns to lose the class prefixes, it was shown that amongst other things there should be nominal or verbal agreement. For compound nouns it was revealed that the presence of the second member in a compound allows for class prefix drop to take place as it acts as a complement to the first member.

Relative to the position of the head, Sesotho compounds were concluded to be mostly left-headed. The head of Sesotho compounds was identified on semantic grounds as the element which transfers its category and other morphosyntactic and semantic properties to the compound as a whole. The semantic analysis of headedness resulted in three categories of compound nouns, namely, endocentric, exocentric and coordinate which were all found to exist in Sesotho.

It was hard to establish the existence of right-headed compounds in Sesotho as only one example of this kind was found. So, this one example could not serve as enough evidence to conclude that such compounds exist. The constraints ALIGN-LEFT (WORD, HEAD) and IDENT-IO (HEAD) were useful for determining the position of the head in Sesotho compounds, as well as in distinguishing between the headed compounds and non-headed ones. The constraint ALIGN-LEFT (WORD, HEAD) stipulated that the semantic head of a compound should be on the left, as it has been mentioned that Sesotho compounds are predominantly left-headed. This constraint differentiated between headed and non-headed compounds in Sesotho in that it favoured candidates which were headed but was violated by those which lacked the head.

Both structures which represent compounds and phrases exist in Sesotho and their differences were accounted for within the Lexical Integrity Principle which clearly distinguished between structures which are syntactic and those which are morphological in nature. The compounds discussed in this study were therefore concluded to be morphological constructions as it was seen that they morphosyntactically behave like

simple nouns. The constraints INTEGRITY (WORD) and DEP-IO played a crucial role in distinguishing between structures which are compounds and structures which are phrases.

As has been seen from the literature review chapter, compounding in other Bantu languages was discussed. In conclusion, a table is presented below to provide a summary of the similarities and differences between Sesotho compounding and compounding in other Bantu languages.

Table 1: Similarities and differences between Sesotho compounding and other languages

Properties	Sesotho	Setswana	Sepedi	Tshivenda	Cibemba	English
ROOT COMPOUNDS	Yes	Yes	Yes	Yes	Yes	Yes
SYNTHETIC COMPOUNDS	Yes	Yes	Yes	Yes	Yes	Yes
HEAD	Left	Left	Left	Left	Left	Right
CLASS PREFIXES	Yes	Yes	Yes	Yes	Yes	No

6.3 Recommendations for Future Research

The notion of headedness brought up issues which were still not well accounted for. For examples, the issue of whether headedness exist in compounding still needs to be studied. This is brought upon the fact that some languages were shown to be mixed headed, that is, consisting of left and right-headed compounds, which then gives rise to the question of whether compounds should be analysed based on headedness.

A broader data of Sesotho compounds may clarify if there really are right-headed compounds in Sesotho, which will then answer the question of whether the head should not be positionally defined, meaning that the head position should be perceived as not fixed.

REFERENCES

- Alonso, R.T. (2011). The morphological structure of old English complex nouns. *Journal of the Spanish Association of Anglo-American Studies*, 33(1), 127–146.
- Altakhaineh, A.R.M. (2016). What is a compound? The main criteria for compoundhood. *ExELL (Explorations in English Language and Linguistics)*, 4(1), 58–86.
- Altakhaineh, A.R.M. (2018). A cross-linguistic perspective on the Right-Hand Head Rule: The rule and the exceptions. *Linguistics Vanguard*, 5(1), 1–11.
- Anderson, S.R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Anderson, S.R. (2000). Towards an optimal account of second position phenomena. In J.F. Leeuw & J. Weijer (eds). *Optimality Theory: Phonology, Syntax and Acquisition*, pp 302–323. Oxford: Oxford University Press.
- Aronoff, M., & Fudeman, K. (2011). *What is Morphology*. Malden, MA: Blackwell Publishing.
- Basciano, B., Kula, N., & Melloni, C. (2011). Modes of compounding in Bantu, Romance and Chinese. *Italian Journal of Linguistics*, 23(2), 203–249.
- Bauer, L. (1998a). When is a sequence of two nouns a compound in English? *English Language and Linguistics*, 2(1), 65–86.
- Bauer, L., & Renouf, A. (2001). A corpus-based study of compounding in English. *Journal of English Linguistics*, 29(1), 101–123.
- Bauer, L. (2003). *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- Berg, T. (2012). The cohesiveness of English and German compounds. *Mental Lexicon*, 7(1), 1–33.
- Bisetto A., & Scalise, S. (2005). Classification of compounds. *Lingue e Linguaggio*, 2(1), 319–332.
- Booij, G.E. (2005). Compounding and derivation: Evidence for construction morphology. In W. Dressler, D. Kastofsky, O. Pfeiffer & F. Rainer (eds). *Morphology and its Demarcation*, pp 109–131. Amsterdam: John Benjamins.
- Booij, G.E. (2007). *The Grammar of Words*. Oxford: Oxford University Press.

