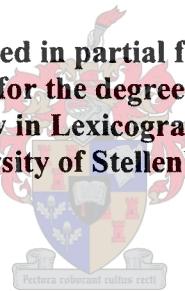


THE NOUN AND THE DICTIONARY IN TSHIVENDA

BY

SHUMANI LEONARD TSHIKOTA

**A thesis presented in partial fulfilment of the
requirements for the degree of Masters of
Philosophy in Lexicography at the
University of Stellenbosch.**



Study Leader: Prof. JA du Plessis

Date Submitted: March 2001

DECLARATION

I, the undersigned hereby declare that the work contained in this thesis is my own original and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature

Date: 8 March 2001

SUMMARY

The lemmatisation of nouns in African and Non-African language dictionaries is investigated with a view to account for the development of a theoretical framework of how nouns could be lemmatised in a Tshivenda monolingual dictionary.

Within the African language dictionaries two traditions exist for lemmatisation of nouns: the stem or root system and the word system. Dictionaries belonging to the Nguni languages enter nouns according to the letter of the root or stem and this has been followed because of the presence of a preprefix in these languages. Languages like Venda, Tsonga and Sotho usually enter nouns in the dictionary under the first letter of the prefix if present.

In African languages the morphological and syntactic category noun is exactly the same. With syntax and morphology the category noun is presented by the root or stem of the noun with prefixes added as inherent semantic and grammatical features of such nouns.

So, the exact entry in a dictionary has to be considered from the view point of the syntactic and morphological category only. Taking these considerations into account a case can be made for the treatment of nouns in dictionaries. In a Tshivenda monolingual dictionary nouns are entered as head of the noun phrases (NPs); thus indicating the syntactic operation of the noun outside the context of a sentence. In other instances the syntactic operations are indicated within the context of a sentence and this information is indicated within a dictionary implicitly or covertly by means of illustrative examples.

Nouns in these dictionaries may appear with two and or more than two arguments. Semantically, the noun **Halwa** in a Tshivenda monolingual dictionary may have two arguments: firstly it may mean “an intoxicating drink” and secondly it may refer to “a container for holding beer”.

Thus the noun in African language dictionaries, Tshivenda in particular, could be represented syntactically morphologically and semantically.

OPSOMMING

Die lemmatisering van naamwoorde in Afrika- en nie-Afrikatale woordeboeke is ondersoek om 'n teoretiese raamwerk te ontwikkel vir die lemmatisering van naamwoorde in 'n Tshivenda eentalige woordeboek.

Binne die Afrikatale woordeboek bestaan twee tradisies vir die lemmatisering van naamwoorde: die stamsisteem en die woordsisteem. Woordeboek vir die Ngunitale skryf naamwoorde in volgens die eerste letter van die stam: Dié stelsel is toegepas weens die teenwoordigheid van 'n voorprefiks in hierdie tale. Tale soos Venda, Tsonga en Sotho neem gewoonlik naamwoorde in die woordeboek op onder die eerste letter van die prefiks, as daar een is.

In die Afrikatale is die morfologiese en sintaktiese kategorie naamwoord presies dieselfde. Die naamwoord word deur die stam of wortel verteenwoordig en van prefikse vergesel ter aanduiding van bepaalde semantiese en grammatikale kenmerke.

Die presiese inskrywing in 'n woordeboek word bepaal deur die sintaksiese en morfologiese kategorie. Baie aandag moet dus gegee word aan die aard van die behandeling van naamwoorde in woordeboeke. In 'n Tshivenda eentalige woordeboek, word naamwoorde opgeneem as kern van die naamwoordstukke ter aanduiding van die sintaksiese optrede van die naamwoord buite die konteks van 'n sin. In ander gevalle word die sintaksiese optrede aangedui binne die konteks van 'n sin en hierdie inligting word in 'n woordeboek implisiet aangedui deur middel van voorbeeldmateriaal.

Naamwoorde in hierdie woordeboek kan met twee en of meer as twee argumente voorkom. Semanties kan die naamwoord [**Halwa**] in 'n Tshivenda eentalige woordeboek twee argumente hê : eerstens kan dit beteken “ 'n dronkmakende bier” en tweedens 'n blik wat bier hou”.

Dus kan die naamwoorde in Afrikataalwoordeboeke, veral Tshivenda, sintakties, morfologies en semanties voorgestel word.

ACKNOWLEDGEMENTS

In the first place, I wish to express my cordial thanks to my supervisor, Prof. J.A. du Plessis, without whose expert supervision this study would not have been a reality. To him I say: “A vha ntshileli!” Mudzimu a vhe navho vha lalame!”

Secondly, my heartfelt thanks go to Prof. R.H. Gouws, our M.Phil in Lexicography course co-ordinator. To him I say: You are wonderful.

In the third place, I am most grateful to those who helped me with the collection of data. In particular I wish to thank the following people: Mr Wilson Silidi, MA. (Stell); Mr Alfred M. Tsanwani, MA (Stell), Dr. N.M. Musehane of the Department of Tshivenda, University of Venda for Science and Technology; Mr C. Mamphwe, Mrs Nokuzola S. Tebhe- Mavhungu for Xhosa Lessons and others. To them I say: “Ndi a Livhuha”

Fourthly my thanks go to my wife, Norah and our two sons, Mukundi and Thendo for their encouragement.

Fifthly, I wish to thank in a special way, Tshifhiwa Radzilani’s Professional Printers at Tshikombani Village who typed the first manuscript and Ms E. Chauke who typed the entire thesis.

Last but not least my thanks go to the Almighty God who gave me strength to continue with my studies in times of hardships.

ACKNOWLEDGEMENT FOR NRF REPORT

The financial assistance of National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the National Research Foundation.

TABLE OF CONTENTS

1.	INTRODUCTION.....	1
1.1	AIM OF THE STUDY.....	1
1.2	METHOD.....	1
1.3	ORGANISATION.....	1 – 2
1.4	THE THEORY OF LEXICOGRAPHY.....	2 – 3
1.4.1	THE THEORY OF THE LEXICOGRAPHIC DESCRIPTION OF LANGUAGE.....	3
1.4.1.1	MAIN TYPOLOGICAL CATEGORIES.....	4
1.4.1.2	STRUCTURE OF LEXICOGRAPHICAL TEXTS.....	6
1.4.2	THE THEORY OF LEXICOGRAPHIC LANGUAGE RESEARCH.....	12 – 14
1.4.3	THE THEORY ON RELATIONSHIP BETWEEN LEXICOGRAPHY AND OTHER DISCIPLINES.....	14
1.4.4	THE THEORY OF ORGANIZATION OF LEXICOGRAPHIC WORK.....	14
1.4.5	THE THEORY ON THE PURPOSES OF DICTIONARIES.....	15
1.5	THE LEMMA.....	16
1.5.1	DIFFERENT TYPES OF LEMMATA.....	16
CHAPTER 2		
2.	THE NOUN IN THE DICTIONARY.....	18
2.1	THE NOUN IN NON-AFRICAN LANGUAGE DICTIONARIES.....	19
2.1.1	THE INCLUSION OF NOUNS IN ENGLISH AND AFRIKAANS DICTIONARIES.....	20

2.1.1.1 THE INCLUSION OF NOUNS IN ENGLISH DICTIONARIES.....	20
2.1.1.2 THE TREATMENT OF NOUNS IN ENGLISH DICTIONARIES...	21 - 26
2.1.1.3 THE INCLUSION AND TREATMENT OF NOUNS IN AFRIKAANS LANGUAGE DICTIONARIES.....	26 – 29
2.1.1.4 THE INCLUSION AND TREATMENT OF NOUNS IN ENGLISH – AFRIKAANS/ AFRIKAANS-ENGLISH TRANSLATION DICTIONARIES.....	30 – 32
2.2 THE NOUN IN AFRICAN LANGUAGE DICTIONARIES.....	32
2.2.1 THE INVESTIGATIVE ANALYSIS OF THE ENTRY AND TREATMENT OF NOUNS IN DICTIONARIES BELONGING TO AFRICAN LANGUAGES WITH SPECIFIC REFERENCE TO TSHIVENDA.....	33
2.2.1.1 NOUNS IN CLASS PREFIXES.....	33
2.2.1.2 DERIVED NOUNS.....	49
CHAPTER 3	
3.1 SYNTAX OF THE NOUN.....	56
3.1.1 THE X-BAR THEORY OF PHRASE STRUCTURE.....	56 - 64
3.1.2 THE STRUCTURE OF THE NP.....	65 – 68
3.1.3 THETA ROLE AND PREDICATE ARGUMENT STRUCUTRE.....	69 – 75
3.1.4 FUNCTIONAL POSITIONS IN A SENTENCE.....	76 – 83
3.1.5 CASE THEORY.....	84
3.1.6 MOVEMENT.....	84
3.2 MORPHOLOGY OF THE NOUN.....	86
3.2.1 NOMINAL INFLECTION.....	86
3.2.2 LEXICAL DERIVATION.....	94
3.3 SEMANTICS OF THE NOUN.....	102

CHAPTER 4

4. LEMMATIZATIONS OF NOUNS IN A TSHIVENDA DICTIONARY... 114

4.1	AIM.....	114
4.2	NOMINAL COMPONENT.....	114
4.2.1	SYNTAX OF THE NOUN.....	114 – 115
4.2.2	MORPHOLOGY OF THE NOUN.....	116
4.2.3	SEMANTICS OF THE NOUN.....	130 – 133
4.3	CONCLUSION.....	134 – 137

CHAPTER 5

5.	CONCLUSION.....	138 – 142
6.	BIBLIOGRAPHY.....	143 - 145

CHAPTER 1

1.1 INTRODUCTION

The main aim of this study is to investigate the inclusion and treatment of nouns in Tshivenda dictionaries. It will be the purpose of this study to explore the categories of nouns with regard to their treatment in non-African language dictionaries, i.e. English and Afrikaans and in African language dictionaries, i.e. Tshivenda, Xitsonga, Sepedi, Zulu, Xhosa, Setswana and Sesotho. This will help to develop a theoretical framework as to how nouns should be treated in African language dictionaries, Tshivenda in particular. This study will also cover the treatment of morphological, syntactical and semantic aspects of the nouns in African Languages, Tshivenda in particular.

1.2 METHOD

In this study various sources will be used which include books on theory of lexicography, books on morphology, syntax and semantics, books on lemmata and the lemmatisation of the nouns. Other related issues will also be used. People who have a know how of African Language Linguistics, Lexicology, Morphology, Semantics and Syntax will be consulted. This study will be limited to the analysis of nouns with special reference to Tshivenda as an African Language. Other African and Non-African Languages will also be analysed.

1.3 ORGANISATION

The thesis has been organised in the following ways:

Chapter 1 - is the introduction. It deals with the aim of the study and the research method involved in the compiling of the thesis. The organisation of the thesis is followed by the theory of Lexicography. The discussion starts from a practical point of view and tries to move towards the development of a theory of Lexicography. Certain guidelines

and principles are explored in order to establish a theoretical framework of lexicography. This is concluded by the discussion on what a lemma is and the different types of lemmata.

Chapter 2 – deals with the treatment of nouns in dictionaries. Here dictionaries belonging to the Non-African and African Languages are critically analysed. From each of these dictionaries, 5 pages are randomly selected and a critical comment on the treatment of nouns as lexical entries is made.

Chapter 3 – deals with the noun in morphology, syntax and semantics. This chapter will focus on the following:

- The noun and morphology
- The noun and syntax and
- The noun and semantics

Chapter 4 - deals with the lemmatisation of nouns in Tshivenda dictionaries. The discussion will be based on the theoretical framework on the treatment of nouns in dictionaries developed from chapter one. This chapter will focus on the following:

- Can the noun be lemmatised with its variants?
- Can the noun be lemmatised on a class prefix system?
- Can the noun be lemmatised on a stem system?
- Are natural nouns the only ones to be lemmatised?
- Are derivational nouns lemmatised?

Chapter 5 – deals with conclusion. Here a brief summary of what has been discussed in the whole thesis is given.

1.4 THEORY OF LEXICOGRAPHY

Lexicography, cf. Singh (1982:3) has been generally defined as the writing or compiling of a lexicon or dictionary: the art or practice of writing dictionaries or the science of

method of compiling dictionaries. Lexicography is not only treated as an art or practice but as a science as well. In order to compile a dictionary both the art and science are needed. As Frans Joseph Hausmann (1986:101) states: the science which is concerned with the making of the dictionary and with the finished product is called dictionary research or metalexigraphy. Based on his argument metalexigraphy can be divided into five fields of interest, namely:

- Lexicographic theory,
- Criticism of dictionaries,
- Research on dictionary use,
- Research on dictionary status and marketing and
- History of Lexicography

And the field of dictionary research which covers the scope of this study is the lexicographic theory. Lexicographic theory can be subdivided into five different areas of interest namely:

- The theory of the lexicographic description of language.
- The theory of lexicographic language research.
- The theory on the relationship between lexicography and other disciplines.
- The theory of organisation of lexicographic activities.
- The theory on the purposes of dictionaries.

Lexicography is a scientific practice aimed at the production of dictionaries so that a further practice, the cultural practice of dictionary use can be initiated. The lexicographic practice developed certain pattern and systems unique to specific types of dictionaries. And this can be regarded as the beginning of a theoretical approach to lexicography. These are criteria which were developed from a well established practice. Establishment of these patterns involves aspects regarding the structure and compilation of dictionaries and different data categories to be included.

1.4.1 The theory of the lexicographic description of language

This theory includes dictionary typology and the structure of lexicographical texts.

1.4.1.1 Main typological categories

Available dictionaries may be categorised according to a typological spectrum. And in order to determine a typological status of a particular dictionary, it must be evaluated according to criteria of a particular model. They can be classified according to clearly definable typological categories, the borders between the categories are often fuzzy. Categories overlap and dictionaries from different categories display some common features. Below is a discussion of main typological categories.

1.4.1.1.1 Linguistic dictionaries and encyclopaedic dictionaries

Linguistic dictionaries are dictionaries which focus on a linguistic treatment of lexical items. According to, Singh (1982:12) this include the lexical or linguistic information which pertains to Linguistic characteristics of the lexical unit, namely: the definition, etymology and grammar aspects. Encyclopaedic information has the following features:

- The inclusion of names of Person, Places and Literary works.
- Coverage of all branches of human knowledge and
- Extensive treatment of facts.

Encyclopaedic dictionaries are dictionaries which focus on extra linguistic materials. They contain primarily non-meaning items. Encyclopaedias can be regarded as the most typical examples of encyclopaedic dictionaries. Encyclopaedic information can be found in general linguistic dictionaries. Encyclopaedic information is included in the dictionary to assist the user in an unambiguous identification of the meaning of a given word.

1.4.1.1.2 Diachronic and synchronic dictionaries

Diachronic dictionaries, cf. Gouws (1997:12) deals with history and development of lexical items and their meanings. These dictionaries can also be divided into historical and etymological dictionaries. Historical dictionaries focus on the changes in the form and meaning of a word in a given period. They are known by chronological ordering of certain entries in the dictionary article, i.e. the senses of a polysemous word and the selection has to be representative of a fixed period.

Etymological dictionaries focus on the origin of lexical items and can be subdivided into restrictive and general dictionaries. Etymological dictionary present the origin of words by tracing the present day words to their oldest forms.

Synchronic dictionaries deal with the description of a language at a given stage of its development without being concerned with the historical changes that occur. Synchronic dictionaries do not necessarily treat contemporary language. Synchronic dictionaries include the public at large. They are general descriptive and translation dictionaries.

The distinction between the general and the restricted synchronic dictionaries is due to the selection of items. The restricted synchronic dictionary selects items which usually represent a well defined subsection of the lexicon of the given language. This category embraces the following example dictionaries: a psychology dictionary, a dictionary of rugby terminology and a philately dictionary. The other criteria which could determine the compiling of the restricted dictionary would be the language of a social grouping which results in a slang dictionary. The general synchronic dictionary selects in full items representing the full lexicon and this results in three subcategories of general synchronic dictionaries, namely: comprehensive dictionaries, standard dictionaries and pedagogical dictionaries.

1.4.1.1.3 Comprehensive dictionaries

A comprehensive dictionary, cf. Gouws (1997:13) endeavours to include the most representative selection of lexical items. This dictionary is also known as an overall descriptive dictionary and it deals with the selection of lexical items and the treatment of these items in a comprehensive way. The treatment of these items is aimed at a comprehensive transfer of information categories and within each one of these categories the treatment has to be extensive. Comprehensive dictionaries are well known for their inclusion of non-standard form items, dialectal items, items with a limited use and the vocabulary of a different social group. It reflects the real language use and provide the vocabulary which is relevant for general communication. Because the dictionary is comprehensive, it includes a substantial amount of encyclopaedic information. And

because they present information on historical development of the target language, they have a diachronic approach. Comprehensive dictionaries have an informative function instead of a normative function.

1.4.1.1.4 Standard dictionaries

Standard dictionaries have a more limited scope and extent and are usually single volume dictionaries. This category includes descriptive, bilingual and multilingual dictionaries. Their aim is to reflect the standard variety of the target language(s) and that is why they are normative in their approach.

1.4.1.1.5 Pedagogical dictionaries

Pedagogical dictionaries cf. Gouws(1997:15) can be divided into two subcategories, i.e. school and learner's dictionaries. School dictionaries help dictionary users who seek information regarding their first language and provide assistance to scholars. The learner's dictionaries are meant to assist foreign language learners. They differ from other dictionaries with regard to content, presentation and geographical procedures. Learner's dictionaries are known for including only those words with an attested high usage frequency in general communication and showing an interaction between components of an article.

1.4.1.2 Structure of Lexicographical texts

Lexicographical texts are organised into a particular structure.

1.4.1.2.1 The major components of lexicographical texts

With Lexicographic practice in place, the format and presentation of dictionaries were conventionalised. The process which took place before theoretical lexicography, cf. Gouws(1997), focussed attention on the aspects regarding the content of dictionaries and the form of the text of the source of reference. To be familiar with the editorial system of the dictionary one needs to take into consideration the sections which constitute the major

components of a dictionary, i.e. the front matter, the central list and the back matter. The central list is arranged alphabetically.

1.4.1.2.2 The structural component of lexicographical texts

The structure of lexicographical texts can be divided into two major categories, i.e. the macrostructure and microstructure. The macrostructure of the lexicographical texts is the collection of lexical items to be included and to be treated in the dictionary. The microstructure of the lexicographical texts includes the different data categories presented in the treatment of lexical items.

1.4.1.2.2.1 The macrostructural component

1.4.1.2.2.1.1 The nature

The macrostructural component is the collection of lexical items included as part of the central list. It contains a list of words but it may not be defined as a collection of words.

1.4.1.2.2.1.2 Macrostructural ordering

Macrostructural ordering can be divided into two types, namely: the vertical (alphabetical) ordering and the horizontal ordering. The vertical ordering and the strict alphabetical arrangement of lemma may be ordered before or after the sinuous lemma file. The horizontal ordering takes place in the sinuous lemma file. The lemmas which are horizontally arranged in the sinuous file are components of either a lemma niche or a lemma nest. And they are known as niched or nested lemmata respectively. Niched and nested lemmata are often found in an article cluster introduced by a lemma which is part of the vertical ordering of the macrostructure. Niched lemmatisation maintains a strict alphabetical arrangement within the niche and with regard to the surrounding main lemmata and does not necessarily display a semantic relation between the lemmata in the niche. Nested lemmata do not maintain a strict alphabetical ordering in the nest.

1.4.1.2.2.1.3 The selection of macrostructural elements

The macrostructural elements, cf. Geeraerts (1984: 42-46), are selected on the basis of the following criteria.

1.4.1.2.2.1.3.1 Geographical criteria

The geographical criteria determine a dictionary written for geographical differentiated sub languages. This involves dialect in the narrowest sense of the word and regional variants in the broadest sense of the word.

1.4.1.2.2.1.3.2 Sociolinguistic criteria

The sociolinguistic criteria determines the social difference leading to linguistic differentiation dictionaries describing the sub-language of particular ethnic group, a socially marginal group and the different kinds of social classes.

1.4.1.2.2.1.3.3 Temporal criteria

Temporary criteria involve the compiling of contemporary dictionaries and historical dictionaries. A contemporary dictionary is a dictionary of new words. A historical dictionary is a dictionary of different periods.

1.4.1.2.2.1.3.4 Interlingual criteria

Interlingual criteria came into place where a dictionary treat one or more languages though its macrostructure may not be multilingual.

1.4.1.2.2.1.3.5 Etymological criteria

Etymological criteria relate to the linguistic origin of words or their expression in a particular language.

1.4.1.2.2.1.3.6 Formal and grammatical criteria

These criteria involve the spelling of words and communication. Formal criteria include spelling of words or their pronunciation. Grammatical criteria may include the grammatical category of the lexical items.

1.4.1.2.2.1.3.7 Stylistic criteria

Stylistic criteria deal with stylistic values of lexical items, such as colloquialism and archaisms.

1.4.1.2.2.1.3.8 Pedagogical criteria

Pedagogical criteria deal with lexical items which express what a particular group of people do not know. It is intended for basic vocabulary or for children of varying ages. It contains items acquired by the potential public.

1.4.1.2.2.1.3.9 Semantic and encyclopaedical criteria

Semantic criteria deal with items which are selected on the basis of the subject they refer to. An example could be a dictionary of terminologies relating to one particular subject. Encyclopaedical criteria deal with systematical inclusion of individually referring expressions. Their macrostructure will consist mainly of nouns with relatively few verbs and will systematically include items that refer to individual persons, places, facts etc.

1.4.1.2.2.1.3.10 Idiolectical criteria

Idiolectical criteria are restricted to vocabulary of one particular person or restricted set of texts. The concordances and biblical dictionaries are examples of this category.

1.4.1.2.2.1.3.11 Frequency criteria

Frequency criteria determine the most or the least frequently used items in different dictionaries.

1.4.1.2.2 The microstructural component

1.4.1.2.2.1 The nature

The microstructural component is comprised by the different data categories presented in the treatment of lexical items. These information categories deal with semantic information and pragmatic information. Semantic information is constituted by the description of meaning and semantic relations. Lexicographical definitions, polysemy and homonymy fall under the description of meaning while semantic inclusion and semantic opposition fall under semantic relations. Lexicographical labels, examples and citations or quotations fall under pragmatic information. Lexicographical definitions, c.f. Zgusta (1971: 257-258) are guided by four principles, namely:

- All words used in a definition should be described elsewhere in that particular dictionary.
- The definitions should not contain words that are more difficult to understand than the definiendum itself.
- The definiendum itself or derivations or combinations there of may not be used in the definiens, unless they are defined separately.
- The definiens should correspond to the lexical category of the definiendum where possible.