- Booij, G.E. (2009). Construction morphology and compounding. In R. Lieber & P. Štekauer (eds). *The Oxford Handbook of Compounding*, pp 201–216. Oxford: Oxford University Press.
- Bloomfield, L. (1933). *Language*. New York: Henry Holt.
- Cole, T.D. (1955). *An Introduction to Tswana Grammar*. Cape Town: Longmans, Green & Co.
- Comrie, B., & Thompson, S.A. (2007). Lexical nominalisation: In T., Shopen (ed). *Language Typology and Syntactic Description*, pp 334–381. Cambridge: Cambridge University Press.
- Creissels, D. (2014). The ‘new adjectives’ of Tswana. In R. Simone & F. Masini (eds). *Word classes: Nature, Typology and Representations*, pp 75–94. Amsterdam: John Benjamins.
- Creissels, D. (2016). Additive coordination, comitative adjunction and associative plural in Tswana. *African Languages and Linguistics*, 2(1), 11–42.
- Delfitto, D., Fiorin, G., & Kula, N. (2011). Syntactic gradients in compounding: Bemba associative nominals vs. Italian prepositional and deverbal compounds. *Rivista Di Linguistica*, 23(2), 301–325.
- Doke, C.M., & Mofokeng, S.M. (1957). *Textbook of Southern Sotho Grammar*. Cape Town: Longman.
- Downing, L.J. (2003). Compounding and tonal non-transfer in Bantu languages. *Phonology*, 20(1), 1–42.
- Dressler, W.U. (2006). Compound types. In G. Libben & G. Jarema (eds.) *The Representation and Processing of Compound Words*, pp 23–44. Oxford: Oxford University Press.
- Eberhard, D.M., Simons, G.F., & Fennig, C.D (eds). (2020). *Ethnologue: Languages of the World*. (23rd ed). SIL International.
- Fabb, N. (1998). Compounding. In A., Spencer & A.M, Zwicky (eds). *The Handbook of Morphology*, pp 151-190. London: Blackwell Publishing.
- Fábregas, A., & Masini, F. (2015). Prominence in morphology: The notion of head. *Lingue e Linguaggio*, XIV(1), 79–96.

- Fanselow, G., & Féry, C. (2002). *Ineffability in Grammar: In Resolving Conflicts in Grammar: Optimality Theory in Syntax, Morphology, and Phonology*: Special issue *Linguistische Berichte* 11: 265–307.
- Guma, S.M. (1971). *An Outline Structure of Southern Sotho*. Johannesburg: Shuter and Shooter.
- Harries, L. (1950). *A Grammar of Mwera*. Johannesburg: Witswatersrand University.
- Hoeksema, J. (1992). The head parameter in morphology and syntax. In D. Gilbers & S. Looyenga (eds.). *Language and Cognition: Yearbook of the Research Group for Linguistic Theory and Knowledge of the University of Groningen*, pp 119–132. Groningen: Universiteitsdrukkerij Groningen.
- Hlalele, B. (2005). *Sethantšo sa Sesotho*. Maseru: Longman Lesotho.
- Illonga, E. (2016). A comparative study of headedness in Ruhaya compounds. *Journal of Education, Humanities and Sciences*, 5(1), 89–100.
- Inkelas, S., & Zoll, C. (2007). Is grammar dependence real? A comparison between Cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics*, 45(1).
- Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- Kalane, M. (2005). *Morphological and Semantic Analysis of Compound Words in Sesotho*. Term paper. National University of Lesotho.
- Katamba, F. (1993). *Morphology*. London: Macmillan Press Ltd.
- Khan, M. (2013). *Word Formation in Urdu: A Linguistic Investigation of Productivity*. Doctoral thesis. Aligarh Muslim University (India).
- Kula, N. (2012). Nominal compounding and associative phrases in Bemba. In M., Brenzinger & A., Fehn (eds). *Proceedings of the 6th World Congress of African Linguistics*, pp 431–444. Cologne: Ruediger Koeppe Verlag.
- Kunene, D.P. (1961). *The Sound System of Southern Sotho*. Doctoral thesis. University of Cape Town.
- Lekhotla la Sesotho. (1984). *Sebopeho Puo 2*. Mazenod: Mazenod book center.
- Lieber, R. (2005). English word formation processes. In R., Lieber & P., Štekauer (eds). *Handbook of Word Formation*, pp 375–422. Dordrecht: Springer.

- Lieber, R., & Štekauer, P. (2009). Introduction: Status and definition of compounding. In R., Lieber & P., Štekauer (eds.). *The Oxford Handbook of Compounding*, pp 3–13. Oxford: Oxford University Press.
- Lieber, R. (2010). *Introducing Morphology*. Cambridge: Cambridge University Press.
- Lusekelo, A. (2019). Endocentric and exocentric compounds in Kiswahili. *SKASE Journal of Theoretical Linguistics*, 16(4), 22–40.
- Mabille, A., & Dieterlen, H. (2011). *Sesuto-English Dictionary*. Morija: Morija Sesuto book depot.
- Machobane, M., Moloi, F., & Demuth, K. (2007). Some restrictions on Sesotho null noun class prefixes. *South African Journal of African Languages*, 27(1), 166–180.
- Mashauri, M.A. (2018). The interface between morphology, phonology and semantics in standard Swahili compound nouns. *Utafiti*, 13(2), 1–22.
- Matthew, P.H. (1974). *Morphology: An Introduction to the Theory of Word Structure*. Cambridge: Cambridge University Press.
- McCarthy, J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics*, 38(1), 169–195.
- McCarthy, J. (2003a). *Optimality Theory in phonology: A reader*. Malden, MA & Oxford: Blackwell Publishing.
- McCarthy, J. (2003b). *Optimality Theory: An Overview*. Oxford: Oxford University Press.
- McCarthy, J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, J. (2008). *Doing Optimality Theory: Applying Theory to Data*. Malden, MA & Oxford: Blackwell Publishing.
- McCarthy, J., & Prince, A. (1993a). Generalized alignment. In G., Booij and J., Van-Marle (eds). *Yearbook of Morphology*, pp 79–154. Dordrecht: Kluwert.
- Mpasha, L.E. (2006). *The Compound Noun in Northern Sotho*. Doctoral thesis. University of Stellenbosch.
- Musehane, N.M. (2004). Root compound nouns in Tshivenda. *South African Journal of African Languages*, 24(4), 236–244.
- Musehane, N.M. (1995). *The Nominal Compound in Tshivenda*. Doctoral thesis. University of Stellenbosch.