Lexicographical definitions can be divided into two major groups, namely descriptive definitions and synonymy definitions. Descriptive definitions include three sub-groups, namely: generally descriptive definitions, genus differentia definitions and circular definitions.

1.4.1.2.2.2 Presentation of microstructural elements

In the presentation of microstructural elements dictionaries do not follow a fixed order with regard to the presentation of these types of information. Dictionaries concentrate mainly on the order of the different types of data. As Geeraerts (1984:51) states that the most important structural principles involved in the presentation of the microstructural elements has to do with the order of the various meanings. Theoretically this may be chronological (oldest meaning first followed by the figurative extensions) or based on the frequency of the sense in question (most common meaning first) in practice these principles will often be combined. In presenting the headword it may not be followed immediately by the list of meanings, but by grammatical, phonetic, etymological or other data. Dictionaries have a common practice of placing collocations, examples, quotations and so on under the definition of the sense(s) they illustrate. The microstructure, cf. Geeraerts (1984:52) may incorporate the supralexicical elements of the lexicon, functioning as syntagmatic information with regard to the alphabetical entry they are part of, on the other hand it may function as a macrostructural indication that the lexical stock of the language contains elements that are supralexicical, i.e. that are longer than single words.

1.4.1.2.2.2.3 The selection of the microstructural elements

The microstructural elements are selected on the basis of the macrostructural criteria discussed above at 1.4.1.2.2.1.3 where several types of data are listed. (cf. Geeraerts 1984: 48-49).

1.4.1.2.2.2.3.1 Orthographic data

This data specify spelling of the lexical items, when compiling dictionaries lexical items need to have uniform or adequate spelling or spelling variants.

1.4.1.2.2.2.3.2 Phonetic data

Phonetic data specify pronunciation. Lexicographers need to choose an adequate system of phonetic writing which will determine the correct pronunciation variant to be included in a particular dictionary.

1.4.1.2.2.2.3.3 Grammatical data

Grammatical data cf. Geeraerts (1984:48), specify macrostructural items by grammatical category or subcategory of the word or expression or morphological properties with regard to case and inflection on the one hand and derivation and compounding on the other hand and syntagmatic data with regard to the collocations and idioms on the one hand, and the syntagmatic patterns on the other hand in which the items appear or may appear.

1.4.1.2.2.2.3.4 Stylistic data

Stylistic data deal with the use of common labels to indicate stylistic values of items. Dictionaries are concerned with euphemistic, formal, humorous, colloquial, pompous and obsolete data.

1.4.1.2.2.2.3.5 Distributional data

This data deal with the distribution of items which belong to one or more particular socio-linguistically differentiated sub-languages. Textually they indicate when an item occurs in a text or corpus and its frequency.

1.4.1.2.2.2.3.6 Etymological data

Etymological data relate to the origins of word forms or word meanings. The word form involves borrowing, purism and neologisms. The word meaning involve historical distribution and semantic change.

1.4.1.2.2.2.3.7 Interlingual data

Interlingual data involve translations given in multilingual dictionaries.

1.4.1.2.2.2.3.8 Illustrative data

Illustrative data indicate illustrations in the dictionary which may be verbal or non-verbal. The verbal illustrations include examples and quotations or citations whereas non-verbal illustrations include pictures, charts and fables. Illustration, cf. Feinauer (1997:7) are used when the verbal definiens does not succeed in defining the definiendum unambiguously.

1.4.1.2.2.2.3.9 Semantic data

Semantic data covers three major subcategories. First they deal with definition of meanings which cover analytical, synonym, morphosemantic, linguistic and encyclopaedic aspects. Secondly, they deal with relations between meaning which embraces metonymy, metaphor, generalization and specification. These help to give translation between the primary literal senses and the derived figurative senses. And lastly deal with relations between lexical items rather than relations between senses with one and the same lexical item.

1.4.2 The theory of Lexicographic language research

The theory of lexicographic language research deals with gathering and processing data. The most important aspects of this theory are data collection and data processing which involve computer assistance. Data collection cf. Chandor (1985:117) is the process of capturing raw data for use within a computer system. It is also known as data gathering and it is related to file creation. In modern business systems the term is often used to imply the capture of information at the instant of a transaction occurring. The theory

deals with how to compile a dictionary base, the collection, the composition, representativity, function and typology of lexicographical corpora relative to dictionary types and concerns with the role played by secondary sources.

There are basically, cf. Schumacher et al (1993:40) six ways to collect data: observations, questionnaire, interviews, documents, tests and unobtrusive measures. In lexicography the task of collecting all the important information for every existing word is through several approaches. One is to hand-build a lexicon specifying only those features that a given system needs and using only the lexical items that are most likely to occur. Data collection could be stored in machine. The assistance of computer is needed. Before storing data it must be selected based on two important things which are, time and the subject or theme. Data collection will also depend on the type of dictionary to be compiled. If a particular dictionary needs grammatical information and if it will need etymological information the data collected must include both grammatical and etymological information. Geeraerts (1991:95) states the following with regards to the selection of data: theoretical lexicography will have to specify how these choices can be rationally justified on the basis of theoretical criteria with regard to the definition of data, and pragmatic criteria with regard to the functional purpose of their incorporation. Data processing is a theory about ways of processing the linguistic data collected, so that a dictionary file suitable for a particular dictionary type or a group of dictionary types is established. In both components the role of the computer must be considered.

Data processing cf. Chandor (1985:123) is the operations performed on data usually by automatic electronic equipment, in order to derive information or to achieve order among files. A data processing system may incorporate clerical functions and ancillary machine operations as well as arithmetic and logical operations performed by a computer. Since the processing of data is usually done by a machine, certain information must be fed to the processor. The lexicon, cf. O'Grady (1996: 647) needs to know about the kinds of structures in which words can appear, about the semantics of surrounding words and about the styles of the text. O'Grady (196:650) provided another example of the way large corpora are useful, see *flounce* from Webster's Seventh New Collegiate Dictionary.

(1) Flounce 1

- to move with exaggerated jerky motions
- to go with sudden determination
- to trim with flounces

The definitions tell nothing about likely subjects. Looking at the corpus data it will yield this information. In the dictionaries 20 occurrences of the verb **flounce** were extracted. Thirteen had subjects that were female, four had subjects that were clothing, one had horses as the subject and two had subjects that were pronouns. The point is that given a good parser, it would be possible to extract automatically all the subjects of a given noun and then look for properties of these subjects.

1.4.3 The theory on relationship between Lexicography and other disciplines

Lexicography can relate to society, other theories and history of lexicography. The establishment of the theory of lexicography must take into consideration how lexicography relate to these other theories or constituent theories such as lexicology, linguistics and grammar. Principles that have been followed in lexicography up to now are established to connect lexicography with the history of lexicography, which principles have been valid for which dictionary types in the past and why, and which principles could apply in future for new types of dictionaries as well (cf. Wiegand 1983:16).

1.4.4 The theory of organization of lexicographic work

Depending on whether it is an alphabetical or a thematic dictionary, cf. Hartmann (1983:13) the work is organised in three fields of activity:

- all the activities leading up to the drawing up of a dictionary plan,
- all the activities involved in establishing a dictionary base and in processing this base in a lexicographical file, and
- all the activities concerned directly with the writing of dictionary texts and thus with the writing of the dictionary.

The dictionary plan includes the written plan of the dictionary in all its aspects and the dictionary base includes the complete linguistic material forming the empirical basis for the production of a language dictionary and it includes the lexicographical corpus as the set of all the primary sources. The lexicographical file is a collection of quotations for potential lemma-signs compiled from the dictionary base.

1.4.5 The theory on the purpose of dictionaries

According to R.R.K. Hartmann (1983:11) dictionaries are written to help users extend the knowledge of the mother tongue, learn foreign language, play word games, compose a report, read and decode the first language texts. He further states that general dictionary serve two purposes, they are either academic or normative or rather referential or overall descriptive. The general dictionary presents the language as it is expected to be and stop if from decay. The referential or overall descriptive dictionary does not have any normative functions but contains words of regionals, social and stylistic variations. A reference dictionary, cf. Hartmann (1983:26) is one behind which does not lie any unified language consciousness. The collected words may belong to heterogeneous speech groups of different periods and which do not in the least form a speech. From the point of view of the coverage of languages, dictionaries can be monolingual and multilingual. Monolingual dictionary, Geeraerts (1991:295), is a dictionary which contains only one object language and in which the explanatory metalanguage is not distinct from the object language. According to Geeraerts (1991:293) the prototypical dictionary, in particular is monolingual rather than multilingual, provides linguistic rather than encyclopaedic information, contains primarily semasiological rather than onomasiological or non-semantic data, gives a description of a standard language rather than restricted or marked language varieties and serves a pedagogical purpose rather than a critical or scholarly one.

As for Wiegand (1977b, 1978, 1983e, 1983f, 1983h) the general purpose for monolingual, bilingual and multilingual language dictionaries are derived from the communicated and cognitive needs of the society; or possibly goals set to stimulate needs.

1.5 The Lemma

The lexical items of a language, cf. Gouws (1991:74) are regarded as potential lemmas; however not all words qualify as lexical items. Words that occur only as components of multiword lexical items should not themselves be regarded as lexical items. Before selecting lemmata for a specific dictionary the lexicographer must specify the linguistic criteria according to which lexical items are identified. These criteria will determine the choice of lemmas.

1.5.1 Different types of lemmata

Different types of lexical items are treated in dictionaries as different types of lemmata. Words are treated as lexical lemmata, subword lexical items as sublexical lemmata and multiword lexical items as multilexical lemmata.

1.5.1.1 Lemmata with a limited lexicographic treatment

Dictionaries do not treat many of the lemmata which fall under this category. These are the variant lemmata, synonym lemmata and self explanatory lemmata.

1.5.1.1.1 Variant lemmata

The variant lemmata includes the variant of a lexical item. The treatment of variant lemmata is given in the article of one of these lemmata. The other lemmata which have a limited lexicographic treatment would be indicated by a cross-reference to the lemma where the full treatment is given.

1.5.1.1.2 Synonym lemmata

Synonym lemmata have a limited treatment in today's dictionaries. Synonym lemmata may be provided as word list. Only synonym with the higher usage frequency will receive a comprehensive treatment and synonym with lesser frequency will get cross-reference. And the synonym with the higher usage frequency will be written first.

1.5.1.1.3 Self explanatory lemmata

Self explanatory lemmata appear grouped together and have internal alphabetical ordering. The dictionary user should be able to retrieve the meaning by looking at the meaning of its components.

1.5.1.2 Lemmata with a complete lexicographic treatment

There are three types of lemmata: lexical lemmata, sublexical lemmata and multilexical lemmata.

1.5.1.2.1 Lexical lemmata

The inclusion of lexical lemmata is determined by the typological criteria. These are words including the simplex and complex lexical lemmata.

1.5.1.2.2 Sublexical lemmata

The sublexical lemmata of cf. Gouws (1997:3) are often included in a dictionary in a haphazard and an arbitrary way. These include the occurrence of stems and affixes. Sublexical lemmata form part of a complete lexical item but cannot function independently.

1.5.1.2.3 Multilexical lemmata

Multilexical lemmata, cf. Gouws (1991a:77) are single lexical items consisting of more than one word. This category is constituted by certain loanwords, idioms and fixed expressions and particle verbs. Loanword groups include words derived or borrowed from other languages without any change in form and pronunciation. An idiom, cf. Procter (1978:xxvi), is a fixed group of words with a special meaning that cannot be guessed from the combination of the actual words used. And multiword particle verbs, cf. Gouws (1991a:80), are single lexical items that should be treated as multilexical lemmata.

CHAPTER 2

2. THE NOUN IN THE DICTIONARY

This chapter looks at an analysis of nouns in selected dictionaries. The aim of this discussion is to analyse the entry and treatment of nouns in these dictionaries. The dictionaries to be analysed belong to African and non-African languages. The non-African languages dictionaries to be analysed are those that belong to English and Afrikaans respectively. In English the following dictionaries will be analysed:

- Collins COBUILD English Dictionary (COBUILD)
 - Longman Dictionary of contemporary English (LDOCE)
- In Afrikaans, the following dictionaries will be analysed:
- Verklarende Handwoordeboek van Afrikaanse Taal (HAT)
 - Basis Woordeboek van Afrikaans (BWA), and
 - The Nasionale woordeboek (NW)

For the discussion involving translation dictionaries, the following dictionaries will be used:

- D.B. Bosman's Tweetalige woordeboek (TW)
- L.C. Eksteen's Major Dictionary/Groot woordeboek (GW),

The African language dictionaries to be analysed cover the dictionaries belonging to Sepedi, Sesotho, Setswana, Tshivenda, Isizulu, Isixhosa and Xitsonga. Since many of the African languages do not have monolingual dictionaries the discussion here will be based on dictionaries which are mainly of the category of Translation dictionaries, namely:

- R. Cuenod's Tsonga – English Dictionary (TE)
- P.J. Wentzel et al's Improved Trilingual Venda Dictionary (VAE)
- N.J. van Warmelo et al's Venda Dictionary (VD)
- C.M. Doke et al's Zulu – English Dictionary (ZE)
- T.J. Kriel et al's Northern Sotho – Afrikaans Pukuntsu Woordboek (NSA)

- T.J. Kriel et al's Pharos Popular Northern Sotho Dictionary (PNS)
- J.W. Snyman et al's Setswana – English – Afrikaans Dictionary (SEA)
- H.W. Pahl et al's Greater Dictionary of Xhosa (GDX)
- R.A. Paroz's Southern Sotho – English Dictionary (SSE)

Each of the above dictionaries will be analysed on the basis of the approach it has used in lemmatising nouns. The main traditional approaches followed by the African language dictionary are the word tradition; i.e. lemmatizing nouns by the first letter of the prefix and the stem tradition; i.e. lemmatizing nouns by the first letter of the stem. The African language dictionaries which used the word tradition are the Tsonga, Sotho and the Venda and the stem tradition is used by the Nguni languages respectively.

2.1 The noun in non-African Language Dictionaries

A noun as described by Encyclopaedia Americana, volume 20 (1992:487) is the name of a person, place or thing and it may be distinguished from other parts of speech by inflection; the formal modification of the end of the word. The Word Book Encyclopaedias, N-O, volume 14, (1987:431; 1991:559) divide nouns into four identifiable groups:

- (2) the proper noun,
the collective noun,
the common noun

The World Book Encyclopaedias (1987:431; 1991:559) define them as:

[...] particular names of titles...; [...] group of person, animal or things...;

[...] subclass of nouns that stand for concepts or ideas that are not observable or measurable... and [...] all the nouns that are not proper.

2.1.1 The entry of nouns in English and Afrikaans dictionaries

Nouns in the macrostructure of English and Afrikaans dictionaries have been entered either vertically or horizontally. The horizontal arrangement of nouns took place in the sinuous lemma file while the vertical arrangement is the dictionary's central list.

2.1.1.1 The entry of nouns in English dictionaries

Nouns in English dictionaries have been entered according to one of the two different principles: Word – by- word or letter – by – letter.

Methods of alphabetization.

Letter-by-letter	Word-by-word
(3) ear	ear
earache	earache
ear drum	ear drum
eared	eared
earful	earful
earliest	ear/be
	earmark
earlobe	ear phone
earlyon	ear piece
early warning system	ear ring
ear mark	earful
ear muff	earliest
earnest	earlyon
ear phone	early warning system
ear piece	earnest
ear plug	
ear ring	

LDOCE (1978: 345-346)

Most of the English and Afrikaans language dictionaries have used the letter – by – letter principle in entering nouns and others used the word-by-word principle in entering compound nouns within article structure of the main lemma. The advantage of the letter-by-letter principle is that users find the lexical item in its own alphabetical place, but in the word-by-word principle users struggle to find a lexical item in a densely populated structure. In the above given two principles, the lexical item [earring] is entered number sixteen in a letter-by-letter principle but it is number eleven in a word-by-word-principle. The advantage of a word-by-word principle is that lexical items are entered together according to the morphosemantic relations. The noun [earring] is grouped together with other lexical items which are not interested in the morphosemantic relatedness of lexical item. The letter-by-letter arrangement of nouns enhance communicative equivalence.

(4) pass n.1.[c]

passable

passage n

passage way

pass away

pass book

LDOCE (1978:792)

The dictionary users who consult LDOCE will not find it difficult to look for nouns in dictionaries using a letter-by-letter principle.

2.1.1.2 The treatment of nouns in English dictionaries

The English dictionaries have used two approaches to enter nouns and other nominal items. Firstly the nominal entries such as inflected nouns belonging to other parts of speech and sense related nouns have been entered together with the main lemmas.

(5) Child/t Saild

/children /s Sildren/

L DOCE (1978:1)

The Longman Dictionary of Contemporary English (LDOCE) has treated the inflection of the noun together with the main lemma. But very few of this type of entry exist in the LDOCE. The noun [child] is a singular form of the lexical item [children]. In LDOCE the lemma [children] is explained in the article of the lemma [child]. It means that LDOCE has only treated the singular forms of nouns in its macrostructural level. According to D.J. Prinsloo et al (1994(4):13):

Such rules are in themselves not user friendly at all and dictionary users are known not to consult them. But in forcing the user to look for plural forms under their singular forms, precious space is saved which can be utilized for other entries.

The disadvantages of this type of entry is that it confuses dictionary users:

(6) Cobbler/ /n. 1. a person who earns money by repairing shoes
z. derog. A careless workman.

Cobblers/ // /n. Br E S1 [u] foolish talk; nonsense
LDOCE (1978:202)

What could happen here is that dictionary users could easily confuse the noun cobblers (for speech) as being the plural of the noun cobbler (for person). Only those that are trained dictionary users will be able to make this distinction on the basis that the LDOCE dictionary has only treated the singular form of the lexical items in its macrostructure.

Except for the irregular plurals the LDOCE dictionaries did not enter any other nominal plural in its macrostructure. This approach is different to the one followed by the Collins COBUILD English Dictionary (Cobuild). The COBUILD dictionary, although it did not enter plural nouns as lexical items did enter them as part of the article structure of the lemma.

- (7) Phone – in, phone – ins...
Photographer / / photographers...

Photon / / photons...

COBUILD (1995:1238)

The plural nouns in the COBUILD dictionary are explained in the entry of the main lemma. The plural nouns cannot be looked for in the alphabetical arrangement of lemmas, except for nouns which in themselves are plural and are indicated as lemmas. This is the same thing with regard to the treatment of derived nouns: those derived from loanwords and those derived from other parts of speech. Nouns derived from loanwords have been treated as lexical items in their own alphabetical places in both the LDOCE and COBUILD dictionaries. This type of treatment of borrowed nouns does not confuse potential dictionary users. The only problem which dictionary users could experience is when dealing with those borrowed nouns which have their phonological forms completely integrated in the new form and the dictionary no longer indicate the language of origin.

- (8) Vox populi/ /lat public opinion
LDOCE (1978:1232)

The LDOCE dictionary have treated the lexical items [cock] as five homonymous entries, each with several polysemous senses. But the COBUILD dictionary (1995:301) treated only 8 (polysemous) senses in one article. This is because COBUILD does not include homonyms as separate lemmas.

The COBUILD and LDOCE dictionaries differ with regard to the treatment of nouns derived from other parts of speech.

- (9) resentment/ /resentments...
reservation/ /reservation...
COBUILD (1995: 1411)

fluctuate/ /v. to rise and fall. to change from
.....action / / n.[u.c (in): there's been some
fluctuation in the rate of her HEARTBEATS.

LDOCE (1978:424)

The COBUILD dictionary have treated the noun [resentment] (deverbative) as a lexical item. But the noun [fluctuation] in the LDOCE has not been treated as a lemma but as an entry within the article structure of the lemma. The noun [fluctuation] could have been treated as lexical item since it does have own phonetic transcription. The approach followed by the LDOCE dictionary is to treat deverbative nouns under the entry of the verb.

The two dictionaries, i.e. the COBUILD and LDOCE have treated nouns with variant spelling in the same way.

- (10) Caiman, cayman/ /.....
LDOCE 1978:141
- Chirp/ /also chirrup...
LDOCE 1978:180
- Gearshift/ /gearshifts; also spelled gear shift...
COBUILD (1995:700)

The approach followed by LDOCE and Cobuild dictionary is to enter nouns with variant spellings together with their variant forms, (1) in the same place with the lemma, (2) indicated by words such as 'other forms of cross-reference.' What is interesting about words with variant spelling is that explanation is given in the entry of the lexical item with high usage frequency. Most of the variant spelling are brought into the LDOCE and COBUILD dictionaries because of the regional English language. Another form of variant spelling in COBUILD dictionary is brought up because of the style of treating compound nouns.

The COBUILD dictionary spelled compound nouns in two ways; (1) as one word (e.g. gearshift) and (2) as two separate words (e.g. gear shift) written apart. The compound nouns of the word [gearshift] may be indicated in the following three ways:

- (11) Gear shift
- Gearshift
- Gear – shift

The point which is interesting is which one has been lematized and if all have been lemmatized how have they been ordered? Can all the three forms be lemmatized in one language dictionary? The answer is definitely, no.

Both the LDOCE and COBUILD dictionaries have treated nouns homonymously and polysemously respectively. Nouns with this type of information are difficult to handle. What the LDOCE dictionary takes as homonym can as well be treated differently by the COBUILD dictionary. This also applies to the treatment of polysemous nouns.

- (12) Cock/ / n. /a. a fully grown male chicken 2 8.
- Cock² v.
- Cock³ n. / the act of cocking (2.3) 2.9
- slight slope; TILT
- Cock⁴ n. small pile of HAY...
- Cock⁵ v.

LDOCE (1978:202)

The LDOCE dictionary have treated the lexical item [cock] with five homonymous meanings, each with several polysemous senses. But the COBUILD dictionary (1995:301) have treated only eight senses of the lexical item. The treatment of sense – related lexical items in the two dictionaries confuses dictionary users.

The other thing which help dictionary users to retrieve meaning with ease is the treatment of nominal labels and phonetic transcription. The COBUILD has provided the entry of the nominal labels in its outer access structure column.

- (13) leadership/ /leadership
1. you refer to people who are in control of a group or organization as the leadership.

COBUILD (1995:943)

The COBUILD has indicated about sixteen various types of nouns while the LDOCE dictionary has only provided the nominal label (n) for all types of nouns. The nominal label is indicated immediately after the phonetic transcription.

- (14) lip/ / N. / [c] one of the edges of the moun...

LDOCE (1978:638)

But both the COBUILD and the LDOCE dictionaries have treated phonetic transcription of nouns immediately after the entry of the lemma.

2.1.1.3 The inclusion and treatment of nouns in Afrikaans dictionaries

Nouns in Afrikaans dictionaries have been marked by the following nominal labels. The Afrikaans language dictionaries have used the symbol nw and s.nw. for naamwoorde and selfstandige naamwoorde (noun) respectively.

- (15) aanhang s.nw. volgelinge; ondersteuning
aanhanging
aanhangel
aanhanklik

HAT (1994:5)

Most of the Afrikaans language dictionaries have entered nouns by the principle of letter – by – letter.

- (16) aanblik I snw
aanbly
aanblyer (-s) s.nw

NW (1988:1)

The noun [aanblyer] has been entered alphabetical in its own place in the NW (Nasionale Woordeboek) dictionary.

Nouns in Afrikaans dictionaries have been lemmatized by the first letter of the prefix plus stem. Most Afrikaans dictionaries are word dictionaries. This means that lexical items in these dictionaries are not treated by their stems but by their full status.

- (17) **familiariteit** s. nw kyk familier
familie (-s) s.nw....
familie betrekking (e) s.nw...

NW (1988:138)

The noun **familie** (family) has been treated in the correct place of its alphabetical arrangement in the dictionary. The alphabetical arrangement of nouns in Afrikaans dictionaries are two- fold: vertical alphabetical arrangement and horizontal alphabetical arrangement.

- (18) **familie** (-e) s.nw. 1. Gesin 2. Groep verwante 3. Groep afstammelingen...4. Afdeling van n rode. 5. Versameling van dinge...
familieskap (by 2); familiekring, -lewe, -portret (by 1); -gelykenis, gebek, -kwaal, -lid, -roman, -trek, -twis (by 2), -wapen (by 3).

The noun *familie* has a number of multiple entries in this article structure. The nouns in the article structure of the lemma *familie* have been organised into four groups according to how they are morphosemantic related. These groupings made the entries not to be alphabetically correct.

(19) (i) **familieskap** (by 2)

(ii) **familiekring,**

-lewe

-portret (by 1);

(iii) **-gelykenis,**

-gebrek,

-kwaal

-lid

-roman,

-trek,

-twis (by 2);

(iv) **-wapen**

The above entries in the nest of the noun *familie* are grouped according to the gena (i) house, (ii) group (iii) house and (iv) group.

This type of treatment of nouns confuses potential dictionary users because the lemmas have been treated and arranged according to a morphosemantic field which dictionary users could not handle. There is no strict alphabetical arrangement between the nouns in the nest. Each of the above groups of lexical items have been indicated with its own alphabetical arrangement. All in all there are four types of alphabetical arrangements of nouns under the entry of the noun *familie*. Quite interesting the Afrikaans dictionaries did not give phonetic transcription of the nouns treated in them. The NW and the HAT dictionaries only provide the inflected part within parenthesis while the BWA dictionary have treated the full inflected form of the noun within parenthesis.

(20) **Gebed** (gebede; gebedjie) BWA (1994:120)

vrk s. nw. (-e)

vuuk s. nw vure

HAT (1994:1252)

The point which is important is to find an approach which is user friendly. Is it the one which enter the inflected part or is the one which enter the full word. The entering of the inflected part saves space and time, but writing the full word does not save space but only helps when dealing with entries which took an irregular form. It would be very difficult for dictionary users to find the correct spelling for the plural form of the noun *vuur* if only the inflected part is indicated. What then will be the inflected part of the noun *vuur*. If lexicographers indicate (-*re*) dictionary users could guess the plural as *vuure* or if they indicate (-*e*), still dictionary users could guess the plural as *vuure*. In a way entering the whole form helps in saving time of making wild guessing by untrained dictionary users.

Nouns derived from loanwords and other parts of speech have been lemmatized in their own alphabetical places in Afrikaans dictionaries.

(21) **tsetsevlieg** (< Tsw)

tsotsi (-s) (< Xhosa)

HAT (1994:1106)

The nouns *tsetsevlieg* (derived from Tswana) and *tsotsi* (derived from Xhosa) have been entered in their own alphabetical places in the HAT dictionary. This type of entries make the retrieval of information easy and dictionary users do not struggle when handling these lexical items.

2.1.1.3 The entry and treatment of nouns in English – Afrikaans/Afrikaans-English Translation dictionaries.

Nouns in Translation dictionaries have been entered as source and target languages respectively. In the source language, the native speaker's nouns are indicated and in the target language translation equivalents have been entered.

- (22) **baard,**
baardaar
baard gras
baardjie
baard (e) loos
baard keep
baard koring
baard man

TW (1979:54)

In its macrostructure, the TW dictionary has in alphabetical arrangement a list of nominal entries. But those who will consult TW dictionary will see inconsistency in entering nouns as lexical items. The TW dictionary has entered the lexical items *baardkeep* and *baardkoring* after the lexical item *baard (e) loos*. Potential dictionary users will not know whether *baard (e) loos* has been entered as *baardeloos* or as *baardloos*. If the TW dictionary has treated the lexical items as *baardeloos* it should have been entered between *baardaar* and *baardgras*. Still, if it was entered as *baardloos*, it should have been entered between the lexical items *baardkoring* and *baardman*. Nevertheless the arrangement of noun in TW dictionaries is by the word tradition. The TW dictionary has

entered each noun per paragraph. But other English -Afrikaans/Afrikaans - English Translation dictionaries have entered many nouns per paragraph. I think this type of entry can only be practised in word dictionaries. The nouns in these entries have been arranged horizontally in the nest.

(23) **Forgecoal**

Forge: - **furnace**, smee - oond; -hammer; namaakbaar;
 -man (...me), smeder, vuurwerker; -poker, vuur
 yster; -namaker; (skrif) vervalser; valsmunter,-
 rake, vuurhark

Forged,

GW (1997: 971)

The above entry of *forge* has a sinuous lemma file. The first lexical item in the file happens to be *forgefurnace* and the last being *forgerake*. Interesting enough all of them cannot be entered alphabetical before the lexical item *forged*. The lexical item *forged* should have been entered between *forgecoal* and *forgefurnace*. This method of entering nouns in the nest saves space but it does have an adverse effect to the dictionary users. Dictionary users will not easily find a noun entered in a sinuous lemma file. In a sinuous lemma file nouns are entered by niche or by nest. Nested lemmatisation is when nouns in the horizontal macrostructure do not maintain a strict alphabetical ordering within the nest due to optional and obligatory differences. Obligatory differences happen due to entering nouns together because they are semantically related.

(24) **Stat** (s) (-le), stable (for horses);.....

Stalagmiet, (-e) stalagmite

Stalakiet, (-e), stalactite

Stale, (fig), steal; n -wil, an iron will

Stat: -besem, stable broom, bassbroom, -boom, stable -bar;

-deur, stabledoor -lantern, stable lantern

Stalles

GW (1998: 535)

The disadvantages of nested lemmatization is that entries are brought together because they are morphosemantically related. Lexical items are not sought and found in their own alphabetical place in the dictionary but in the nest. In the nest of the lexical item *stal*: the nouns *stalbesem*, *stalboom* and *staldeer* should have been treated before the entry of the lexical item *stale* (steel). But because they are semantically related to the noun *stallantern* which could only be entered after the lexical item *stale* (steel), the nouns were entered together without maintaining alphabetical arrangement with the main lemma. This type of entries is not user friendly. The dictionary users will not be able to retrieve meaning with ease. The above entries should have been entered alphabetically correct in their own places in the dictionary . Lexicographers enter nouns in the sinuous lemma file because they want to save space and to complement the article of the main lemma; and although not so often to provide a direct semantic relation between lemmata. The need to save space by dictionaries prompted them to enter the entries in truncated form.

2.2. THE NOUN IN AFRICAN LANGUAGE DICTIONARIES.

Nouns in African languages cf. Du Plessis (1997:21) are specified for a certain noun class and through prefixes which are also known as noun class prefixes. Nouns in the African languages are not specified for gender; which means that it is very difficult in the African languages to determine the feminine and masculine by using a noun. In Spanish, for example, cf. Butterfield et al (1997 : 188) all nouns are either masculine or feminine, whether denoting people, animals or things and gender is largely unpredictable and has to be learnt for each noun. The discussion of the noun in African languages will focus on the entry and treatment of nouns in African language dictionaries and the dictionaries to be discussed here are mainly of the category of the Translation dictionaries, indicated above.

All the African language dictionaries i.e. Sepedi, Sesotho, Setswana, Tshivenda, Xhosa, Xitsonga and Zulu, will be discussed on the basis of how they have entered and treated nouns in their respective dictionaries.

2.2.1 The investigative analysis of the entry and treatment of nouns in dictionaries belonging to African languages with specific reference to Tshivenda.

The Tshivenda dictionaries to be analysed are Van Warmelo's Venda dictionary (1989) and Wentzel/Muloiwa's Improved Trilingual dictionary of Venda Afrikaans-English (1982). Other African language dictionaries to be analysed in a comparative basis with the Venda dictionaries are the following:

- C.M. Doke et al's Zulu -English dictionary (ZE)
- H.W. Pahl et al's Greater Dictionary of Xhosa (GX)
- J. W. Snyman et al's Setswana- English -Afrikaans Dictionary (SEA)
- R.A. Paroz's Southern Sotho -English dictionary (S,S,E)
- Pharo's Popular Northern Sotho dictionary (P ,N,S)
- T.J. Kriel and E.B. van Wyk's N.Sotho -Afrikaans-N, Woordeboek (NSA)

From each of these language dictionaries 5 pages were randomly selected and an investigative analysis will be done on the basis of how each one of them have treated nouns with **class prefixes**, nouns with **invisible class prefixes**, **deverbative** nouns, nouns derived from **classes 15, 20 and 21** and those nouns with Morphophonological alternations in prefix. The discussion will explore the traditional lemmatisation approach used by Tshivenda dictionaries in lemmatising nouns and the advantages and disadvantages of these lemmatisation approaches with regard to each specific language dictionary.

2.2.1.1 Nouns in class prefixes

Nouns in Tshivenda are classified according to the noun class prefixes when such prefixes are present. And these nouns belong to a noun class system,:

(25)	class 1[MU-]	MU-THU
	class 2[VHA-]	VHA-THU

class 1a [-]	KHO-TSI
class 2a[VHO-]	VHO-KHOTSI
class 3[MU-]	MU-RI
class 4[MI-]	MI-RI
class 5[LI-]	LI-IVHA
class 6[MA-]	MA-IVHA
class 7[TSHI-]	TSHI-NONI
class 8[ZWI-]	ZWI-NONI
class 9[N-]	KHOLOMO
class 10[DZI-N]	KHOLOMO
class 11 [LU-]	LU-FHANGA
class 14 [VHU-]	VHU-LUNGU
class 15[U-]	U- SHUMA
class 16[FHA-]	FHA-SI
class 17[KU-]	KU-LE
class 18[MU-]	MU-RAHU
class 20[KU-]	KU-BUDZI
class 21[DI-]	DI-KOLOMO

The two Venda dictionaries, i.e. N.J. Van Warmelo's Venda dictionary (VE) and P.J. Wentzel et al's Improved Trilingual dictionary of Venda - Afrikaans –English (VAE) have entered nouns according to the above noun class prefix system, when such prefixes are present, and by the initial sounds when such class prefixes are not present.

2.2.1.1.1 Nouns classified according to class prefixes.

Nouns in Tshivenda dictionaries may be treated according to visible noun class prefixes, e.g. Class 1: (MU)

- (26) a. **Murungi** (v ha-) kleremaker/dressmaker, tailor

VAE (1982)

b. **Muthavhi** 1 (cf.- thavha) one who stabs, cf. mubai

VE (1989)

Both Van Warmelo (1989) and Wentzel/Muloiwa (1982) have entered nouns belonging to the above category by the prefixes system. The nouns(26.a) **murungi** (tailor) and (26.b) **muthavhi** (one who stabs) have been entered by the noun class prefix **Mu-**. The entry and treatment of nouns have been done through a noun class prefix system. Nouns in these two Venda dictionaries which can be entered by the noun class prefix system are the following:

- | | | |
|------|-----------------|-----------------------|
| (27) | class 3 : Mu- | (Muri; tree) |
| | class 5 : Li- | (Linngo; mango) |
| | class 7 : Tshi- | (Tshinoni; bird) |
| | class 9: N- | (Nngu; sheep) |
| | class 11 : Lu- | (Lufo; wooden spoon) |
| | class 14 : Vhu- | (Vhutshilo; life) |

This treatment of nouns in Tshivenda in (26) above could also be experienced in other African Language dictionaries such as R. Cuenod's Tsonga-English dictionary (TE), Kriel/Van Wyk 's Pukuntsu (NSA) and Pharos' Popular Northern Sotho dictionary (PNS).

- (28) a. Moroki,snw dev kl.l LHL...

NSA(1989)

- b. Rendzo 11 (or riendzo, pl. tiendzo)

cf. -endza, journey.

TE (1967)

- c. Moroki, one who sews

PNS (1997)

The two (28ac) Northern Sotho dictionaries and the (28b)Tsonga – English dictionary have as well treated the entry of the nouns by the noun class prefixes. Nouns such as (28) can be treated by their prefixes. There is a similiar treatment of nouns belonging to this category by the NSA (1989) and PNS (1997) dictionaries, the TE (1967) dictionary, the VAE (1982) and VE (1989) dictionaries. All the above mentioned five dictionaries have lemmatised nouns from a noun class system by their class prefixes. This is not true with the two Nguni dictionaries, i.e. The greater Xhosa dictionary (GX) by H.W. Pahl and Doke/Vilakazi's Zulu English dictionary (ZE), the Southern Sotho English dictionary (SSE) and the dictionary of Setswana - English - Afrikaans (SEA). Nouns belonging to a noun class system were not lemmatised by their class prefixes, but by their stems:

(29) a. **Um.. Siki** b/n l/z : Umuntu o sika, e thunga impahla :
dressmaker, tailor, kleremake, snyer, kleremaakster
(GX)

b. -Siki (Umsiki 2.63.9. abasiki) n.
[< sika]

1. Cutter, one who cuts out patterns
2. Tailor, dressmaker

ZE (1990)

The noun (29a) **Umsiki** in GX dictionary has been entered as it exists in the language today, and not according to an underlying stem initial consonant; but the alphabetical arrangement of the noun **Umsiki** is through the stem system. In the ZE dictionary the noun (29b) **Umsiki** has been entered not as it exists in the language today but according to an underlying stem initial consonant. This system of treatment of nouns as stems is in contrast with the system of treatment of nouns as words which is practised by (26) two Venda dictionaries, (28b) the Tsonga dictionary and the two (28a,c) Northern Sotho dictionary.

The Sesotho dictionary which is in line with the Nguni dictionaries is the Southern Sotho English dictionary (SSE) by R.A. **Paroz** (1988). This dictionary enters nouns as they exist in Sesotho language today, but the alphabetical arrangements of nouns is through a stem system

- (30) a. ***molora (me)*** n., ash, ashes
 b. loti (***ma***) n. cl.3, mountain, high mountain
 With gorges; mountain range.

SSE (1988)

The SSE dictionary has treated nouns with noun class prefix when the noun class prefixes are visible and without noun class prefix when noun class prefixes are not visible. The noun (30a) **Molora** has been treated with its singular class prefix **Mo-** but the noun (30b) [loti] has been treated without its singular class prefix [Le]. This could as well be experienced in the Dictionary of Setswana - English - Afrikaans (SEA).

The SEA dictionary as well has treated nouns by their stems.

- (31) **Lodi (lodi)** l.n. Mo-Me- whistling //
 2. rel Mo-, melodious, painted

SEA (1998)

The noun (31) **Molodi** (whistling) has been lemmatised according to the stem.

2.2.1.1.2 Derived nouns in classes 15, 20 and 21

In Tshivenda classes 15, 20 and 21 are meant for infinitives, diminutives and augmentatives respectively:

(32)

Class 15	Class 20	Class 21
Infinitive	Diminutive	Augmentative
u-shuma	Kwana [<nwana]	Dana [<nwana]
u-tshimbila	Kuri [<muri]	Liri [<muri]

From the above examples it could clearly be seen that diminutive and augmentatives are derived nouns. Nouns such as (32) **kwana/dana/kuri/liri** were derived from other nouns. The infinitives (32) **U- shuma** and **U- tshimbila** were derived from verbs.

- (33) a. **Kudambo (zwi-)** riviertjie/rivulet
Kuduna (zwi-) kwaal, siekte/ailment
Kuitele (zwi-) manie, van doen, prosedure/way of doing; procedure
Kukalana (zwi-) klein kleipotjie/small clay pot.
Kunko (zwi-) klein potjie/small pot.
Kupali (zwi-) partikel, deeltjie/partied.
Kupida (zwi-) partikel, deeltjie/partied
VAB (1982: 26)
- c. U- zwa (musanda lang of chief only)
Dikolomo 21,
Dithu 21,.....
Kudi 20 (cf mudi) small village
Kudu 20 (cf nndu) small hut
Kudambo 20 (cf mulambo) small river, spruit.
VE (1989)

Both Van Warmelo (1989) and Wentzell Muloiwa (1982) have lemmatised nouns derived from classes 15, 20 and 21 respectively. The VAE dictionary has indicated nouns in class 20 which were derived from other nouns. But very few nouns derived from other parts of speech are indicated in the VAE dictionary .

- (34) class 20 **ku- : kuitele; [< -ita]**
 kunko; [< -ka]

It is similar with the VE dictionary which has treated more of the nouns derived from other nouns than those derived from other parts of speech; such as verbs.

- (35) class 15 u- : **u zwa**; [<-zwa]
 class 20 ku- : **kudima**; [<-lima]

The treatment of [**u-zwa**] by Van Wannelo (1989) confuses dictionary users. **U zwa** is an infinitive from class (15) and it is the first of its kind in the Venda dictionaries by Van Warmelo (1989). This could be the error of Van Warmelo (1989) when he tries to lemmatise [**U-zwa**]; infinitive of class (15) [**U-**]. **U zwa** in the language of the royal family means to speak, etc, but it must stand as [**-zwa**].

The other African languages do not have noun classes (20) and (21). But they all have the infinitive classes.

(36) a.

	Sotho	Tsonga	Nguni
Infinitive class	Ho-	Ku-	-ku-
Examples	Ho-sebetsa	Ku-tivha	Ukusebenza

b. **Uku-sebenza** nz/v work, labour, toil.

GX (1989)

With the exception of the GX dictionary , the Sotho dictionaries, i.e. NSA. (1989), PNS (1997) ,SEA (1989) and the SSE (1988); the Tsonga dictionary (i.e. TE 1967) and the Zulu - English (i.e. ZE 1990) have not treated any lexical item with an infinitive class.

2.2.1.1.3. Nouns classified according to initial sounds.

In Tshivenda certain nouns belonging to a noun class system have invisible noun class prefixes. These nouns are simply classified according to the initial sounds. Such nouns usually appear in the following classes :

- (37) class 1a: **Malume**
class 5: **Dzedze**
class 9: **Thanda**

Both Van Wannelo (1989) and Wentzel / Muloiwa (1982) have treated nouns belonging to this category.

- (38) a. **dzima (Ma-)** naam; naam word; opskrif; title/name; noun; title
Dzedze (Ma-) sandvlooi/huis/sandflea/louse
Malume (Vho-) swae; oom/brother-in-law/uncle.
VAE (1982)

- b. **Dzedze** 5 1. lice on dogs, fowl
2. sandflea
Malume 1 (pl vho, lit. "male mother")
1. Mother's brother
2. Wife's brother (also mulamu)
3. Wife's brother's son
VE (1989)

Nouns without visible noun class prefixes have been treated by their initial sounds. This treatment of nouns does not pose any problem with the word dictionaries such as those which belong to Tsonga and Northern Sotho, language

- (39) a. **Tatana** 1, father
Dzedze 5, flea
Hosi 9, chief king
TE (1967)
- b. **Malome** my maternal uncle

Tata, father

Tate, father

Kgosi chief, king

PNS (1997)

c. **Tata**, snw. Leenw. Kl la LH vader

Tate, snw. Kl la LH. Vader

Kgosi, snw. Kl. 9 ...koning, Hoofman, vors, meneer; ...

Malome, snw. Kl. la LHL .oom aan moederskant.

NSA(1989)

There is no problem with the treatment of nouns without visible noun class prefixes in the above dictionaries (38 a + b) and (39 a + b + c). All of the above language dictionaries treat nouns as they exist in language today. The problem is experienced with stem dictionaries such as the two Nguni dictionaries, the southern Sotho English dictionary and the dictionary of Setswana -English -Afrikaans. These four dictionaries are stem dictionaries and the entry of nouns is through a stem tradition. Nouns which appear in class 9 are difficult to handle in stem dictionaries.

(40) a. **-Khosi** (inkosi, amakhosi) king, paramount chief

ZE (1990)

b. **in-kosi** bin 9/10: paramount chief

GX (1989: XXXV, adopted From paragraph 5)

c. **Malome** (bo) my maternal uncle.

(n) tate (bo) my father, father sir

SSE (1988)

In GX dictionary the noun (40b) inkosi could have been entered in the alphabetical position of KO and not of KHO as in the (40a) ZE dictionary. The ZE dictionary has

entered inkosi according to an underlying stem initial consonant transformed into an ejective consonant under the influence of a preceding **n**. Both Pahl (1989) and Doke et al (1990) have treated the entry of the noun **inkosi**; but differently. Doke et al (1990) has used an underlying form **khosi** and Pahl (1989) has used the noun **inkosi** as it exists in the language today. This indicates the failure of stem dictionaries to accommodate nouns of classes 9 and 10 as lexical entries. These type of entries are better dealt with in word dictionaries. Another serious problem is the entry of the noun (40c) **Malome** in SSE dictionary. The SSE dictionary has lemmatised the noun (40c) **Malome** according to the letter of alphabet L. To Paroz (1980) the **Ma-** before **-lome** in **Malome** is a class prefix and **-lome** (suffix) is a stem. In order to make the entry in line with other stem entries Paroz (1988) had to lemmatise it as a noun with visible class prefix [Ma-]. This type of entry could have been avoided in a word dictionary where it could have been lemmatised by its first letter.

2.2.1.1.4 Indication of class prefixes

The second of the paired prefixes is not directly represented in these Venda dictionaries. Van Warmelo (1989) only indicates the number of the singular prefix after the entry while Wentzel/Muloiwa (1982) give the plural prefix after the entry:

- (41) **Malume** 1 (pl. vho; lit "male mother")
 1.Mother's brother 2. wife's brother,...
- Murungo** 3 (cf. -runga) seam
- Dzedze** 5 1 lice on dogs fowls sandflea.
- Thanda** 9 (cf. danda) 1 (cl 9) long pole 2. (cl.10) pl. of lutanda
- Lufu** 11 (pl. mpfo 10) 1 wooden spoon for stirring 2 canoeman's
 Paddle
- Vhutshilo** 14 (cf. -tshila) life
- Kudambo** 20 (cf. mulambo) small river, spruit.