- Nakov, P. (2013). On the interpretation of noun compounds: Syntax, semantics, entailment. *Natural Language Engineering*, 1(1), 1–40.
- Nishimura, K. (2013). *Morphophonology of Japanese Compounding*. Doctoral thesis. University of Tokyo.
- Onwukwe, C. (2018). Constraints-based explanation of nominal compounds in Ibeme Igbo. *Ebonyi Journal of Language and Literary Studies*, 1(2), 99–111.
- Orgun, C.O. (1996). *Sign-Based Morphology and Phonology with Special Attention to Optimality Theory*. Doctoral thesis. University of California.
- Plag, I. (2003). *Word formation in English*. Cambridge: Cambridge University Press.
- Pretorius, L., Viljoen, B., Pretorius, R., & Berg, A. (2008). Towards a computational morphological analysis of Setswana compounds. *Literator*, 29(1), 1–20.
- Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction to Generative Grammar*. Oxford: Blackwell Publishing.
- Ralli, A. (2013). Compounding and its locus of realization: Evidence from Greek and Turkish. *Word Structure*, 6(2), 181–200.
- Rao, R. (2015). On the phonological status of Spanish compound words. *Word Structure*, 8(1), 84–118.
- Rose, Y., & Demuth, K. (2006). Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua*, 116(7), 1112–1139.
- Sebotsa, M. (2016). *La Contribution de la Traduction a L'expansion Lexicale du Sesotho*. Doctoral thesis. Universite de Lyon.
- Scalise, S., & Fábregas, A. (2010). The head in compounding. In S. Scalise & I. Vogel (eds.). *Cross-disciplinary Issues in Compounding*, pp 109-125. Amsterdam/Philadelphia: John Benjamins.
- Schadeberg, T. C. (2003). Derivation. In D., Nurse & G., Philipson (eds). *The Bantu Languages*, pp 71-89. London & New York: Routledge.
- Spencer, A. (1991). *Morphological Theory*. Oxford: Blackwell Publishing.
- Teuber, D. (2012). *Alignment constraints in Optimality Theory: Two examples*. Accessed: <https://archives.bukkyo-u.ac.jp/rp-contents/ER/0019/ER00190L036.pdf>
- Tongco, M.D.C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research & Applications*, 5(1), 147–158.

APPENDIX 1: Sesotho Noun + Noun compound examples

Noun + Noun	Literal meaning	Ncl.	Gloss
thota-moli	high flat ground + plant used for making ropes	9	name of airport
thaba-motloang	mountain + kind of plant	9	name of a mountain
nonyana-tšepe	bird + iron	9	aeroplane
buka-ntsoe	book + word	9	dictionary
boema-fofane	standing place + aeroplanes	14	airport
seana-marena	one who takes an oath + chiefs	7	Basotho blanket
'maka-litaba	one who causes + matters	1	trouble maker
mafika-lisiu	rocks + grain baskets	6	place name
tšupa-nako	that which points + time	9	watch
pula-maliboho	one who opens + fords	9	pioneer
setima-mollo	that which extinguishes + fire	7	fire extinguisher
seila-tsatsi	one who avoids + sun	7	beautiful woman
maapara-nkoe	those who wear + leopard	6	warriors
thaha-meso	yellow crowned-bishop + dawn	9	food prepared by daughter in-law very early in the morning
sethiba-mathe	one who impedes + saliva	7	last wife in polygamy
thita-tsebe	unploughed ground (thite) + ear	9	bone behind the ear
lehlaba-phieo	one who pierces + kidney	5	traitor
seroala-nkhoana	that which carries + small bucket	7	helicopter
thabana-mokhele	small mountain + name of a person	9	name of a mountai
sebala-makhulo	one who counts + pastures	7	vagabond

thaba-bosiu	mountain + bosiu	9	name of a mountain
thaba-moorosi	mountain + name of a chief	9	name of a mountain
mophara-thetana	that which smears + knotted strings won by women (traditional)	3	type of bread
moletsa-phala	one who blows + whistle	1	referee in sports
molaola-sephethepethe	one who controls + traffic	1	traffic policeman
sebopeo-puo	structure + language	7	Sesotho grammar book
motsoala-bana	one who bears + children	1	father
molika-letšana	that which surrounds + small lake	3	the plant; Scleria Diererleni
molika-seolo	that which surrounds + ant heap	3	the rush Scirpus Burkei
molia-nyeoe	that which causes a fall + court case	3	Basotho traditional hat
moroka-pula	one who praises + rain	1	rainmaker
monya-mafi	one who defecates + sour milk	3	traditional medicine
tlali-mothoana	lightning + small person	9	lightning out of an almost cloudless sky believed to be caused by man
motsoka-sephali	one who wags + lash	1	leader
motšoara-teu	one who handles + riem used for leading the front oxen	1	leader of a span of oxen
motjoli-matšana	snipe + small lakes	3	kind of a snipe
molomo-khaba	mouth + spoon	3	African spoonbill