VE (1989)

The numbers indicated after the entry of the nouns represent their singular form. With these numbers in (41) above Van Warmelo (1989) wants to indicate that the noun in question is in its singular form. This is different to the technique used by Wentzel/Muloiwa (1982).

- (42) **Malume** (vho-) swear; oom/brother in law; uncle
Murungi (vha-) kleremaker/dressmaker; tailor
Dzedze (Ma-) sanfdvlooi, luis /sand flea; louse
Thanda (dzi-) stock, paal; timmerhout/pole; stick; timber
Lufo (Mpfo) lepel/ladle, wooden spoon.
Vhutshilo (geen mv/no plural) lewe/life
Kudambo (zwi-) riviertjie/rivulet

In contrast to Van Warmelo (1989) who indicates the number of the singular prefix after the entry, Wentzel/Muloiwa (1982) indicate the plural prefix after the entry. The advantage of the technique used by Wentzel/Muloiwa (1982) is that it covers two entries where as the technique used by Van Warmelo repeats the information displayed by the noun. The noun **dzedze** is singular and writing the number 5 which represents the singular prefix is providing information redundancy in Van Warmelo's dictionary. It's like the dictionary is teaching noun class system which is not the task of the dictionary. In Wentzel/Muloiwa (1982) the noun [dzedze] is a **lexical** item and writing the plural prefix [Ma-] after it is an indication of another lexical item known as **madzedze**. But in order to save space and time Wentzel/Muloiwa (1982) had to write the plural prefix after the lexical entry. Another disadvantage of this system by Van Warmelo is inconsistency in indicating the number of the noun class prefix.

- (43) a. **Musadzi** 1 (cl.1) woman; wife. 2 adj. cl.3 of -sadzi " female".
 VE (1989)

Van Warmelo (1989) indicates inconsistently, the number of the singular prefix after the designation of the polysemous senses. After the noun **Musadzi** the number 1 does not

represent the noun class prefix but it represents the first polysemous sense of the noun **musadzi**. The noun class prefix is indicated after the sense number within parenthesis. This is in contrast with other entries where numbers indicating the noun class prefix were entered immediately after the entry. The other African language dictionaries which use numbers to indicate class prefixes are Pahl's Greater Xhosa dictionary (GX), Kriel/Van Wyk's Pukuntsu, N Sotho - Afrikaans dictionary (NSA) and Cuenod's Tsonga - English dictionary (TE).

- (44) a. **Um. Siki** b/n 1/2 ...tailor
In. tloni b/n 9/10 hedgehog
GX (1989)
- b. **Moroki**, snw dev kl. 1 LHL
NSA(1989)
- c. Xitlhoni 7, = nhloni, hedgehog
TE (1967)

All the three (44) dictionaries have entered numbers which represent the singular noun class prefixes. In addition the GX dictionary has entered both the numbers which represent the singular and the plural forms respectively. The numbers indicating the singular forms are repeating information represented by the nouns entered in these dictionaries. But the number 2 in GX dictionary represent the plural prefix for the plural form of the noun **Umsiki**. For dictionary users, to understand this system used by the GX dictionary, they have to know what number 2 represents in noun class system. This is another way of teaching noun class system by the GX dictionary. Before one retrieves meaning he has to know and understand the noun class system. In this case users must know that number 2 represents [aba] in the Nguni noun class system. The other dictionaries use the class prefixes and have entered them after the nominal entries.

- (45) a. **sese**, mo- me-, dress
SEA (1989)

- b. **roka**..... ..
 Moroko (me) n. beer dregs
 SSE (1988)

The two dictionaries have entered the plural prefix after the entry of nouns (45a) **Mosese** and (45b) **Moroko** respectively. The SEA dictionary has entered the plural prefix **Me-** after the nominal entry (**sese, Mo-**). And the SSE dictionary has entered the plural prefix **Me-** after the nominal entry **Moroko** which has been treated in the article structure of the lexical item (45b) **Roka**. The ZE dictionary have entered both the singular and the plural nouns after the nominal stem.

- (46) a. **-siki (umsiki, 2.6.3-9. abasiki)**...
 ZE (1990)
- b. **Moroki**, one who sews
 PNS (1997)

The re-entry of the complete form (46a) **Umsiki** does not save space and time. Even the entry of the noun (46a) **abasiki** could have been indicated by the plural prefix. There is no reason why a complete noun form is provided. The complete noun form should only be provided when there is an irregular spelling form. The (46b) does not indicate any noun class prefix after the entry of the noun. The noun is treated as it exists in language today without any reference to a class prefix. The SSE dictionary does indicate the class prefix. The SSE dictionary writes the plural prefix after the entry of the noun. At times the plural prefix is indicated by the number. The only problem with SSE dictionary is that class prefixes are numbered differently.

- (47) Mohwebe, n cl 2 muddy water, muddle
 SSE (1988)

The number representing the singular noun class prefix **Mo-** in (47) is 2 instead of 3 and the number 3 represents noun class prefix **le-** which is usually represented by 5. This is

because SSE dictionary has one number for each pair of noun class prefixes. The SSE dictionary used class 2 for both singular and plural class prefixes Mo- and Me-. This is not in line with other African language dictionaries which give each class prefix a separate number. Class prefixes Mo- and Me- will be represented by numbers 3 and 4 respectively.

2.2.1.1.5 Morphophonological alternations in prefixes

In Tshivenda the plural forms of nouns are indicated in two ways. Both Van Warmelo (1989) and Wentzel/Muloiwa (1982) give the full plural in brackets after the entry and these plural forms are again entered separately in the dictionary as lexical items.

(48) a **bango (mapango)** groot paal/big pole
 mapango cf. bango

Gopwa (magopwa/makopwa) armholte/armpit
 Makopwa cf. gopwa

Ino/lino (mano) tand/tooth
 mano cf. lino.

Kwana (zwana) kindjie/small child
 Zwana cf. kwana

Luhura (khura) heining/fence, hedge
 khura cf. luhura

Nwaha (minwaha) jaar/year
 minwaha cf. nwana.

VAE (1982)

b. **bango** 5 (pl. mapango) pole..

Mapango 6 pl. of bango

Gopwa 5 (pl mag, mak) armpit

Makopwa 6 pl. of gopwa

Ino 5 (pl mano) tooth.

Mano 6 (pl. of ino) teeth.

Kwana 20 (pl. zwana) small child

Zwana 8 pl. of kwana small child.

Luhura 11 (pl. khura 10) outer Fence or stockade round a
village or garden

Khura 10 pl. of luhura

Nwaha 3. (pl. minwaha)year

Minwaha 4. pl. of nwaha

VE (1989)

The above treatment (40) of nouns in both Van Warmelo (1989) and Wentzel/Muloiwa (1982) increase noun redundancy and potential spaces for other important lexical items are occupied .Both Van Warmelo (1989) and Wentzel/Muloiwa (1982) provide the entry of plural nouns with a cross-reference to the singular nouns. In VAE and VE dictionaries, the plural nouns are not explained but are explained through the entry of both the singular and plural nouns. The entry of both the singular and plural nouns as lexical items is found in word dictionaries rather than in stem dictionaries. The word dictionary uses various class prefixes in the lemmatisation of nouns while a stem dictionary uses the initial letter of the nominal stem for both singular and plural nouns.

- (49) a. **Moreki**; buyer
Bareki; buyers ,customers.
PNS(1997)
- b. **Mogwe**, snw k/1 LH : swaar ,skoonseun
Bagwe, snw k/2 LH : skoonouers
NSA(1989)

The (49) word dictionaries lemmatise plural nouns in one way. Plural nouns are indicated only under the nominal entry .But stem dictionaries lemmatise plural nouns under or within the stem entry.

- (50) a. **Um sabi** b/n 1/2 fugitive
GX(1989)
- b. **-ndima** (indima, **izindima**) n.[(lima)] plot of ground set
for cultivation
ZE(1990)
- c. **le nwabo** (ma) n **vha meleon**
SSE(1988)
- d. **sepedi**, mo- ba- person who travels on foot.
SEA (1989)

2.2.1.2. Derived nouns

This paragraph will explore the entry and treatment of nouns derived from parts of speech such as nouns, verbs, diminutives and locatives.

2.2.1.2.1. Nouns derived from nouns.

In Tshivenda nouns may be derived from other nouns by affixation.

- (51) a. those derived from locative suffix -ni
 b. those derived form diminutive affixes ku-, nyana and -ana
 c. those derived from gender suffixes such as -lume 5 kadzi, -nyana and -ana

Both Van Warmelo (1989) and Wentzel/Muloiwa (1982) have treated nouns with derivational affixes. These are nouns formed by affixing prefixes and suffixes.

- (52) a. **hayani**, huis toe, tuis/ at home
khofheni, in/on the face
nwananyana (vhananyana) meisie/girl
ngwana lam/lamb
malume (vho) swaer; oom
brother in law; uncle
mufumakadzi (vha) mevrou/woman
 VAE (1982)
- b. **Haya** (pl. mah) home, in general; locat **hayani** home
Khofheni 15 (loc of khofhe) face, visage
Nwananyana 1 (pl. vhananyana) girl; daughter
Ngwana 9 (cf. nngu) lamb
Malume 1. (pl. vho; lit "male mother")
Mufumakadzi 1. married woman of rank.
 VE (1989)

Wentzel/Muloiwa (1982) have treated nouns derived from derivational affixes as lemmas in their own alphabetical places of arrangement in the dictionary . Van Warmelo (1989)

has used two approaches in treating nouns such (52b) **hayani** and **ngwana**. Van Warmelo (1989) has treated (52b) nouns as lexical items and as part of other lexical items. In Van Warmelo (1989) the noun **hayani** has been treated as part of another lexical item (52b) **haya**. The other African language dictionaries which have treated nouns in similar way to Wentzel/Muloiwa (1982) are the following:

- (53) a. **Morakeng**, at the cattle ranch tseleng, on/in the road
Ngwananyana, a little girl
Tselana, small road
Malome, my maternal uncle
Mohumahadi, lady, queen, Mrs.

PNS (1997)

- b. **ngwanenyana** snw. Dim kl I klein meisie
tselanathokwana kl. 9 sypadjie.
malome, snw. K1. la. LI + L oom aan moederskant
mohumagadi, snw. Fem k11. LH1 + LL: dam
vrou, eggenote.

NSA(1989)

The (53) two Northern Sotho dictionaries have treated nouns derived from derivational affixes as lexical items. The majority of the African language dictionaries treated nouns similar to Van Warmelo's approaches. Most of these dictionaries have treated these nouns as lexical items and as entries within article structures of other items.

- (52) a. **Xitiko**. 7 fire place. Loc.
Xitikweni, center of hut floor.

TE (1967)

- b. **intaba** ncp dim. In. tatyana loc entabeni
entabeni L + d loc < intaba
in. tatyana b/ n ncp < intaba: dim < intaba

in. tlungu b/n a/10 = hlungu l + d/

loc. Entlungwini (sorrow, grief, sadness, distress)

GX (1989)

- c. **mo.raka** (me) n. cattle post; loc **moraka, morakeng**
o hole, at a distant cattlepost; **meraka, merakeng**.

Tsela' (di) n. road, way path; manner, method;

Journey.....

..... loc **tseleng**; dim. **Tselana**. /tsela/

ngwana (bana) n. d. 1. Child.....

.....

ngwanana (banana) n. d. 1.

Girl.....

.....loc. **Ngwananeng**, at the home of a girl; dim

ngwananyana, little girl

Malome (bo) n. my maternal uncle

SSE(1988)

The above dictionaries (54 a + b + c) have treated nouns derived from derivational affixes as lexical items and within the article structure of other lexical items. The GX dictionary has treated the nouns (52b) entabeni and (52b) **intatyana** as lemmas and as entries with the article structure of the lexical item **intaba**. In contrast to dictionaries discussed above Doke/Vilakazi (1990) have entered the nouns derived from affixes differently.

- (54) a. -ana dim suff. added to nouns and qualificatives to indicate;
i. a small a small object.
e.g. intwana a little thing.
ii. a young one, e.g. **inkonyana** (calf)
< inkomo

iii. a small a few, e.g.

izinsukwana (a few days)

< izinsuku

b. kazi aug suffix

c. kazi fern. suffix.

Malume (umalume/omalume) lit. My male mother.

ZE (1990)

2.2.2.1.2.2. nouns derived from verbs.

Both Van Warmelo (1989) and Wentzel/Muloiwa (1982) have treated deverbative nouns.

But the entry and treatment of deverbative nouns in these dictionaries are different.

(56) a. **Mushumi** (vha) werker/worker

muloro (mi) droom/dream

pfunzo (dzi opvoeding/education.

VAE (1982)

b. **mushumi** 1. (cf. -shuma) worker

muloro 3. (cf. -lora) dream

pfunzo 9 (cf. -funza) teaching

VE (1989)

The VAE (1982) dictionary has treated the following deverbative nouns (56a) **mushumi**, **muloro** and **pfunzo**. These nouns are derived from verbs such as **-shuma**, **lora** and **-funza** respectively. The VAE dictionary has entered and treated these (56a) nouns without giving reference to the verbs from which they are derived. In Wentzel/Muloiwa (1982) derivations are treated in their own alphabetical places in dictionary. Derivations are not treated within the article structure of other lexical items from which they are derived. The verbs are also treated in their own alphabetical places of arrangement in the

dictionary. The lexical items **mushumi** (worker) and **shuma** (work) have no reference to each other in Wentzel/Muloiwa (1982). This is in contrast with Van Warmelo (1989) who has treated the deverbative nouns in a similar way to Doke/Vilakazi (1990)

- (57) -sebenzi (umsebenzi /abasebenzi) [< -sebenza] worker
 phupho (I (li) phupho/amaphupho) n. [< -phupha] dream.
 Fundo (imifundo, sg only) [< -funda] education, learning
 ZE (1990)

Both Van Warmelo (1989) and Doke/Vilakazi (1990) indicate the derivative after the entry of the noun and the class prefix designation (Van Warmelo) and after the part of speech designation (Doke/Vilakazi). The other dictionaries which indicate deverbative nouns and parts of speech from which they are derived are the Southern Sotho - English dictionary (SSE) Pharos' Popular Northern Sotho dictionary (NSA) and the Greater Xhosa dictionary (GX).

- (58) a. **dira** v.t. to do.....
**tiro**, n. action, deed; use
ruta, v.tt. to teach, to preach.....

muroti (ba) n. teacher
 minister, preacher, pastor, missionary.
thuto (di) n. teaching, doctrine, dogma, religion
 lesson, subject; sermon; education, instruction
 SSE(1988)
- b. **mosomi**, snw. Dev kl. 1. LI + HL: werker.
thutho, snw. Dev kl 9 HL: les leerstelling,
 onderwysers, onderrig, stdie, kursus, verhandeling
toro², snw. Dev. K/9 I + 1 : droom.
 NSA(1983)

- c. **uku. Sebenza** nz/v: work, labour
toil; do manual or mental
work; perform a service.
b/n Y 1/2 **umsebenzi**; 3/4 1 ncp/dim
umse to yenzana td/loc
emsebenzini. works, labour
um. sebenzi. b/n 1/2 employee, workman, labourer
GX (1989)
- d. **mosomi**, worker, labourer.
Toro, dream
Thuto, teaching, lesson, doctrine education.
PNS (1997)

The SSE dictionary treats deverbative nouns with reference to the verb from which they are derived. The SSE dictionary has treated the nouns (58a) **tiro**, **moruti** and **thuto** under the verb or within the article structure of the verb **dira**. This indicates that the three nouns have a reference to the verb **dira**. The NSA and GX dictionaries have different methodologies for indicating derivative nouns. The NSA use the symbol **dev** after all the nouns derived from verbs but it does not indicate the verb from which the nouns is derived. After the nouns (58b) **mosomi**, **thuto** and **toro**² there is this symbol dev to indicate that all these nouns are deverbatives. The GX dictionary like the SSE dictionary have entered deverbative nouns within the article structure of the verb. The noun (58c) **umsebenzi** is a deverbative noun entered within the article structure of the verb (58c) **ukusebenza**. The Pharos' Popular Northern Sotho dictionary (PNS) like the VAE dictionary has entered deverbative nouns without any reference to the verb from which they were derived. The (58d) nouns **mosomi**, **toro** and **thuto** have been entered without any reference to the verbs **soma**, **lora** and **ruta** respectively. Treating nouns with reference to the verbs from which they were derived has a minor advantage and a disadvantage. The approach does not reduce noun redundancy but treat nouns in a complex situation where it could be difficult to retrieve meaning. The dictionaries which treated nouns from verbs were trying to give the meaning of nouns through that of verbs.

Dictionary users do not understand the thematic sorting of meaning but look for the place where the noun is alphabetically arranged in the dictionary. Providing the meaning of the noun through that of verb may be necessary but that is not what the dictionary are written for: meaning must be derived from the lexical item concerned. The meaning of (56) **mushumi** must not be derived through the meaning of [shuma]. Meaning must be derived from each and every lexical item respectively.

CHAPTER 3

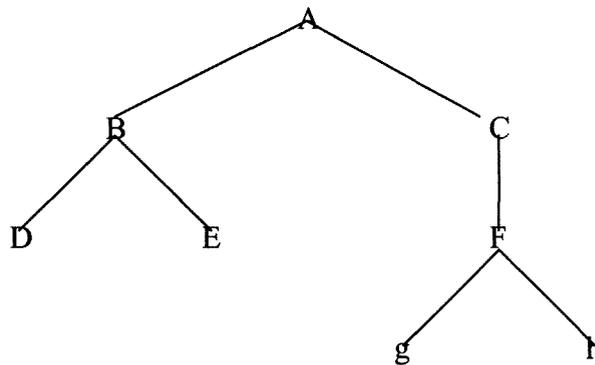
3.1 SYNTAX OF THE NOUN

The aim of this section is to investigate the syntax of the noun with regard to x-bar theory.

3.1.1 THE X-BAR THEORY : the noun is the head of the NP.

Before discussing the existence of the noun as the head of the noun phrase (NP), there are general properties of the phrase structure which need to be stated. In Tshivenda a phrase may consist of one or more constituents. A phrase **A** may have constituents **B** and **C**. And constituents **B** and **C** may consist of constituents **D**, **E** and **F**, but when there is no need to specify the internal structure of **F** for the purposes of a particular discussion, this part of the tree is simplified into a triangle connecting the final element

(59)

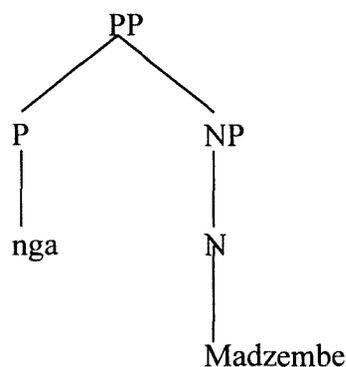


The phrase structure of the sentence is a hierarchy, with each constituent successively consisting of other constituents, until only non-expandable items are left. The 'consist of' relationship can be expressed as rewrite rules ($A \rightarrow B C$) or as bracketing ($[A,B,C]$). An item that appears above another item in the tree is said to dominate it. **A** in the tree dominates every thing from **B** - **F**; **B** dominates **D** and **E** and **C** dominates **F**. Because **A**

comes immediately above B and C, B above D and E and C above F, thus A, B and C immediately dominate items that appear immediately above them. But A does not immediately dominate D, E and F. B and C and D and E are regarded as sisters. Each of them is dominated by the same element. This form of phrase structure employed is X-BAR SYNTAX. X-bar syntax is distinctive in claiming that every phrase conforms to certain requirements. Thus a noun phrase (NP) such as **itshi tshinoni** contains a head **tshinoni**. An essential requirement of X-BAR SYNTAX is that the head of the phrase belong to a particular category related to the type of phrase. A noun phrase (NP) contains a noun (N) as the head. The head of a phrase is not related arbitrarily to the phrase type; it is not chance that an NP contains an N rather than V. The general principle that all phrases contain a particular type of a head can be formalised: as $XP \rightarrow \dots X \dots$. The X in both places stands for the same category : any phrase XP must have a head X of the same type -NPs contain N heads. It is a principle of X-BAR theory that phrases have heads of same category as the phrase itself. Here the discussion looks at the noun (N) as the head of the noun phrase (NP). Of the four lexical phrases used in X-BAR SYNTAX , structure of the noun phrase exists in three of them: verb phrase (VP) noun phrase (NP) and the prepositional phrase (PP). Below is an example of a phrase structure for each type of lexical phrase.

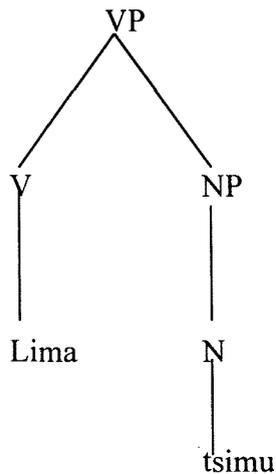
(60) Vhana avha vha lima tsimu [nga madzembe]

These men plough the field with handhoes.



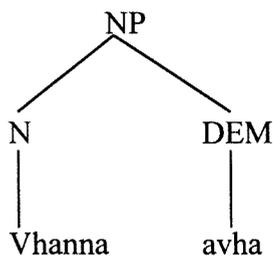
Within the prepositional phrase (PP) **nga madzembe** (with handhoes) there is a noun phrase (NP) **madzembe**. In the NP **madzembe** the lexical category which is the head of the phrase is the noun (N) **madzembe**.

- (61) Vhana avha [vha lima tsimu] nga madzembe .
 These men plough the field with handhoes .



Within the Verb Phrase (VP) (61) **vhalima tsimu**, exists the noun phrase (NP) **Tsimu**. In the NP **Tsimu** the lexical category which is the head of the phrase is the noun (N) **Tsimu**.

- (62) [Vhana avha] vha lima Tsimu nga madzembe.
 These men plough the field with handhoes.