molisa-linotsi	herder + bees	3	lesser honeyguide
kolobe-moru	pig + forest	9	wild pig
motjeba-hlohoana	that which shakes + small head	3	kind of traditional beer
motlosa-ngaka	that which removes + doctor	3	bad excuse
sethoena-moru	that which hops + forest	7	cape robin-cha
setona-mahloana	that which widens + small eyes	7	cape white-eye
molisa-lipela	herder + rock jumpers	3	drakensburg rock jumper
selahla-marungoana	one who throws away + small spears	7	runaway warrior
morotoana-phokoana	small urine + young male goat	3	traditional medicine
mohlaba-lerumo	that which pierces + spear	3	bedding gras
sehoete-moru	carrot + forest	7	kind of root
sehamela-poli	that which milks for + goat	7	traditional medicine
monyatalane	that which defecates + unripe grain	3	traditional medicine
seshoa-bohloko	that which dies + pain	7	black nightshade
mmusa-pelo	leader + heart	3	the plant; Indigofera eriocarpa
seja-matekoane	one who eats + dagga	7	smoker of dagga
seballoa-kutloisiso	that which is read + comprehension	7	reading comprehension
moja-lefa	one who eats + inheritance	1	heir
molia-ngoetsi	that which causes a fall + daughter in law	3	kind of traditional beer
khoho-noka	chicken + river	9	common moorhen

khoho-moru	chicken + forest	9	a kind of wild fowl
thabana-liphoofu	small mountain + elands	9	name of a mountain
bosolla-hlapi	the state of wandering about + fish	14	overseas
tšoene-motho	monkey + person	9	ugly person
sehata-marikhoana	one who steps on + trousers	7	useless person
mocha-hlama	that which burns + dough	3	kind of traditional bread
sekoalla-bosulu	that which closes off + fart	7	jumpsuit
seloma-basali	that which bites + women	7	kind of insect
neano-taba	exchanging of + affair	9	information exchange
matšela-nokana	those that cross + small river	6	food from another village
tšita-baloi	that which unables + witches	9	traditional medicine
tšita-baholo	that which unables + elders	9	difficult matter
mahlabisa-lihlong	that which causes + shame	6	shameful behaviour
sekha-metsi	one who fetches + water	7	female child
mochana-pula	nephew + rain	9	little rain
'moko-taba	bone marrow + affair	3	theme
mohloa-tšepe	grass + iron	3	carrot seed grass
mohloa-mafika	grass + rocks	3	the grass; Agrostis lachnantha
setsamaea-naha	one who walks + country	7	vagabond
tsie-balimo	locust + ancestors	9	kind of locust
lefula-tšepe	one which grazes + iron	5	springbok shooter

sehlaba-thebe	that which pierces + shield	7	Lesotho national park
boea-batho	a place gone to + people	14	death
sefala-bohoho	That which scrapes + pot remains	7	a kind of star
masisa-pelo	those that move + hearts	6	grieves or sorrows
mafihlelana-khotla	those that arrived + village court	6	children of same age; peers
mofata-seliba	one who digs + a well	1	pioneer
mopheha-khang	one who cooks + quarrel	1	quarrelsome person
setla-bocha	that which comes + state of being new	7	newcomer
sefela-peloana	that which gets finished + small heart	7	short tempered person
sesita-hlohoana	that which unables + small head	7	deep sleep or torpor
molula-setulo	one who sits + chair	1	chairperson
seola-molora	that which collects + ash	7	newly emerging breast of a young girl
thaba-linoha	mountain + snakes	9	name of a mountain
leboea-bochabela	North + East	5	name of a place
leja-hlapi	one who eats + fish	5	Englishman
bochaba-tsatsi	state of rising + sun	14	direction where the sun rises from
sesotha-molala	twist + neck	7	cattle disease
sesala-hae	one who stays + home	7	man who stays at home while others go to war
moetsa-libe	one who does + sins	1	sinner
letea-tšoene	one who strikes + baboon	5	successful warrior

sebitsa-banna	that which calls + men	7	muti used to attract men
marema-tlou	those that chop + elephant	6	elephant hackers
fito-ntja	knot (lefito) + dog	5	a badly made knot
sebolaea-likokonyana	killer + insects	7	insecticide
ntja-mme	dog + mother	1a	friend
nohana-metsana	small snake + small villages	9	Earthworm
lehana-puso	one who refuses + governance	5	rebel
lenyeka-thipa	one who licks + knife	5	pauper
maapara-kobo	those who wear + blankets	6	Basotho
mantša-tlala	those that bring out + hunger	6	food
molula-qhooa	sitter + top	1	one who sits at the top
motima-hlaha	one who extinguishes + wild fire	1	problem solver
seepa-pitso	one who digs + gathering	7	chief
tika-motse	that which surrounds + village	9	traditional medicine
ngoana + tšoene	child + baboon	9	small piece of meat left over

APPENDIX 2: Sesotho Noun + Adjective compound examples

Noun + Adjective			
bitso-lebe	name (lebitso) + bad	5	name of a place
molala-motšo	neck + black	3	turtle dove
thaba-putsoa	mountain + grey	9	name of a mountain
maru-matšo	clouds + back	6	name of a dog
lijo-thollo	food + grains	10	grain food
ntloana-tšoana	small house + blackish	9	prison
morena-motšo	chief + black	3	kind of muti
thaba-chitja	mountain + round	9	name of a mountain
tona-kholo	minister (letona) + big	9	prime minister
monko-monate	smell + nice	3	scent
phakoana-tšoana	small kite + blackish	9	black shouldered kite
thabana-ntlenyana	small mountain + a bit nice	9	name of a mountain
phokojoe-khoaba	jackal + white streaks	9	kind of a jackal
maliba-matšo	deep places in the river + black	6	name of a river
phatla-nchocho	forehead + protruding	9	protruding forehead
ntlo-kholo	house + big	9	head office
motse-moholo	village + big	3	city
ngope-tšoeu	donga (lengope) + white	9	name of a place
mosali-mofubelu	woman + red	1	kind of muti
hlooho-putsoa	head + grey	9	elderly person
seema-noosi	one who stands + alone	7	loner

nko-phara	nose + flat	9	one with a flat nose
khohlo-ntšo	gorge + black	9	mane of a place
maoa-mafubelu	those that fall + red	6	kind of muti
poho-tšehla	bull + yellow	9	traditional medicine
telu-putsoa	beard + grey	9	old man
tšana-talana	small lake (letšana) + greenish	5	name of a lake
tša-kholo	lake (letša) + big	5	name of a place

APPENDIX 3: Sesotho Verb + Noun compound examples

Verb + Noun			
tima-lebone	extinguish + lamp	1a	name of a person
phutha-mahlasoa	collect + careless people	9	the plant: Aristida congesta; steek grass
thetsa-balisa	deceive + herders	9	red capped lark
cha-bosiu	burn + night	9	a blister that appears at night
thula-mabota	collide with or hit + walls	9	type of dance
thala-boliba	frolic + deep hole in a river	9	an insect:Gyrinus or Whirligig
phutha-lichaba	collect+ multitudes	9	name of a place
qhala-maboli	scatter + field mice	9	one who scatters field mice

APPENDIX 4: Sesotho Noun + Adverb compound examples

Noun + Adverb			
moja-morao	one who eats + last	1	person who eats lastly
seoa-holimo	one who falls + up	7	person who arrives unannounced
moja-pele	one who eats + first	1	person who eats firstly
sebuela-hole	that which speaks + far	7	loud speaker
sebonela-hole	that which sees + far	7	telescope
seqhomela-nkong	that which jumps + at the nose	7	insect; cricket
moshejoa-hang	one who is looked at + once	7	ugly person
mahata-mmoho	those who tread + together	6	people of the same affiliations
letsoa-ntle	one from + outside	5	foreigner

APPENDIX 5: Sesotho Noun + Associative marker + Noun compound examples

Noun + Assc + Noun			
koloi-ea-mollo	vehicle + of + fire	9	train
koena-ea-thaba	crocodile + of + mountain	9	species of mint
bolepo-ba-metsi	moss + of + water	14	water moss
boea-ba-ntja	hair + of + dog	14	Bulbostylis striatella plant
bohobe-ba-linonyana	bread + of + birds	14	Digitaria monodactyla plant
sechaba-sa-mafika	people + rocks	7	grass; Ficinia gracilis
sehoete-sa-noka	carrot + of + river	7	the plant; Convolvulus hastatus
lefito-la-molala	knot + of + neck	5	nape
leihlo-la-khomo	eye + of + cow	5	the plant; Cucurlogo plicata
leme-la-khomo	tongue (leleme) + of + cow	5	the plant; Berkheya setifera
leoto-la-khoho	foot + of + chicken	5	the plant; Xysmalobium Parviflorum
letsoai-la-balisana	salt + of + herders	5	the plant; Epilobium Hirsutum
litsi-ba-ntja	hair (bolitsi) + of + dog	14	the plant; the grass; Hordeum capense
mahlo-a-konyana	eyes + of + kid	6	the plant; Lobelia Presli
masapo-a-noha	bones + of + snake	6	a kind of bead