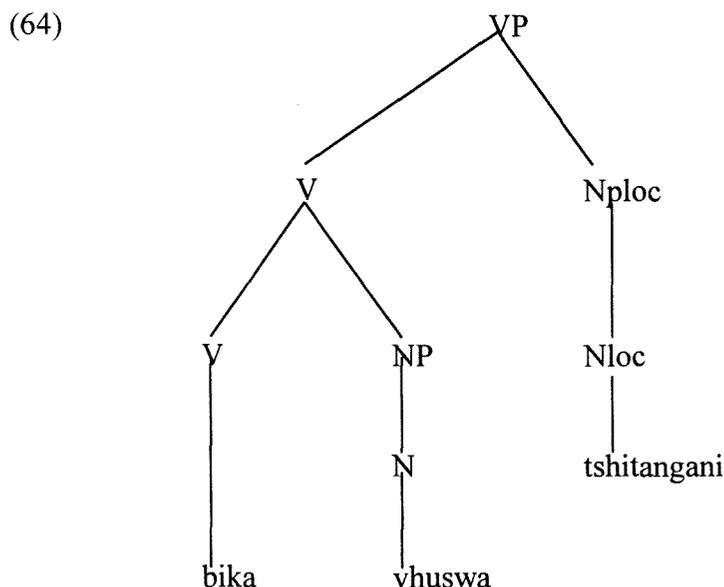


The noun [vhanna] is the head of noun phrase (NP) (62) **vhanna avha**. In all the noun phrases, the head is the noun. Although the head appears with the demonstrative [avha] in [vhanna avha] the N remains the head.

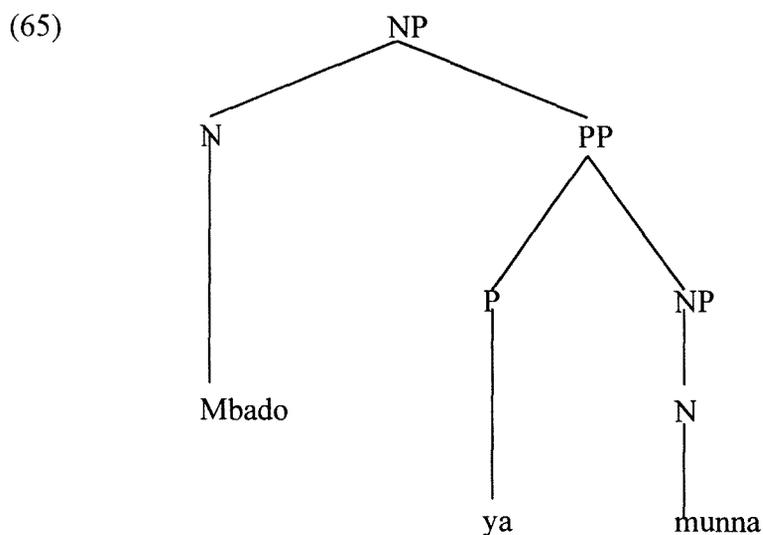
Each phrase consists of a head and other elements, which are phrases themselves; each phrase is built up around a lexical category such as N from which it takes its name and its main properties.

(63) Miri mihulu i thusa kholomo nga murunzi
(Big trees help cattle with shade)

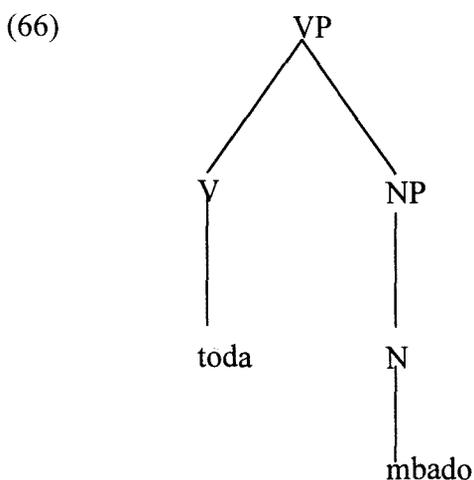
The sentence (63) above demonstrate how the structure depends on lexical heads of phrase. The first NP **miri mihulu** contains a head N, which is the lexical item **miri**. The phrase structure of the sentence is linked to the lexicon not just because the sentence eventually consists of actual words, but also because the heads of the phrase within the sentence must be lexical categories and must suit the particular structure of which they form part. An NP is a phrase which contains a head N and a NP is in a sense the elaboration of the properties of its head N. The x-bar theory claims that the phrase level, i.e. XP where X starts for any of the categories and the category level, i.e. X, are insufficient to capture all the details of phrase structure, a further intermediate level is required.



The NPloc **Tshitangani** is detached from the head V, and the close relationship between V and object NP can be expressed through sisterhood in the V phrase. But VPs are only one type of phrase. X-bar syntax extends the same solution to all other phrases; they all have an intermediate level. By convention the top level category has two bars; a VP is V", an NP is N", and so on. Each double bar phrase contains a single -bar category head; a V", N" contains N' and so on. The principle can be shown in the NP **mbado ya munna** (the axe of the men) the lexical category which is the head of the phrase is the N **mbado**.



The VP (Verb Phrase) **toda mbado** (want axe) has as its head the V **toda** and contains within it an NP **mbado**.



The AP (Adjective phrase) **zwavhudi nga maanda** contains an adverb head **zwavhudi** and **nga maanda**.

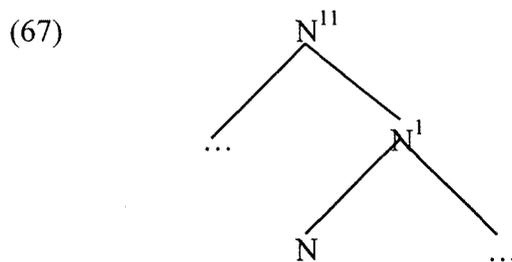
In this way:

$X'' \rightarrow \dots X' \dots$

X stand for any category that can be the head of a phrase. Each single –bar phrase in turn contains a Zero-bar lexical category head ;thus a V' contains a V and N' contains and N and so on. The principle can be put as:

$X' \rightarrow \dots X \dots$

Where X stands for the same category as before. A zero bar is the head itself ,usually written as X rather than X'. The skeleton of a NP is:



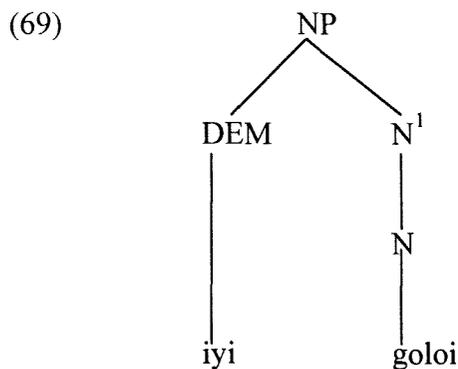
The above structure shows that N build into N' and from there into N'', V' and V'' and so on.

- (68) Uyu mukalaha u -rwa mukegulu
(This grandfather beats grandmother)

The above sentence can introduce the notion complement. The transitive V **rwa** is closely related to a following object in the VP such as **Mukegulu**. Complement are always

complete phrases in themselves; the complete NP **Mukegulu** is a complement in the VP **rwa Mukegulu**.

Phrases contain a third main element in this structure which is called the specifier, While complement belong along side the X in the phrase i.e. are sister of X specifiers belong along side the X'. The demonstrative for example is a specifier in the NP **iyi Goloji**. The structure of **iyi Goloji** is therefore



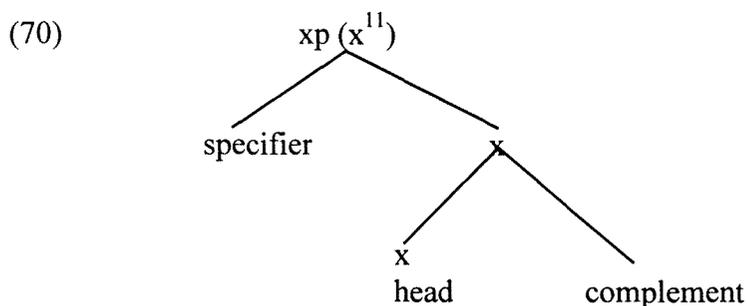
The over all principle is that a double -bar phrase may consist of a head X' and possible specifiers,

$X'' \rightarrow X'$ specifier

Or

$X'' \rightarrow X'$ specifier X'

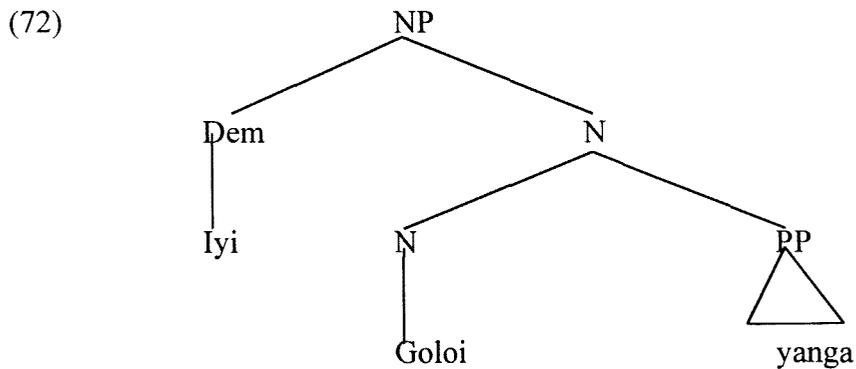
The principle does not spell out the order of specifier and head X: putting the X-bar principles together and labelling the positions gets the structures:



Lets us flesh this out with the NP:

- (71) Iyi goloji yanga
(This car of mine)

This is N' (NP) having a specifier consisting of the dem **iyi**, and an N' **goloji yanga**; this N' in turn consist of an N, the lexical item **goloji**, and a complement PP **yanga**. Both the specifier and complement position in the phrase are filled.



In addition a phrase may have adjuncts i.e.

- (73) uyu munna u reila goloji badani
(This man drives the car at the road)

Where the NPloc **badani** is an adjunct

This can be seen in the relative clause such as

- (74) Vhasidzana [avha vha shuma] u wana tshelede
(Girls who works get money)

The NP has a head N **vhasidzana**. The relative clause [**avha vha shuma**] cannot be a specifier; it cannot be a complement since the N **vhasidzana** does not specify a complement.

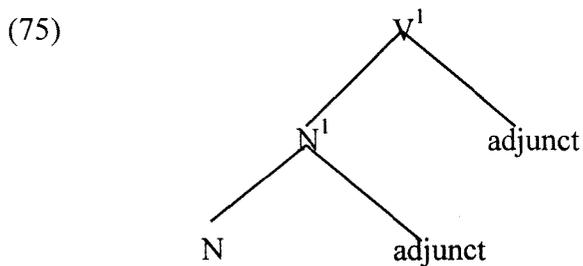
This adjunct can be expressed as

$N' \rightarrow N' \text{ Adjunct}$

The other way is to expand it into an N and its complement:

$N' \rightarrow N \text{ complement}$

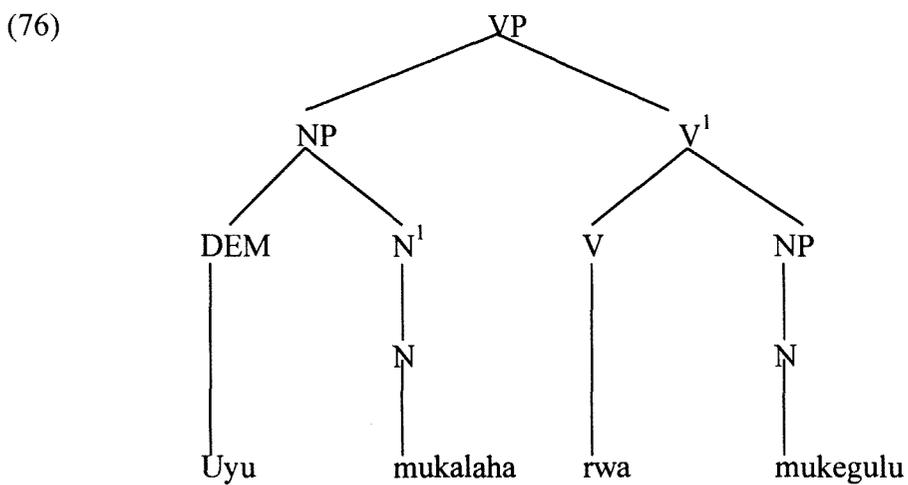
Each adjunct is added to the NP by recursively expanding an N^1 into another N^1 with the adjunct as its sister:



Uyu mukalaha u-rwa mukegulu

(This grandfather beats do the old grandmother)

Can now be represented as:



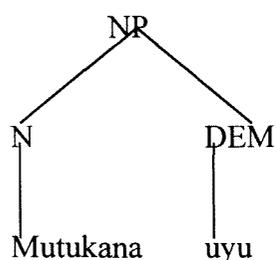
3.1.2. The structure of the NP

Noun phrases appear with a noun as head. These nouns may appear with various nominal modifier as complements.

3.1.2.1. DEMONSTRATIVE [DEM]

(77) Mutukana uyu

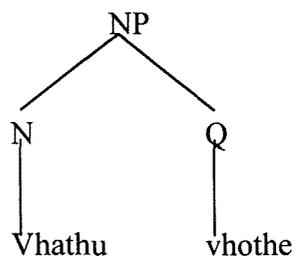
(This boy)



In (77) the Noun Phrase (NP) is (**Mutukana uyu**). The noun (N) Mutukana is the head of the noun phrase and uyu is the demonstrative.

3.1.2.2. QUANTIFIER

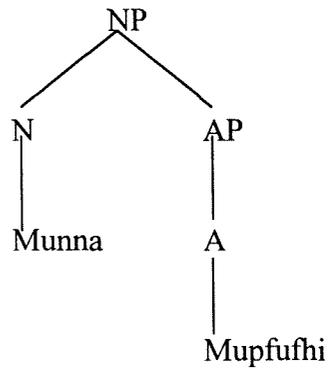
(78) VHATHU VHOTHE



In the above example **vhothe** is the quantifier and the head of the noun phrase is the noun **vhathu**.

3.1.2.3. ADJECTIVAL PHRASE [AP]

- (79) Munna mupfufhi
(Short man)

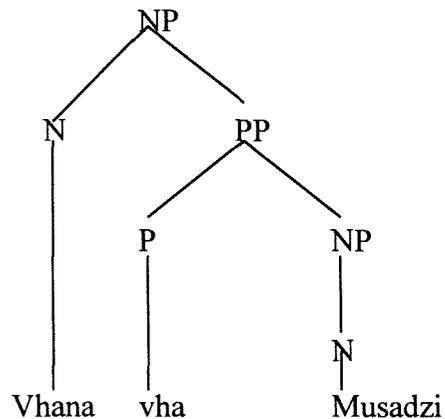


Munna is the head of the noun phrase [Munna mupfufhi] and mupfufhi is the adjective .

3.1.2.4. PREPOSITIONAL PHRASE [PP]

3.1.2.4.1 Possessive

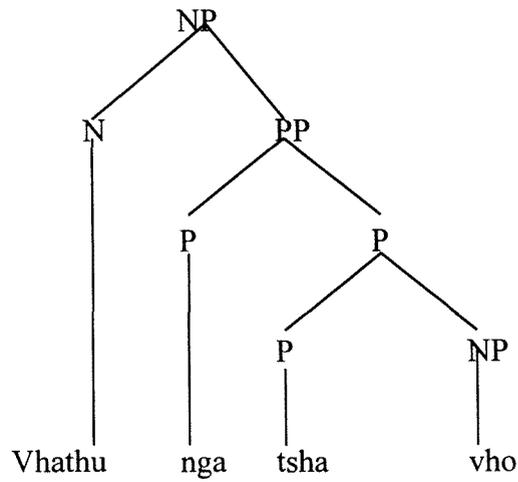
- (80) VHANA VHA MUSADZI



The head of the noun phrase in the above example is **vhana**.

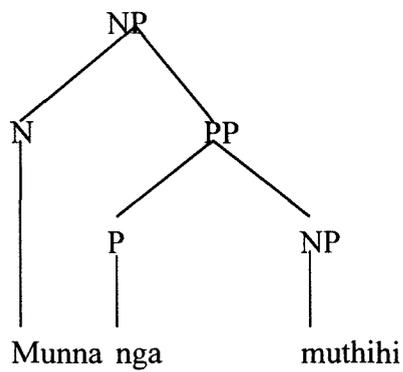
3.1.2.4.2. REFLEXIVE EMPHASIZER

(81) VHATHU NGA TSHA VHO



3.1.2.4.3. EMPHASIZER

(82) MUNNA NGA MUTHIHI

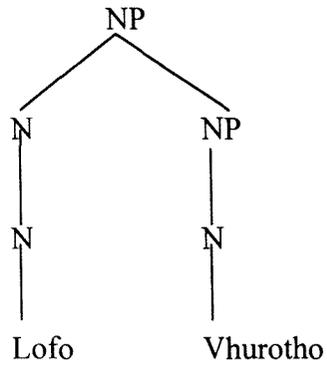


Munna is the head of the noun phrase [munna nga muthihi]

3.1.2.4.4. NOUN PHRASE [NP]

Descriptive Possessive

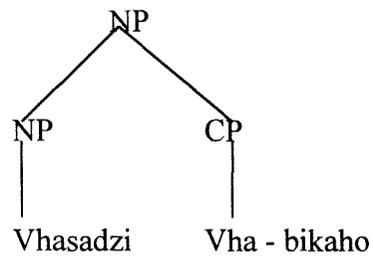
(83) Lofu ya Vhurotho



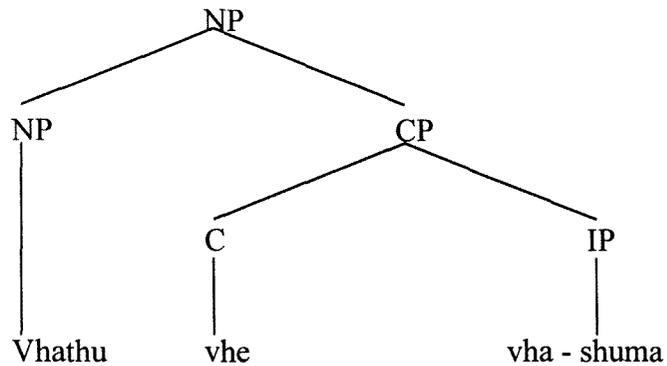
3.1.2.4.5. COMPLEMENTIZER PHRASE [CP]

Relative clause

(84) Vhasadzi vha bikaho



(85) Vhathu vhe vha shuma
Vhathu vhane vha shuma



3.1.3. Theta role and Predicate Argument Structure

The NPs that are assigned theta -roles (θ -role) are called arguments. Arguments could be NPs and clauses which refer to person, things and places. The term Theta role can be a synonym of the term. The lexical representation is that of predicate argument structure (PAS) and there is no mention of semantic content of θ -roles within this lexical representation. An argument may have a specific semantic relationship with its predicate and here a lexical semantic representation exist. In the lexical syntactic representation the predicate argument structure (PAS) of a verb indicates the number of arguments it takes. A predicate is described according to the number of argument that it takes and each argument will have a specific variable corresponding to such an argument or alternatively, such variable may contain semantic labels such as agent, theme.

- (86) a. **Na:** X (a variable)
 theme (a semantic label)
- b. **La:** X(Y) (variables)
 agent (theme) (semantic labels)
- c. **Fha:** X(YZ) (Variables)
 agent (recipient theme)(semantic labels)

The above (86a + b + c) represent predicates: (86a) is ONE PLACE PREDICATE, (86b) is a TWO PLACE PREDICATE and (86c) is a THREE PLACE PREDICATE. The assignment of θ -role is governed by projection principle and the θ criterion. The projection principle ensures that a verb may only subcategorise for complements that it θ - marks. The criterion imposes a one-to-one association between θ -roles and arguments: each argument bears one and only one θ -role. Each θ -role is assigned to one and only one argument. Each variable in the P AS must correspond to an NP (syntactic constituent):

- (87) [Nwana] ula [Vhurotho]
Child eats bread

The variable X corresponds to the NP **nwana** while the variable Y corresponds to the NP **Vhurotho**. The X variable is an **agent** and the Y variable is a theme. These variables are found in PAS of 1a and the two arguments represented by **nwana** and **vhurotho** will each have one θ -role. There are about three modes of θ -roles assignment: by a verb, a preposition and a VP via predication and θ -role assignment gives the association between the NPs +n the argument position of verb in the syntax and the variables in -the PAS of the verb. The NP arguments of a verb in the syntax are not the same and the NP argument assigned a theta role by the VP via predication must be outside the maximal projection of the verb. The remaining arguments are internal to the maximal projection of the verb. If a verb assigns a theta role to a position, it θ - marks that position and the argument that appears in the position subcategorized by a verb is called the internal argument. When a verb assigns a theta role directly to the NP the direct NP argument is formed and when a preposition assigns a theta role to the NP an indirect NP argument is formed. The internal NP arguments are assigned their θ -roles in the syntax under government. The manner of θ -role assignment is part of the lexical properties. The lexical representation of a verb must include a specification of how each NP argument is assigned its theta role together with the number of arguments of each verb.

- (88) a. **vhea** x < y loc z > (variables)
agent < theme loc location > (semantic labels)
- b. [Musadzi] u vhea [bodo] [tshivhasoni]

The structure at (88a) has three variables. The verb **vhea** is a three place predicate with the NP arguments in brackets. In (88a) the variable X is outside the brackets in the PAS and it represents the external argument; the NP musadzi in (88b). The variables Y and Z in (88a) are inside the brackets representing the internal argument: one being direct and the other one being indirect. The indirect one Z being assigned its theta role by a locative

morpheme **-ni**. The manner of theta role assignment must indicate variables in the **PAS** which are direct, indirect) internal or external arguments. The lexical syntactic representation with variables in (88a) establishes correspondence between positions in P AS and positions in syntax; but makes no reference to the semantic content of the NP arguments. In the place of variables X, Y and Z theta role labels can be given:

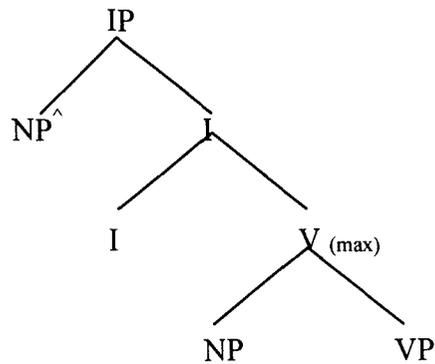
(89) vhea: [Agent Theme Location]

There are no criteria for determining which theta role label an argument may have but an appropriate lexical semantic representation can be found in the lexical conceptual structure of Jackendoff (1990). He gives a list of the combinations of thematic roles and NP argument structures

(90) Actor/Theme [Mashudu] u a tuwa
 Actor [Agent, Beneficiary: [vhanna] vha thusa [vhasadzi]

The subject of **tuwa** can be assigned both the actor and theme theta roles. In assigning these two theta roles there must be a semantic decomposition of the verb into a structured complex of elementary predicates, each with an array of arguments. One argument of the complex predicate may correspond to two or more arguments of the individual predicates. Linking rules then map from LCS to the PAS which function as a bridge between the LCS and syntactic structure. Expanded inflection fall within the ambit of the X-bar theory and various functional categories have been recognised. These categories establish dependency between parts of a sentence and they are represented as beads projecting X-bar phrases. Morphophonosyntactic formatives corresponding to a functional category become syntactically heads of maximal projections. These developments have affected external arguments such that the d -structure position of the subject has to be evaluated.