Guided questions for National University of Lesotho Students' interviews

This study is intended to collect Sesotho nouns. You have been approached to help the researcher to collect Sesotho words. The following questions are based on the areas which the researcher wants to find words.

Sesotho nouns

1. Can you give me Sesotho nouns? Let us start with parts of the head.

For questions 2, 3, 4 you first give examples from your home place and its surrounding areas, then move on to examples from other places in general that you may know.

2. Can you give place names?

3. Do you know any names of rivers?

4. What about mountains, do you know any names of mountains?

5. Do you know of any Sesotho traditional food?

6. Can you name Sesotho traditional household items?

7. Can you name modern household items?

8. Can you give names of Basotho blankets you know?

9. Do you know names of Basotho hats?

10. I have a list of names I would like for you to tell me in Sesotho:

Leader

Laptop

Bicycle

Aeroplane

Toothpaste

Television

Turtledove

Wristwatch

Remote control

Loudspeaker
Flash drive
Toothbrush
Handbook
Wheelchair
Airport
Railroad
Fire truck
Firefighter
Earthquake
Landline phone
Cellphone
Supermarket
Headphones
Pianist
Referee in football
Traffic police
Sunglasses
Taxpayer
Debit or credit card
Rainbow
Lightening
Textbook
Ambulance
Bookshelf

APPENDIX 7: Questionnaire for traditional doctors

Interview questions for Traditional doctors

This study is intended to collect Sesotho nouns. You have been approached to help the researcher to collect Sesotho words with regard to your practice.

1. Am I right to refer to you as a traditional doctor?
2. What does a traditional doctor do?
3. Do you produce your own traditional medicines?
4. Do you give names to such medicines?
5. Can you give the names of the medicines?
6. Can you tell me the names of the plants you regard individually as traditional remedies and the ones you use for making your own traditional medicines.
7. Do you ever use exotic plants; if you do, can you tell their names?
8. What about wild animals, do you use them for medicines? If so, can you give their names? And maybe give the names of their body parts that you use.
9. Do you know of any exotic animals?
10. Are there any other specific items that I have not mentioned that you use in your practice? If so, can you tell me their names?

APPENDIX 8: Ethical clearance



Faculty of the Humanities

15-Jun-2018

Dear Miss Mokoaleli

Ethics Clearance: **Sesotho compounding**

Principal Investigator: **Miss Maseanakoena Mokoaleli**

Department: **Linguistics and Language Practice (Bloemfontein Campus)**

APPLICATION APPROVED

With reference to your application for ethical clearance with the Faculty of the Humanities, I am pleased to inform you on behalf of the Research Ethics Committee of the faculty that you have been granted ethical clearance for your research.


Your ethical clearance number, to be used in all correspondence is: **UFS-HSD2018/0502**

This ethical clearance number is valid for research conducted from 15-Jun-2018 to 15-Jun-2019. Should you require more time to complete this research, please apply for an extension.

We request that any changes that may take place during the course of your research project be submitted to the ethics office to ensure we are kept up to date with your progress and any ethical implications that may arise.

Thank you for submitting this proposal for ethical clearance and we wish you every success with your research.

Yours Sincerely



Dr. Asta Rau
Chair: Research Ethics Committee
Faculty of the Humanities

Dekaanskantoor: Fakulteit Geesteswetenskappe
Office of the Dean: Faculty of the Humanities
T: +27(0)51 401 2240, E: humanities@ufs.ac.za
Flippiegroenewoud Building / Gebou, FGG106
205 Nelson Mandela Drive/Rylaan | Park West/Parkwes, Bloemfontein 9301 | South Africa/Suid-Afrika
P.O. Box/Posbus 339 | Bloemfontein 9300 | South Africa/Suid-Afrika | www.ufs.ac.za