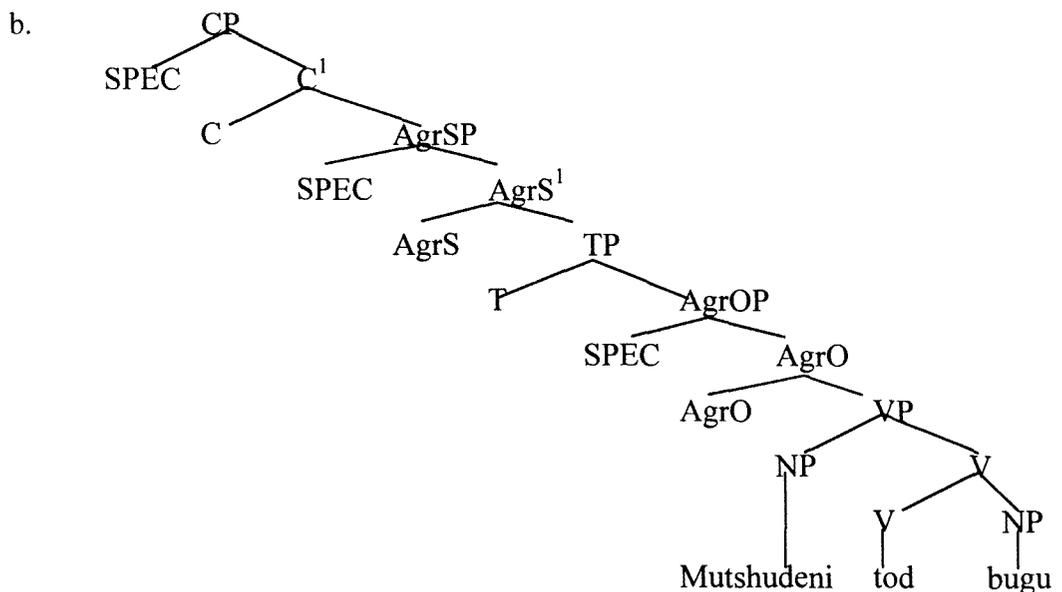
(91)



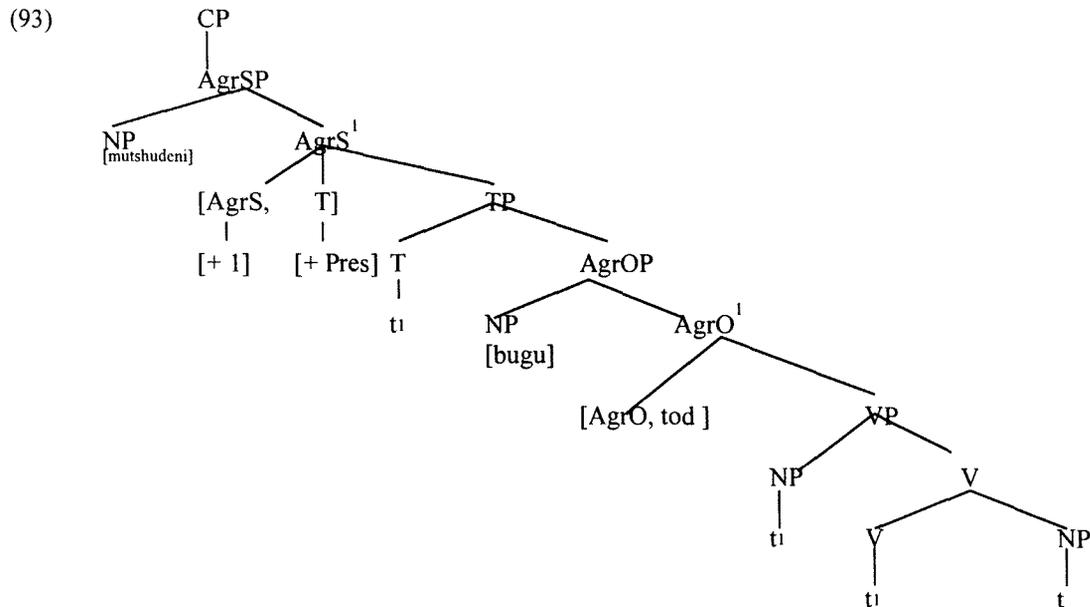
The NP* is the canonical or D - structure position of the subject. The NP[^] is the S-structure position of the subject. A subject generated in position NP* must move to position NP[^]. This external NP* argument has to be in that position as a conditions for theta assignment. A theta assignment relation between X and Y requires sister hood between X and Y: for a subject NP to receive a theta role it must have as a sister VP. In the above structure the NP will not receive any theta role because all the inflectional categories have been placed between NP[^] and the maximal projection of the verb and there is no sister hood between the NP argument and the VP. With the NP internal subject hypothesis this position has been refined (Chomsky) (1992).

(92) a. [Mutshudeni] u toda [bugu]

Student wants book



The above structure has the external NP argument appearing in the NP with its theta role assigned by the VP. The external NP argument [Mutshudeni] has been generated within the VP. It moved from this position to the SPEC position next to AgrS. The V and the internal argument NP [bugu] moved as well together with inflectional categories, leaving traces in the S- structure



According to Williams (1994) a number of distinguishable arguments A1 ... An, exist for each verb. William's theta theory is concerned with a relation between a verb and a noun phrases. The NP in a sentence must be an argument of a verb and the relation has three features:

(94) a. Obligatory relation

The NP in a sentence must be an argument of some verb.

1. [Maluta] u funa [Mulalo]

Maluta loves Mulalo

2. [Maluta u a funa]

3. [Pro,] u a funa

4. Ndi a vhea/ndi vhea [Malegere] [tafulani]

5. Ndi a la/ndi a la zwliwa

b. Unique Relation

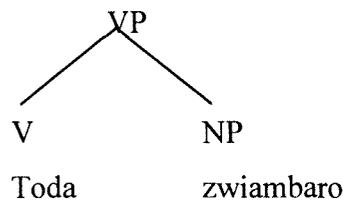
One NP may not be assigned two theta roles.

[Maluta] u toda [fhethu ha u dzula]

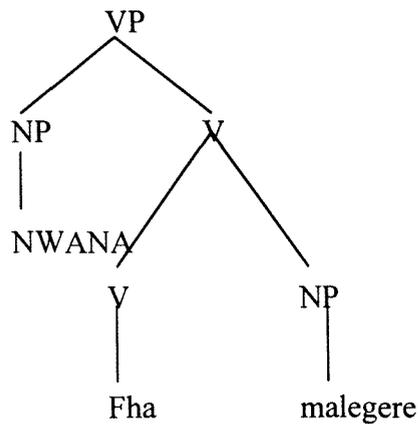
Maluta wants a place to stay

c. Locally structured relation

The verb and its arguments must be sisters



In (94a) the verb **funa** is a two place relation where the NPs Maluta and Mulalo are arguments of the verb **funa**. The NP must be an argument of a verb. In (94a2) the subject argument is **Maluta**, if it is missing (94a3) it may be represented by an empty pro with subjectival agreement and the missing subjects are interpreted as definite and non subject arguments (94a4) are optional. The object argument of a verb (94a5) may be optional. In (94a5) the object is missing in **ndi a la** but present in **ndi a la zwiliwa**. Missing objects are interpreted as generic; i.e. class of objects which can be eaten. In (94b) the unique relation exist because **Maluta** is assigned the subject argument while the NP [fhethu ha u dzula] is assigned the object argument by the verb **toda**. Here each phrase is assigned only one theta role. In (94c) the NP zwiambaro is a complement of the verb **toda** as well as its sister. The only difference is when a verb has two object argument NPs may not be sisters of the verb.



The verb **fha** has two object arguments; i.e. [Nwana] and [Malegere]. The NP argument **nwana** above is not a sister of V but of V'. The relation between **nwana** and V' is not local. A local relation exist between **fha** (V) and **malegere** (NP). **Malegere** is a sister to the verb **fha**. In this case M-command is necessary; a verb and its argument must be dominated by the same maximal projection. The VP [**fha nwana malegere**] is a maximal projection dominating the verb **fha** and its two arguments **nwana** and **malegere**.

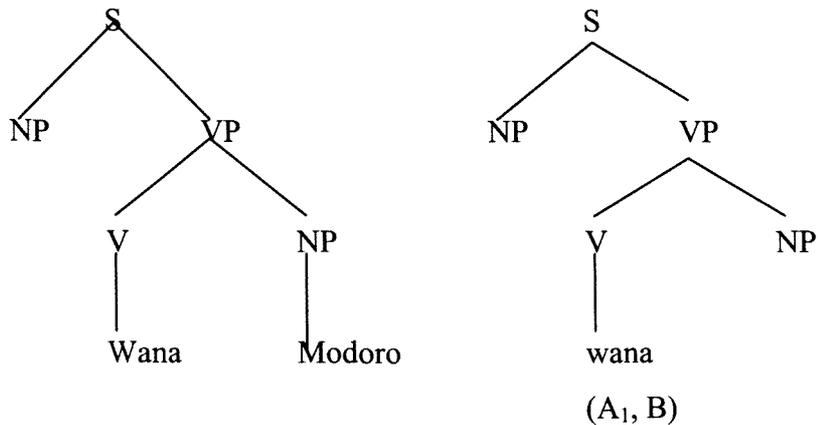
The lexical theta role assignment may take place from left to right (theta directionality parameter).

(96) Mbulaheni u nanga maswiri

The verb **nanga** assigns a theta role to its complement **maswiri**. The verb is on the left side of the complement **maswiri**.

The external argument: the subject argument is a sister of the maximal projection of the verb.

(97) a



Mukundi 97 a is the subject argument and this NP is a sister of VP which is the maximal projection of V. Mukundi is an external argument because it is located external to the maximal projection of the VP. The VP is a one place predicate because it is assigning only one theta role to the subject argument. The head of the VP is **wana**. But in 97 b the verb wana assigns two theta roles and **A** is the underlined external argument and **B** indicates internal argument.

3.1.4 Functional positions in a sentence

Grammatical functions in sentences are the subject, the object and the complement.

3.1.4.1 The Subject

There are various arguments which appear in the subject position.

3.1.4.1.1 An Agent

The subject argument is the willing agent of the action that participates internally in the action expressed by the predicate. The subject argument attributes the property of will or intention to the agent. The agent argument is the intender of the action denoted by the verb. The subject argument with intransitive verbs can be divided into two subcategories:

A. One argument of intransitive verbs with external agent.

- (98) 1. Mukalaha u a shuma
Old-man Agr-Lpres -Work-Pres
The old man is working
2. Nwana u a lila
Child Agr-Lpress-Cry-Pres
The child is crying

The non-motion verbs 98 have a subject argument which is agent.

b. Verbs with an internal argument

- (99) 1. Vhanna vha-do-livh-a mushumoni
Men Agr-Future-go-to-work
Men will go to work
2. Vhasadzi vha-ts-a doroboni
Women Agr-go-pres in-town
Women are going down to town

Verbs above in 99 have an internal argument which is a locative noun phrase .The locative noun phrase is not an object of the verb

3.1.4.1.2 An Experiencer

Thematic role of experiencer relates to mental states, to experience, feel or perceive some psychological state. Predicates with an experiencer are also known as psych verbs

- (100) a. **Mudededzi o-sinyuwa**
[0 < Agr + perf - o]
teacher Agr – perf - Angr - perf
The teacher is angry
- b. **Vhana vha - a - nala**
Children Agr - L - pres - sulk - pres
Children are sulking

3.1.4.1.3 A THEME

A theme argument occurs with a verb of motion or location where the terms motion or location can be concrete or abstract. There are three sub-categories of verbs which have a thematic role of theme in the subject position:

- (101) a. Weather verbs: in Tshivenda
no weather noun may appear in the subject position but the agreement morpheme **hu** occurs with all weather verbs.

Hu - a - penya
Agr L Press - flash - pres
It flashes

- b. Stative verbs: appear in the perfect tense
With a present tense meaning.

Goloi yo - **luga** [yo < Agr + o]
Car Agr - perf - ready
The car is ready

c. Location /Motion Verbs

Have two arguments; external argument is either an **agent** or a **theme**

And an internal argument interpreted as location

(102) (i) Ndo - dzula tshiduloni

I - sit - perf on - chair

I am sitting on a chair .

(ii) Vhurifhi vhu - **bva** - poswoni

Letter Agr come from at Post Office

The letter comes from the Post Office

3.1.4.1.4 A Patient

Patient is affected by action and have two sub categories, namely:

(103) a Event verbs:

Mulilo u a **duga**

Fire Agr - L Pres - Bum - Pres.

The fire is burning

b Body Parts Subjects:

Zwilonda zwo **fhola** [zwo < Agr to]

Wounds Agr – Perf - heal

The wounds are healed

3.1.4.2 The Object

The object has various syntactic properties.

A. Word Order

In Tshivenda the object always follows the verb. Tshivenda is an SVO language.

(104) [Mukundi] [u la] [maswiri]
S V O
Mukundi eats oranges

In the above sentence [Mukundi u la maswiri) the object is [maswiri)

B. objectival Agreement

The objectival agreement may appear on the verb in transitive verbs.

(105) Ndi nga -mu -rwa Mutukana
I can beat him the boy

In the above sentence *Ndi nga - mu - rwa mutukana* the object **mutukana** appears together with its agreement **mu** on the verb.

c. Passive Sentences

The object may be moved to the subject position in passive sentences

a. [Aya masimu] a a linwa

These fields are ploughed

The object aya Masimu has been moved to the subject position in 106

d. Ditransitive verbs

Ditransitive verbs appear in predicate which have two internal arguments. The two internal arguments of ditransitive verbs have the grammatical function of direct or indirect object.

(107) Mukundi u nea Thendo] [Maswiri]

Mukundi gives oranges to Thendo

Considering the word order of the objects Thendo and Maswiri, the indirect object is the noun phrase Thendo. The direct object maswiri then follows the indirect object Thendo. In Tshivenda in surface structures the direct and indirect object may be interchanged without any difference in meaning.

(108) a Mukundi u nea Thendo Maswiri

b Mukundi u nea maswiri Thendo

This type of interchange does not give rise to difference in meaning. The indirect object Thendo has the feature [+ animate] while the direct object maswiri is [-animate]. Inanimate NPs are usually not recipients. But when both objects are [+ animate] the word order cannot be changed and it must remain as it is:

(109) Khotsi u fila [musidzana] [munna]

Father gives the girl a husband

If munna becomes the recipient meaning of (109) above can be changed. In ditransitive verbs a predicate may have one direct object, through argument reduction.

(110) Bannga i hadzima [tshede]

The bank is lending / borrowing money

Argument reduction of the direct object is not common but verbs may be found with only an indirect object.

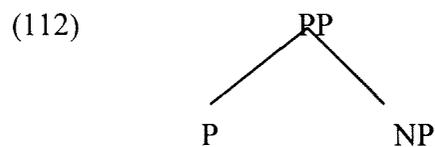
- (111) Ndi khou hama kholomo [mafhi]
I am milking the cow

3.1.4.3 Complement of a preposition

In Tshivenda there are prepositions: nga, na and kha.

3.1.4.3.1 The preposition NGA-

The preposition nga - can combine with any noun phrase in a structure like the following:



Noun phrases will appear as complements of the preposition **nga-**

- (113) a. [np DEM N]
A tho ngo mu vhudzisa tshithu [nga havho munna]
I didn't ask him a thing about that man!
- b. [np[npDEM N] NP]
O vha a siho [nga duvha helo la nndwa]
He was not present on that day of the fight.
- c. [np [np DEM N] S]
Ndi toda u mu tsivhudza [nga hetshi tshithu tshe a da nga tsho]
I want to advice him about this thing which he came about:

3.1.4.3.2 Preposition NA

a. [NP DEM N]

(114) Vho vha vho da [na heyo mbudzi]

b. [np [np DEM N] NP]

(115) O vha o da [na tshelede heyo ya malume]

c. [np [npDEM N] S]

d.

(116) Ndi to da u mu rwa [na hetsho tshigidi tshine a vhulaha nga tsho]

3.1.4.3.3 Preposition Kha

a. [NP DEM N]

(117) Tshinoni tsho kavha [kha hoyo muri]

b. [NP [NP DEM N] NP]

(118) Vho vha vho edela [kha nndu heyo duvha la lupenyo]

c. [NP [NP DEM N] S]

(119) Ndi toda u mu vhulaha [kha heyo nndu a dzula kha yo]

3.1.5 CASE THEORY

Case is a nominal category controlled by syntax and expresses relations between nouns and nouns and verbs in a phrase. In Tshivenda case is marked by free adpositions. Mukundi u nwalela khotsi vhurifhi nga penisela nduni.

(120) [Mukundi] appears in the nominative to indicate its syntactic role as subject. [Vhurifhi] is in the accusative to indicate that it plays the role of direct object. [Khotsi] is in the dative because it is the indirect object. Penisela and nduni are in the instrumental, the former because it represents the means of writing and the letter because it represents the place at which the letter was written. In Tshivenda case is never expressed by simple affixation. It is a set of categories which are determined solely by syntax and has nothing to do with the lexicon.

3.1.6 MOVEMENT

3.1.6.1 Movement for the passives

The passive affix - w - can be added onto the verb, leaving the subject position empty; designated as pro. The agreement in inflection will then receive the feature [+ existential] which will be spelled out as hu. This agreement is then coindexed with the empty pro so that both may have the feature [existential]

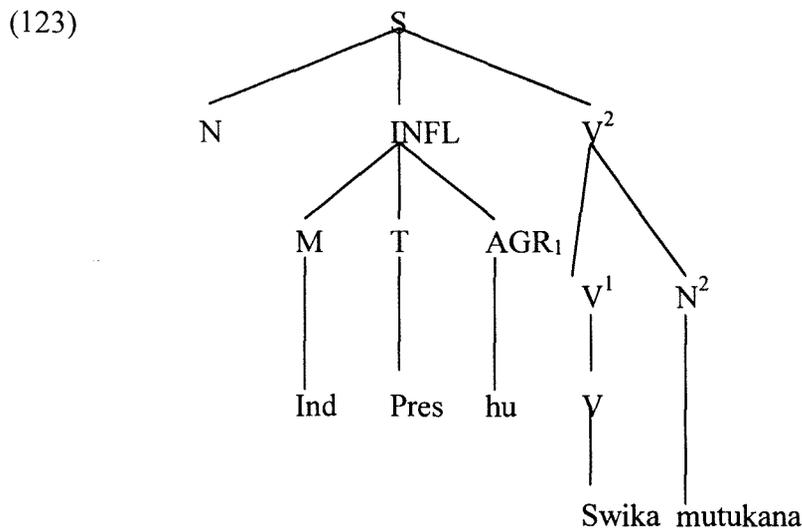
(121) Hu - a - linwa
Agr : L Pres - work - pass - press
There is being ploughed

3.1.6.2 Subject Inversion

The subject in a sentence with intransitive verb may be moved to a post -verbal Adjunct position:

- (122) a. Mutukana u a swika
The boy arrives
- b. Hu swika mutukana
There enters the boy

A movement transformation adjoins *mutukana* to the VP. **Hu** - with the feature [existential] is then inserted under agreement of inflection. This adjunct **Mutukana** and *pro* with **hu** - constitute the members of a chain and they are coindexed. The agreement and *pro* are coindexed with the feature [existential] and *pro* and **mutukana** are coindexed through syntactic indexing indicating a chain for case and theta role assignment.



Mutukana has been entered in an adjunct position where the N^2 is a sister of V^1 the typical adjunct position. The subject **pro** has been left empty after the movement of **mutukana**

3.2. MORPHOLOGY OF THE NOUN

3.2.1 Nominal inflection

The discussion of nominal inflection will include case, noun class, agreement and number

3.2.1.1 Case

for the discussion involving case refer to 3.1.5.

3.2.1.2 Noun class

In Tshivenda noun class is a nominal category Gender is not a nominal category in Tshivenda but only It noun class. The noun classes of Tshivenda are the following:

(125)	1.	[MU-]	MU-THU
	2.	[VHA-]	VHA- THU
	1A.	[-]	KHOTSI
	2A.	[VHO -]	VHO-KHOTSI
	3.	[MU -]	MU- RI
	4.	[MI -]	MI -RI
	5.	[LI -]	LI -IVHA
	6.	[MA -]	MA -IVHA
	7.	[TSHI -]	TSHI -NONI
	8.	[ZWI -]	ZWI -NONI
	9.	[N -]	KHOLOMO
	10.	[DZI -N -]	KHOLOMO
	11.	[LU -]	LU-FHANGA
	14.	[VHU -]	VHU- LUNGU
	15.	[U-]	U-SHUMA
	16.	[FHA-]	FHA-SI

- | | | |
|-----|---------|-------------|
| 17. | [KU -] | KU- LE |
| 18. | [MU-] | MU-RAHU |
| 20. | [KU -] | KU- DU |
| 21. | [D1 -] | D1 -KHOLOMO |

3.2.1.2. Morphological structure of Tshivenda nouns

In Tshivenda nouns may be recognised by prefixes as it is indicated below.

3.2.1.2.1 Class 1/2: MU / VHA

(126) **Mu** - tukana (boy) and **Vha** - tukana (boys)

Mu - nna (man) and **Vha - nna** (men)

In forming plurals, the plural morpheme **Vha-** replaces the singular morpheme **Mu-**

3.2.1.2.2 Class 1a / 2a : θ/vho

These classes represent titles of people and respect of elderly people.

(127) θ Malume (uncle) and Vho - Malume (uncles)

θ Khotsi (father) and Vho - khotsi (fathers)

These two classes are the subclasses of classes 1/2 : Mu/Vha singular nouns which start with Ma -, ne, nya and ra -are found in these subclasses.

(128) a. Mavhulaise > vho ma mavhulaise (cause other people's death)

b. Nemudi > vho nemudi (owners of the village)

c. Nyanenge > vho nyanenge (beloved one)

d. Ramabindu > vho ramabindu (businessmen)

Class I a = θ has no singular class prefix , but class **2a** has a plural class prefix Vho. The plural morpheme Vho is affixed to the nouns of class 1 a to make them plurals.

3.2.1.2.3 classes 3/4 Mu/Mi

These classes embrace natural things

(128) **Mu-** roho (vegetable) and **Mi-** roho (vegetables)

Mu- shumo (work) and **Mi-**shumo (works)

These classes are formed by the singular classes prefix **Mu** -and the plural classes prefix **Mi**-.The plural noun is formed by replacing the singular class prefix **Mu-** by the plural class prefix **Mi**-.

3.2.1.2.4 Classes 5/6 : Li / Ma

These classes cover things which are either concrete or abstract.

(129) **Li-** ivha (dove) and **Ma-** ivha (doves)

Duvha (day) and **Ma-** duvha (days)

The singular class is indicated by a class prefix **Li-**. There are nouns which do not have visible class –prefixes such as duvha. But in plural they take **Ma-**. The plural is formed by the plural class prefix **Ma-**. The plural class prefix **Ma-** can be used in forming plural nouns in other singular classes such as:

- | | |
|---------------|--|
| (131) Class 9 | θ khotsi (chief) > Ma- hosi (chiefs) |
| Class 11 | Lu - nwalo (letter) > Ma - nwalo (letters) |
| Class 14 | Vhu - tshilo (life) > Ma - tshilo (lives) |
| Class 20 | Kutshimbilele (manner of walking) |
| > | Matshimbilele (manner of walking) |

Class 21 **Di-** thu (agre person) > **Ma** -dithu
(Ogre people)

3.2.1.2.5. Classes 7/8: **Tshi/Zwi**

(132) **Tshi** - noni (bird) and **Zwi** - noni (birds)
Tshi - du (small house) and **Zwi** -du (small houses)

These classes sometimes indicate small things. The plural nouns belonging to these classes are formed by replacing the singular class prefix **Tshi** – by the plural class prefix **Zwi**-.

3.2.1.2.6 Classes 9/10: **N/DZ** – **N**

These classes indicate names of animals and things.

(133) **θ** kholomo (cattle) and **Dzi** -kholomo (cattle)
N - ngu (sheep) and **DziN** -ngu (sheep)

These classes represent nouns which occur with or without singular class prefix **N**- .The plurals are formed by the plural class prefix **Dzi**-.

3.2.1.2.7 Class 11: **LU**

(134) **LU**- fo (wooden spoon)
LU- fhang (knife)
LU- tanda (small things)

The plurals of noun in this class are formed by the plural class prefixes **Dzi**- (10) and **Vhu**- (14) respectively.

- (135) Lu - fo > **dzi - mpfo**
Lu - tanda > **Vhu - tanda**

3.2.1.2.8 Class 14: Vhu -

- (136) **Vhu - tshilo** (life)
Vhu - dele (cleanliness)
Vhu -tali (wisdom)

This class forms the plural of other classes.

3.2.1.2.9 Class 15: U-

- (137) U bika (looking)
U lima (ploughing)

It is the infinitive class.

3.2.1.2.10 Classes 16/17/18 : Fha-1 Ku-1 Mu-

- (138) **Fha - si** (down)
Ku - le (far)
Mu - rahu (back)

These are locative classes

3.2.1.2.11 Class 20: Ku

- (139) **Kwana** (small child)
Ku - budzi (small goat)

This is a diminutive class. Nouns in this class are formed by affixing class prefix **Ku-** to other nouns. The plural nouns are formed by the plural class prefix **Zwi-**.

- (140) Kwana > Zwana
Ku - budzi > Zwi – budzi

3.2.1.2.12 Class 21 : Di

- (141) Dana (ogre of child)
Dithu (ogre of person)
Dikolomo (big cattle)

This is an augmentative class. The plural nouns are formed by the plural class prefix **Ma-**

- (142) Dana > Madana
Dikolomo > Ma - di – kolomo

This class has two ways of forming plurals. There are plurals nouns formed from those singular nouns without visible class prefix Di - like dana. And there are those plurals nouns formed by those singular nouns with double plural class prefixes:

- (143) Dana > Ma - dana
Dithu > Ma -di – thu

The plural nouns are formed by the plural class prefix Ma- (cl.6)

3.2.13 AGREEMENT

In Tshivenda Agreement (an inflectional category) is found as a category in the following instances:

Subjectival agreement (AgrS)

- (144) [Vhatukana] vho swika
(Boys they -came)

Objectival agreement (AgrO)

- (145) Ndi a v ha vhona [vhasadzi]
I see them the women.

Prepositional agreement.

- (146) Ndi do tuwa ngayo .[goloji]
I will go with it the car

Nominal modifiers

Nominal modifiers in Tshivenda have some form of agreement with their head nouns. Noun classes which show AgrS, AgrO and Agreement with nominal modifiers such as demonstratives and adjectives in Venda.

- (147) Class
1. [Uyu mutukana mulapfu] o mu -toda [uyu musidzana]
 2. [Avha vhatukana vhalapfu] vho vha toda [avha vhasidzana]
 3. [Uyu muri mulapfu] wo i a nwa [iyi mitshelo]
 4. [Iyi miri milapfu] yo I anwa [iyi mitshelo]
 5. [Ili linngo litswuku] lo li dadza [ili pulasitiki]
 6. [Ayo manngo matswuku] o a dadza [aya mapulasitiki]
 7. [I tshi tshinoni tshitswu] tsho tshila [itshi tshikoli]
 8. [Izwi zwinoni zwitswu] zwo zwila [izwi zwikoli]
 9. [Iyi nngu tshena] yo i thukhula [iyi thambo]
 10. [Idzi dzinngu tshena] dzo dzi thukhula [idzi dzithambo]

- 11 [Ulu lusunzi lusekene] lwo lu luma[ulu luvhuda]
- 14 [Uvhu vhusunzi vhusekene] ho vhu luma[uvhu vhuvhuda]
15. [Uhu U-tshimbila] ho - mu - netisa
16. [A fha fhasi] hu - a - fhisa
17. [Ngei kule] huh- ah- bvafhisa
18. [Afho murahu] hu - a - ofhisa
20. [Uku kuduna kudenya] kwo - ku - vhulaha [kumange]
21. [Ili dithu lihulu] lo - li-mila [linowa]

Agreement is not an inherent category but a derived category determined by noun class. It is an inflectional (syntactic) rather than a lexical category.

3.2.1.4 Number

The common functions of number are singular, dual and plural. In Tshivenda number is another grammatical category generally taken to be inflectional and number is never marked with a free morpheme but by a bound morpheme .Number is reflected in agreement and it is referential only among nouns. In Tshivenda number can be discerned in the following instances as an inflectional category:

- (148) a. In agreement of the first and second person with an empty pro:

[pro;] Ndi a u toda [pro ;]

I want you

[pro;] Ri a vha toda [pro;]

We want you

- b. In imperative plurals

[pro;] tuwa - ni;

c. A model of nominal inflectional categories

Nominal categories in Tshivenda are divided into three inflectional and one lexical category. Case is determined solely by syntax and has nothing to do with the lexicon. Agreement interprets the lexical category, Noun class. Number is an inflectional category which appears with an empty pro. Agreement is required in syntax given the fact that noun class is lexical, not a syntactic category.

3.2.2 LEXICAL DERIVATION

Lexical derivation is found in the lexicon. In lexical derivation we find **noun class** and **human** as features. Lexical derivation operates on the grammatical category features which are found in the lexicon. There are four different types of lexical derivations in Tshivenda, namely: feature value switches, expressive derivations, transposition and functional derivations. For our scope of study I will concentrate on feature value switches and expressive derivations.

3.2.2.1 FEATURE VALUE SWITCHES

Du Plessis (1997 : 109) states that the most common type of L. derivational rule required of a description of natural languages is a simple SWITCH or TOGGLE, which resets the :1: or other values of grammatical features. This type of rule no doubt operates both in syntax and the lexicon, determining the values on the features of the I -categories as well as the values of inherent lexical features. Feature value switches may be expressed by nominal affixes.

3.2.2.1.1 FEATURE VALUE SWITCHES AND GENDER

Languages generally have a rule which derives feminine correlates from unmarked Masculine nouns, for example Tshivenda **Ndou** 'elephant' (unmarked masculine) | Ndoukadzi (female elephant). In Tshivenda marked masculine, such as **Munna** whose

gender cannot be changed are lexically subcategorised as [- feminine + masculine] to distinguish them from UNMARKED masculine like **MUTSHUDENI** 'student' which bear the features [+ feminine + masculine]. Marking helps to distinguish those Masculine nouns which may be feminised from those which may not and it also characterises the semantic potential of unmarked Masculine to refer to both females and males. If unmarked masculine nouns have both their feminine and Masculine value set as " + " [+ Feminine + Masculine. Languages must share a rule of feminisation such as [+ feminine + Masculine] → [+ feminine -masculine]. In other words, this rule optionally resets the value of [+ masculine] in the context [+ feminine] to [-masculine. For any category A Jackson (1932,1939), a marked of that category signifies the presence of A while the unmarked term makes " no statement of " A. The unmarked grammatical form of A refers indiscriminately to A and B and while the marked grammatical form refers exclusively to B. The unmarked form is then the default form and gender in Tshivenda may be indicated in the following ways:

(149) a. different words for male and female

[+ male -female]	[-male + female]
munna	musadzi
(man)	(woman)
khotsi	mme
(father)	(mother)

b. female with **-kadzi**

[+ male, female]	[-male + female]
Ndou	ndoukadzi
(male elephant)	(female elephant)
mudededzi	mudedekadzi
(teacher)	(lady teacher)

- c. Female with **- nyana**
Nwana + nyana > nwananyana (girl)
- d. Male with **- lume**
Malume (uncle)
Tshitundulume (lavish spender)

3.2.2.1.2 FEATURE VALUE SWITCHES AND LOCATIVE

The locative nouns in Tshivenda are divided into three classes. The old locative class nouns of classes 16,17 and 18 which do not function with their prefixes any longer, locative nouns with a suffix -Ni and various place names.

- old locative class nouns: fhasi, kule and murahu
- locative affix -Ni : thavha –Ni
- place names: Tshavhalovhedzi

These locative nouns above are the heads of a maximal projection locative noun phrase [NPloc] in syntax. The feature value switch is [:1: locative]:

(150) [-locative]	[+ locative]
Makhado	Makhado
Dorobo	doroboni

Feature value switches ,i.e. gender and locative, may be expressed by nominal affixes.

3.2.2.2 EXPRESSIVE DERIVATIONS

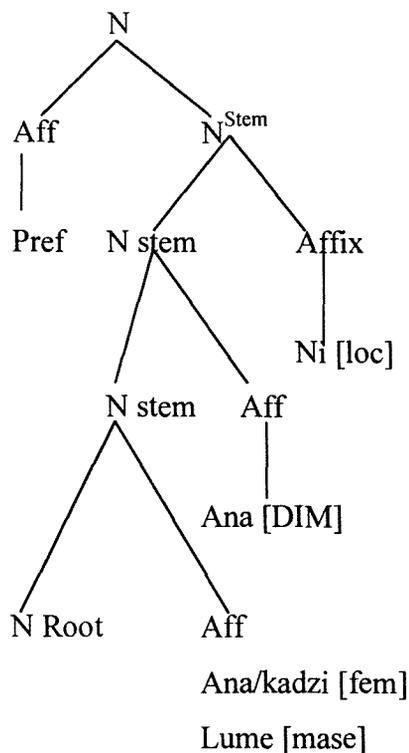
Expressive derivation reflects at least five functions namely: Diminutive, Augmentative, Pejorative, Affectionate and Honorific. They do not change the meaning or lexical class of the lexemes over which they operate: but generate nouns from nouns; verbs from verbs

and adjectives from adjectives. Expressive derivations express prejudices of the speaker as to whether the referent is SMALLER, LARGER, MORE LIKEABLE or MORE THREATENING than other member of its semantic category. The categories of expressive derivations are the before not like other grammatical categories; while Augmentative and Diminutive may express such functions as LARGE (x) and SMALL (x) they do not seem to mop onto grammatical functions because expressive derivations are always optional and subjective. Reference could be made to a person of normal size with a diminutive in one context and an Augmentative in another context with emotional effect. Pejorative noun variants indicate the speaker's dislike of reference while an Affectionate form reflects the speaker's favour. Most often, the Affectionate and pejorative forms conflate with diminutive and Augmentative forms respectively. Expressive derivations may be expressed by nominal affixes.

3.2.2.3 NOMINAL AFFIXES IN VENDA

Nominal affixes in Tshivenda expressed feature value switches. (i.e. gender and locatives) and expressive derivations.

(151)



3.2.2.3.1 GENDER AFFIXES: -KADZI/ANA/ -NYANA

In Tshivenda there are two gender affixes in which nouns can be classified, i.e. FEMININE and MASCULINE. Each has affixes by which they are identified, e.g. Feminine affixes -kadzi, -ana and -nyana and the masculine affix -lume. Underneath follow nouns (feminine and masculine affixes).

(152) a. Nouns with feminine affix –kadzi

Mufumakadzi (woman, wife, lady)

Tshilikadzi (widow, deceased woman)

Ndoukadzi (female elephant)

Khulukadzi (can which has had several calves)

b. Nouns with feminine affix- ana

Tshibwanana (female beast without horns)

Khunwana (female ,hed all, head of cattle)

Tshekana (black and white she -goat)

Tswana (female had ofcattle,black allover)

c. Nouns with feminine affix – nyana

Nwana + -nyana > nwananyana (female child, girl)

d. Nouns with masculine affix -lume

Malume (uncle)

Tshitundulume (lavish spender)

Ndoulume (elephant bull)

Gukulume (cock)

3.2.2.3.2 DIMINUTIVE AFFIXES: -ANA/ -NYANA

Diminutive nouns can be expressed by the following affixes:

(153) a. nouns with diminutive affix- ana

mutukana (small boy)

mbohwana (young bull)

thavhana (small mountain)

mulambwana (rivulet)

b. Nouns with diminutive affix – nyana

Muthothonyana (a bit of gravy)

Dzemanyana (a little habit)

Swigiri nyana(a small glanting of sugar)

Musidzanyana (small young girl)

3.2.2.3.3 LOCATIVE AFFIX -NI

Locatives can be expressed by the following locative nouns: murini (at/to under the tree)

(154) Mutukani (to at the young boy)

Kholomoni(to/at the head of cattle)

Dakaloni) at/during happiness)

3.2.2.4 NOMINAL FUNCTIONAL DERIVATION

According to du Plessis (1997 : 157) morphology has certain means to specify relations among arguments in an argument structure and one such relation is control. The suffixes

[-I] and [-o] are found in nouns derived from verbs. These suffixes specify control of an argument of the predicate.

(155) [Mu -shum -I]
 [x, y] [+ human]

The suffix [-I] , represented above by [x] specifies control of the external argument of the predicate.

(156) [Mu- shum - o]
 [x, y] [-human]

The suffix [-o] represented above by [y] specifies control of the internal argument of the predicate. In Tshivenda, noun class prefixes have certain contributions to the derivation both the noun class prefix and the suffixes [-i] and [-o] have to be taken into account when dealing with deverbatives. Such derivations may then be classified in two major categories with the feature [+ human] where [+ human is spelled out as [I] and [-human] as [-o]. The following features have to be established when dealing with the noun class prefix in deverbatives.

3.2.2.4.1 NOMINAL DERIVATIONS FROM VERB WITH SUFFIX [-I]

(157) a. person names: class /
 mushumi worker
 mubiki cook
 Affixes: Mu- : definite person name
 : human

3.2.2.4.2 NOMINAL DERIVATIONS FROM VERBS WITH THE SUFFIX [-0]

(158) a. Class 3 : event/manner

mushumo : work

murungo : sewing

Affixes: mu-: definite event, manner

-o : [-human]

b. Class 7: theme (concrete)

Tshigwevho : judgement

Tshirendo poem "

Affixes: tshi-: definite theme

-o : [-human]

c. Class 9 : Result / Resultative (abstract)

Ngudo: lesson

Khano: harvest

Ndimu: ploughing

Affixes: n-: definite result/resultative

-o : [-human]

d. Class 11 : abstract

Lufuno : love

Affixes: Lu-: definite abstract

-o : [-human]

- d. Class 11 : event

Lumalo : marital gift

Lugungulo: grumbling

Affixes: Lu- : definite event

-o : [-human]

3.2 SEMANTICS OF THE NOUN

3.3.1 Levels of representation

There are four levels of representations in a generative lexicon (computational system)

- (159) a. Argument structure: specification of number and type of logical arguments, and how they are realised syntactically.
- b. Event structure: Definition of the event type of a lexical item and a phrase. Sorts include STATE process and TRANSITION and events may have sub-eventual structure.
- c. Qualia structure: Modes of explanation composed of Formal, Constitutive, Telic and Agentive role.
- d. Lexical Inheritance structure: Identification of how a Lexical structure is related to other structures in the type lattice and its contribution to global organization of a Lexicon.

The four levels of representations are connected by a set of generative devices. These generative operations have semantic transformations all involving well formedness conditions on type combinations namely:

- (160) a. Type coercion: where a lexical item or phrase is coerced to a semantic interpretation by a governing item in the phrase, without change of its syntactic type.
- b. Selective Binding: Where a lexical item or phrase operates specifically on the substructure of a phrase without changing the overall type in the composition.
- c. Co-composition : where multiple elements within a phrase behave as functors, generating new non-lexicalized senses for the words in composition. This also includes cases of underspecified semantic forms becoming contextually enriched, such as manner composition failure transcription and light verb specification.

The three semantic transformations help us to capture the semantic relatedness between syntactically distinct expressions. Argument, event and qualia structures must conform to the well-formedness conditions defined, by the type system.

3.3.1.1 ARGUMENT STRUCTURE

Du Plessis (1999 :281) states that the argument structure for a word can be seen as a minimal specification of its lexical semantics. Lexical items, nouns in particular could have four types of arguments, namely:

- (161) a. True Arguments: syntactically realized parameters of the lexical item
- b. Default Arguments: Parameters which participate in the logical expressions in the qualia , but which are not necessarily expressed syntactically.

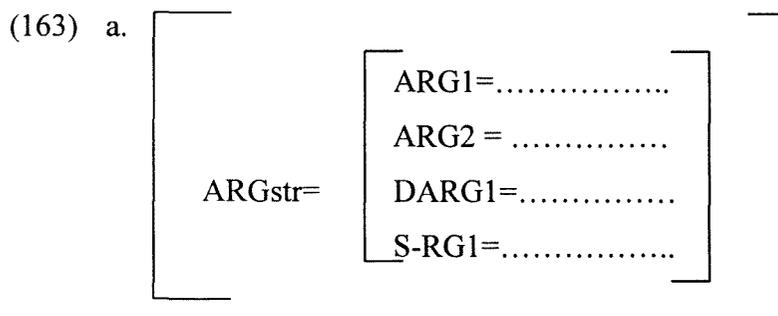
- c. Shadow arguments: Parameters which are semantically incorporate into the lexical item. They can be expressed only by operational of subtyping or discourse specification.
- e. True adjuncts : Parameter which modify the logical expression but are part of the situational interpretation, and are not tied to any particular lexical items semantic representation including adjunct expressions of temporal or spatial modification.

3.3.11.1 TRUE ARGUMENTS

True arguments define parameters which are necessarily expressed at syntax.

- (162) [Idzi mbudzi] dzo nona
 'These goats are fat'

The arguments for a lexical item e.g. ARG1, ARG2, ARGn are represented in a list structure where argument type is directly encoded in the argument structure i.e. ARGstr where Dr ARG is a default argument and S.ARG is a shadow argument.



The argument structure of nouns depends on the number of different senses which a specific noun may have. Nouns such as the following will appear with only one argument:

- (164) Tshidula (frog) [ARG1= animal]
Dafula (table) [ARG1 = artifact]

Nouns may have more than one argument, i.e. they may have two or more different senses. The noun [dzingadevhe] may refer to a deaf individual or a small insect

- (165) [ARG 1 = animate, individual]
[ARG2 = animate, insect]

3.3.1.1.2 DEFAULT ARGUMENTS

- (166) a. Nda vhumba ndou nga mavu/vumba
I formed an elephant out of clay/mud
- b. Ndi thata nndu nga matombo
I built a house out of stones

The above is an indication of the material product/alternation. The material (maw, matombo) is optional, its status as an argument is different from the created object (ndoul nndu). The material! product pairs above are called default arguments. They are necessary for the logical well - formedness of the sentence, but may be left unexpressed in the surface syntax. Default arguments can be satisfied by full phrasal expressions such as a pp with nga or as a phrase incorporated into a true argument.

- (167) a. Nda vhumba ndou ya mavu/vumba
I formed a wooden elephant
- b. Nda fhata nndu ya matombo
I built a stone house

Verbs like swika (arrive) may also appear with default arguments

(168) Nda swika [hayani]

I arrived at home

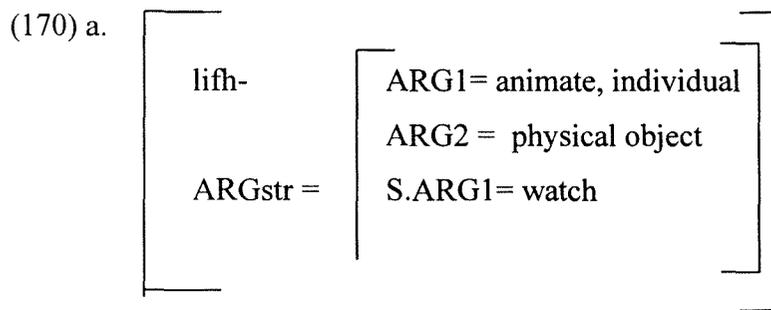
The locative phrase [hayani] is an example of a default argument.

3.3.1.1.3 SHADOW ARGUMENTS

(169) Ndo lifha watshi [nga mbudzi]

I paid for the watch with a goat

Appearing within brackets, shadow arguments refer to semantic content that is not necessarily expressed in syntax and are expressible when the expressed arguments stand in a sub-typing relation to the shadow argument. These arguments are expressed in very specific conditions. In above the paying of the watch could have been done by anything but the specific instrument is a goat and not e.g. Money. These shadow arguments may be represented as follows with the verb lifh- in above as an example.



3.3.1.1.4 TRUE ADJUNCTS

(171) Nwana o - kokovha [nduni]

Appearing in brackets, the true adjuncts are associated with verb classes and not individual verbs. The ability of these verbs to be modified by locative modifier as in (171) above is inherited by virtue of the verbs classification as an individual event.

3.3.1.2 QUALIA STRUCTURE

Qualia structure is the relational force of a lexical item; where a generative lexicon analyses all lexical items as relational to a certain degree. The manner in which this property is expressed functionally will differ from category to category and between semantic classes. It specifies four essential aspects of a word's meaning (or qualia):

- (172) a. Constitute: The relation between an object and its constituent parts;
- b. Formal: That which distinguishes it within a larger domain;
- c. Telic: Its purpose and function;
- d. Agentive: Factors involved in its origin;

There are two general points that should be made concerning qualia roles:

- (173) a. Every category expresses a qualia structure;
- b. Not all lexical items carry a value for each qualia role;

The first point deals with how a generative lexicon provides a uniform semantic representation compositionally from all elements of a phrase. The second point allows us to view qualia as applicable to particular semantic classes.

- (174) a. Vothi (door)
- [formal = property]
- [constitute = phys obj]
- [Telic = open/close]
- [Agentive = create]

- c. Halwa (beer)
 - [formal = liquid]
 - [constitute = intoxicating]
 - [Telic = drink]
 - [Agentive = brew]

- c. Dzingandevhe (deaf person)
 - [formal = physically disadvantage]
 - [constitute = deaf]

- d. Vhuswa (porridge)
 - [formal = food]
 - [constitute = mass]
 - [Telic = eat]
 - [Agentive = cook]

Lexical items may have ability to cluster multiple senses. This is referred to as lexical conceptual paradigm (lcp). The lcp indicates that something inherent in the semantics of a noun such that it is able to project any of three separate senses of the noun in distinct syntactic and semantic environments and it characterizes a lexical item as a meta -entry. Nouns such as newspaper appear in many semantically distinct contexts and kp treats them as logical expressions of different aspects to the meta-entry for newspaper. LCP indicates clearly that syntactic information is inheritable between lexical items. a lexical item associated with an lcp will have three senses. For the noun can construction, consider the following three senses.

- (175)
- a. The house; construction was finished in two months.
 - b. The construction was arduous and tedious
 - c. The construction is standing on the next street.

The above sentences make reference to the process and the result of the process respectively. Nouns like the following ,may have various senses:

- (176) a. thavha (mountain)
[mountain, large heap –Icp]
- b. dzingandevhe
[deaf individual, insect –Icp]

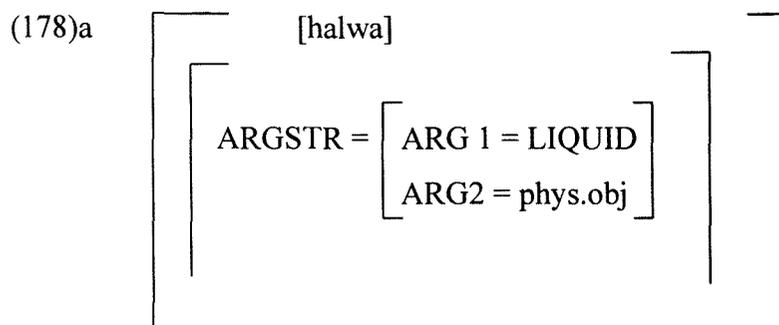
3.3.1.2 LEXICAL INHERITANCE STRUCTURE.

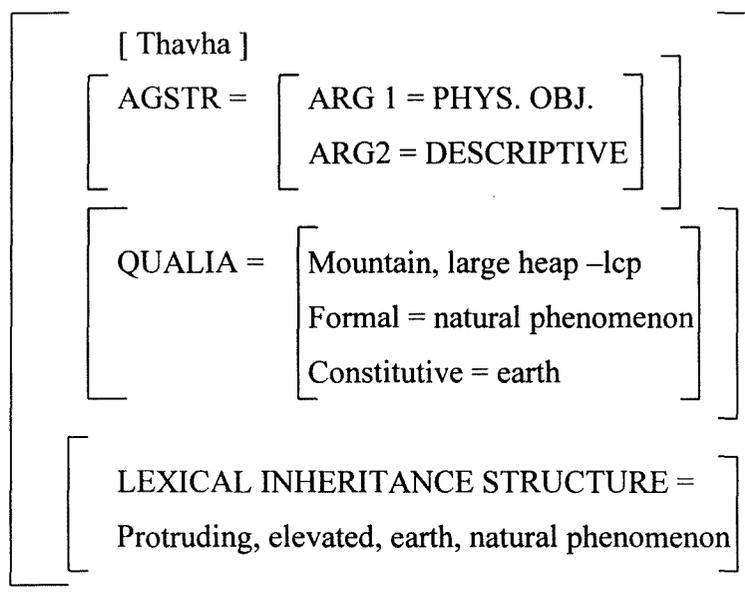
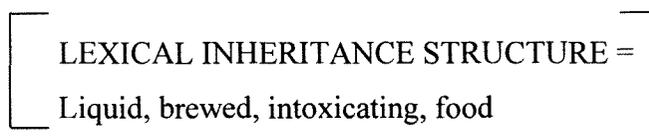
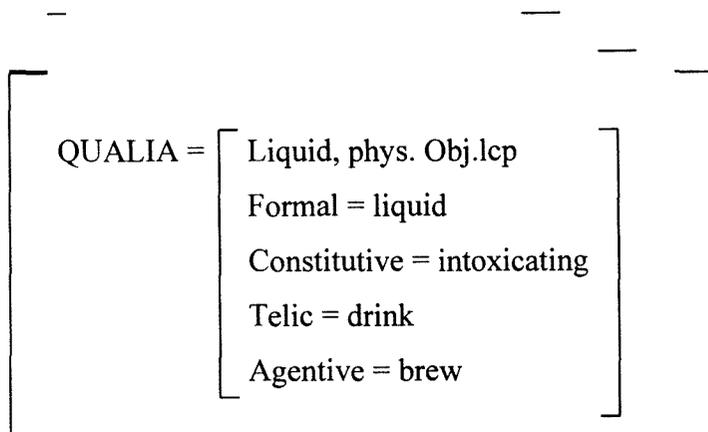
Semantic concepts are organized from specific level to generic level. Nouns have a relative small number of generic concepts.

- (177) a. Veini (wine)
Liquid, intoxicating, food
- b. Nngwe (leopard)
Carnivore, wild animal, vertebrate, animal
- c. maine (doctor)
doctor, professional traditional, medical worker, person.

3.3.1.4. THE INTERACTION OF SEMANTIC LEVELS

Below is an integration of four levels of representation i.e. argument, event, qualia and inheritance, into one system for lexical semantic representations.





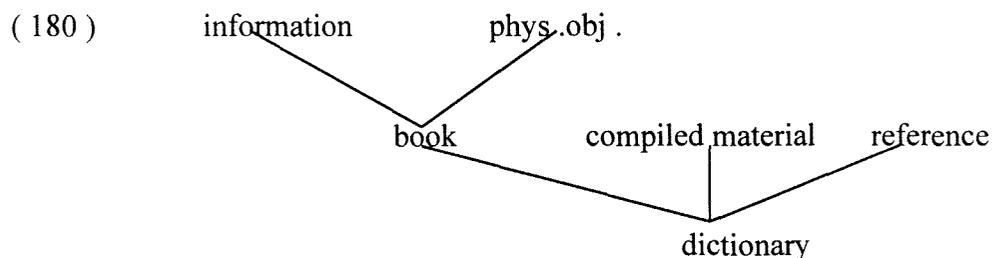
3.3.2. THE SEMANTICS OF NOMINALS

Lexical items may inherit from multiple parents. A scheme which allows multiple inheritance is thus needed.

Firstly it will be necessary to use inheritance relations which appear as an equative relation between two objects.

- (179) a. A book is a physical object.
 A book is information.
- b. A dictionary is a book.
 A dictionary is reference.
 A dictionary is compiled material.

These sorts of inheritance may be indicated as follows:



The above is an indication of descriptive type of relations without a theory to explain how to assign structure to lexical items or to specify lexical relations between lexical items in terms of links between only certain aspects of their representative lexical structures. A lexical item inherits information according to qualia structure it carries. Secondly a qualia structure is needed together with inheritance relation stipulated above.

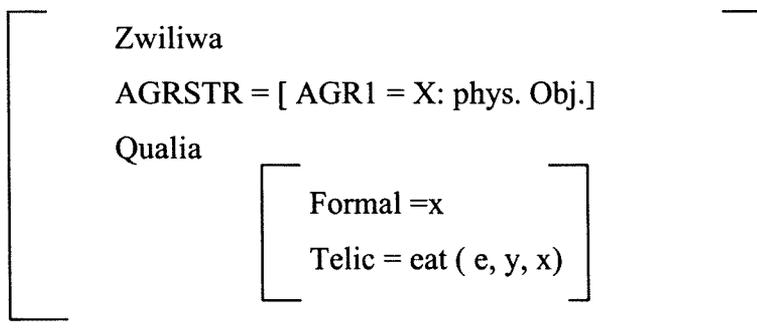
- (181) a. Book is formal.
 Book is telic.
 Book is agent.
- b. Dictionary is formal.
 Dictionary is telic.
 Dictionary is agent.

And thirdly inheritance through qualia will then give the following relations.

- (182) a. Bugu is formal physics object.
 Bugu is telic information.
 Bugu is agent information.
- b. Thalusamaipfi (dictionary) is a formal book.
 Thalusamaipfi (dictionary) is telic reference.
 Thalusamaipfi (dictionary) is agent compiled material.

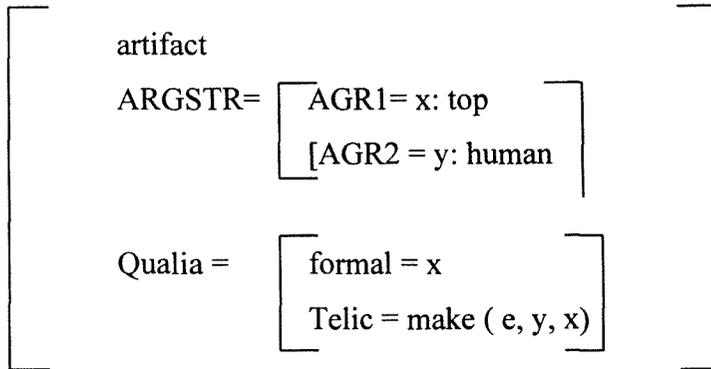
Such unified types as indicated above may be illustrated as follows within a structure of a lexical item such as zwiliwa (food)

- (183) a. inheritance relation
 Zwiliwa (food) is a physical object.
- b. Qualia structure
 telic (its use or function) : activity eating formal (that which distinguishes it within a larger domain) : phys. Obj .
 Agentive (its origin) : make



The physical object zwiliwa must be edible. Another example is concerned with artifacts i.e. objects which were created, made, manufactured or brought about by human activity and artifacts has a default argument (N- ARG) which is human.

(184)



In order to get its origin both the formal value of a physical object and the agentive value of an artifact are combined and the resulting lexical representation is then a unified type.

CHAPTER 4

4. LEMMATIZATION OF NOUNS IN A TSHIVENDA DICTIONARY

4.1 AIM

The aim of this chapter is to examine how nouns could be lemmatised in a Tshivenda dictionary.

This section will try to establish which information with regard to syntax, morphology and the semantics of the noun could be included in the lemma of the noun and how much of this information should be included in a dictionary.

It will further establish whether the various affixes in Tshivenda may occur freely with all nouns or whether they may have a limited distribution.

Lastly, it will establish as to which information with regard to derivations should be included in a dictionary.

4.2 NOMINAL COMPONENT

This section will cover the discussion on syntax, morphology and semantics of the noun.

4.2.1 Syntax of the Noun

It has already been discussed in section 3.1 of chapter 3 that when a noun appears in a sentence it has to satisfy the following:

- (185)
- a. It is the head of the NP,
 - b. It has certain theta-roles,
 - c. It occupies certain functional positions in a sentence,
 - d. It is assigned a case, and

- e. It may be moved in a sentence.

The question with regard to the dictionary is whether any such information as above (185) need to be included within the lemma of a noun. Following on what I have said in (185) there are four kinds of grammatical information that we might expect to find in a dictionary entry. This information which is of a syntactic nature may be provided explicitly or implicitly.

The syntactic operation of the lexical item in relation to other words could be indicated by means of citations or invented examples. The most important task of the examples is to explicate the syntactic operation of the lemma by placing it into syntactical context.

For example, verbs are traditionally marked as “transitive” or “intransitive” but this is insufficient information to enable the dictionary user to build acceptable clauses. The dictionary entry needs to specify in detail which clause patterns a verb may enter, what complementation a verb may take, which items are obligatory or optional or deletable. Additionally, the dictionary needs to specify whether a verb cannot undergo a particular rule of the grammar, if it is unable to form “progressive” tenses or to enter the “passive” construction. Verbs determine to a large extent which other elements may be present in its clause.

What is true for the specification of the syntactic performance of verbs is also true to a lesser extent, to nouns. In dictionaries, verbs are marked by heads of their phrasal structures. In dictionaries, nouns as well, are marked with labels that are primarily of a syntactic nature.

- (186) a. muthu, n. person.
b. kholomo, n. cattle.

In both (186a) and (186b) the label [n], which is the head of the NP, designate the lemma as a noun. This label is of syntactic nature. In syntax the noun [N] is regarded as the

head of the noun phrase [NP]. Consequently, the lemmas [muthu] and [kholomo], is each marked with the label [n] to show their syntactic operations, i.e. that they are heads of the NPs. This is one area where the syntactic operation of the noun may be indicated. And it is the only area where this syntactic operation is indicated outside the context of a sentence.

Other syntactic issues are indicated within the context of a sentence. Within a dictionary this information may be provided implicitly or covertly by means of illustrative examples that may form part of a dictionary entry. These examples, often given in the form of sentences, show typical combinations and various uses of the noun in the context within which it is placed.

(187) Miri, n. trees.....

 (miri mihulu i thusa kholomo nga mirunzi:
 big trees help cattle with shade).

It is within this context where a noun may be represented in dictionary specifying what a noun does satisfy in a sentence, i.e. as in (187) above.

But the fact that a noun may function syntactically may not be stated explicitly but implicitly. In dictionaries this is made possible by illustrative examples invented by lexicographers or their co-workers, who made it possible for nouns to satisfy their syntactic performances.

4.2.2 Morphology of the noun

This section will examine two issues of morphology which are of importance, i.e. nominal inflection and derivation. It will further be established as to which material, in terms of the above issues of morphology, should be included in a dictionary entry of a noun.

4.2.2.1 Nominal inflection

In Tshivenda there are three inflectional categories for nouns, namely: case, noun class and agreement. For this section noun class will be considered since it represents an inherent feature of nouns. It has already been established in chapter 2 that the African languages has a two-fold tradition with regard to the noun class prefixes. Some dictionaries follow a stem or root entry while others include the prefix with the root or stem as a word entry. The tradition in Venda dictionaries has been to concentrate on the noun as word, i.e. nouns are entered within a dictionary according to the first letter of the noun class prefix of the noun.

- (188) a. Mu-lilo (3) 'fire'
 lu-funo (11) 'love'
 b. nwana 91) 'child'
 duvha (5) 'day'/'sun'

On the other hand nouns in the Nguni languages were entered alphabetically according to the letter of the root or stem. Such a procedure has specifically been followed because of the presence of a preprefix in these languages. Depending on the vowel in the noun class prefix, all nouns in the Nguni languages with preprefixes begin with either i, u or a.

- (189) a. iphupho 'dream'
 b. umsiki 'tailor'
 c. abantu 'people'

This will mean that all nouns may only be entered under one of these three letters. The other languages like Venda, Tsonga and Sotho do not have these preprefixes, so that it was not considered a problem for dictionaries. These languages thus usually enter nouns in the dictionary under the first letter of the prefix if present.

(190) a. Venda

Tshi-ungulela 'heart burn'

Vhu-dudo 'heat'

b. Tsonga

Xitlhoni 'hedgehog'

Xitiko 'fireplace'

c. Sotho

Mogwapa 'biltong'

Lerapo 'bone'

However, this practice was not consistently followed, see i.a. the dictionaries of Paroz (1988) and Snyman (1990).

(190) a. Thitho, 1. N. bo-, warmth

tho, 1.n. bo-, personality (SEA, 1990).

b. mo.horo(me) n., burrow, lair

le.supi (ma) n., ruin (SSE, 1988)

This question can only be answered in a principled way with regard to the syntactic and morphological category of noun. Within the X-bar theory as explicated in chapter 2, the noun will be a category with no bars which represents the head of a noun phrase (NP). This syntactic category is shared by the morphology because the morphological and syntactic category noun is exactly the same. It will thus be necessary to look for the answer within these two components which deal with nouns. The answer cannot be apparent from the tradition of dictionaries in the African languages because then there is no consistency apparent as has been shown above.

The three major lexical items in Tshivenda always appear with prefixes and sometimes suffixes. The formal category adjective in Tshivenda has to appear within a syntax with at least a noun class prefix.

- (192) a. mu-denya
b. tshi-vhi
c. zwi-lapfu

The adjective may also appear with a suffix such as – ana.

- (193) a. muhulwane
b. tsethana
c. tswana

However, all the dictionaries in the African languages always enter the adjectives only with their roots and no prefixes and suffixes.

- (194) a. Venda
-hulu ‘big’
-tuku ‘small’
b. Sotho
-kgolo, ‘big’
-sweu ‘white’
c. Nguni
-fuphi ‘short’
-khulu ‘big’
d. Tsonga
-tsongo ‘little’
-kulu ‘big’

There are clear reasons for this, because all adjectives may appear with all the noun class prefixes and such dictionary entries will thus add information which is apparent and which is not needed.

The category verb also appears with a number of prefixes and suffixes, both inflectional and derivational. But the verb is never entered in a dictionary with its inflectional morphemes such as agreement, tense, mood, negative and aspect.

(195) a. agreement:

1. subjectival agreement

vha- a- shuma

2. objectival agreement

vha- a- mu- toda

b. tense:

1. present tense

[vha- a- lima]

2. perfect tense

[vho- lima]

3. future tense

[vha- do- shuma]

c. mood:

1. Indicative (factivity)

Vhana [vha- a- lima]

2. Subjunctive (necessity)

Ndi toda uri [vha- lime]

3. Potential (possibility)

[Vha- nga- lima]

d. negative:

a/nga/sa/si/ngo/songo

lima/lime

e. aspect:

vha	Progressive: kha di	lima
	vho	
	a di	
mbo di		
vho	no	

The verb is always entered as a root or stem with at least a suffix [-a] which does not express any grammatical function.

- (196) a. -tshe-a 'cut'
 b. -gw-a 'dig'
 c. -lim-a 'plough'

Most dictionaries however do include derivation affixes such as the applicative and the causative.

- (197) a. -fela (applicative of -fa) die for, suffer for. (VE, 1989)
 b. -kubugela (appl. < kudugu) charge at, rush at. (SEA, 1990)

or/and the causative such as the following:

- (198) a. -limisa, help or compel to hoe,
 help with hoeing, ploughing (VE, 1989)
 b. -rekisa, (rekisa) caus, sell, sell out,
 cause or help to buy, be tray (SEA, 1990)

In the same way nouns appear with the inflectional morphemes which are known as noun class prefixes and noun class is considered as a morphological category of the noun in Tshivenda.

With syntax and morphology the category noun is represented by the root or stem of the noun with the prefixes added as inherent semantic and grammatical features of such nouns. A lexical entry for a noun in Tshivenda may thus look as follows:

(199)	<div style="display: flex; flex-direction: column; justify-content: space-around;"> -thu -V, +N +1, 2 </div>	<div style="display: flex; flex-direction: column; justify-content: space-around;"> phonological form category noun class semantics </div>
-------	--	--

Such a lexical entry within a lexicon may then be translated in a variety of ways within a dictionary of the language. The following possibilities may be considered:

(200) a. The root with features such as above.

b. The root with the prefix of class 1 sorted alphabetically under the first letter of the root:

e.g. mu-thu

The plural may then be added: muthu (vha)

The noun class feature, e.g. 1/2 may in certain instances be necessary to consider the derived category agreement from a noun class prefix which is not visible

c. The noun may be entered under the first letter of the prefix: mu-thu.

The exact entry in a dictionary has to be considered from the view point of the syntactic and morphological category only. Furthermore it has to be consistent with all the dictionaries in the specific language to guarantee user- friendliness of the dictionary.

Taking these considerations into account, a case can be made for the treatment of nominal inflection as in (200b) above, i.e. entry in the dictionary with the root or stem as base together with its category, the two inflectional prefixes and its semantics. This will conclude the role of nominal inflection in lemmas.

4.2.2.2 Nominal lexical derivation

With noun class as an inflectional morpheme it is apparent that all nouns have to appear with some reference to a class prefix because noun class will always determine agreement on the verb and the nominal modifier.

However, with lexical derivation it will be necessary to establish whether the various affixes may occur freely with all nouns or whether they have a very limited distribution. The answer to this issue will determine whether these derivations need to be entered in a dictionary.

Here, two lexical derivations will need attention; namely: nominal lexical derivational affixes and nominal functional derivations.

4.2.2.2.1 Nominal lexical derivational affixes

Within this category, affixes will be taken into consideration.

(201) a. gender: -lume -kadzi
 Malume ‘uncle’
 Mufumakadzi ‘woman’

b. Locative suffix: - ni

c. Diminutive suffix: -ana

 Mbudzana ‘a baby goat’
 Thorwana ‘small grain’

It seems as if only (201a) above need to be considered because only specific nouns may appear with these gender affixes. The locative and diminutive affixes may appear almost with all the nouns (except for class 1a/2a and the locative). Each noun which may appear with the gender affixes need to be specified in such an entry.

(202)	ndou N 9/10 -lume, kadzi] phonological form category noun class gender
-------	-----------------------------------	---

A dictionary will dictate how such a lemma for the above lexical entry may look.

(202) Ndou, n. 9/10 elephant, [+lume/kadzi]

Alternatively more than one lemma may be necessary but it will be imperative to include all the above information in a dictionary.

4.2.2.2.2 Nominal Functional Derivations

Here, two such derivations need to be considered, namely: derivation from nouns and derivations from verbs.

4.2.2.2.2.1 Derivations from Nouns

In Tshivenda certain nouns are derived from nouns by means of class prefixes.

Nouns derived with such noun class prefixes could express:

- (203)
- a. Augmentative/derogatory meaning,
 - b. Diminutive/pejorative meaning,
 - c. Meaning dealing with locality, and
 - d. Meaning dealing with language/culture

The meaning expressed by the nouns derived with certain noun class prefixes is apparent in the original noun. Augmentative and diminutive nouns or nouns with derogatory sense are derivations which can be realised with all nouns.

- (205) a. class 5: augmentative [li -du]
b. class 5: derogatory individual [li -sadzi]
c. class 7: diminutive [tshiduna]
d. class 7: pejorative [tshisagana]
[tshikedzi]

The above nouns (205 a + b + c + d) are possible with all nouns. Their inclusion in the dictionary depends on their usage frequency. If they have high usage frequency they could be included in a dictionary. This will help not to include all nouns found in this category.

But an exception could be made to derivations with variant spellings.

- (206) a. class 20 diminutive:
nwana > kwana
b. class 21 augmentative/derogatory:
nwana > dana

The above (206a+b) derived nouns have a variant spelling. This could help in including them in a dictionary. But the rest of derived nouns found in this category which can be realised with all nouns, should be included on merit.

Derived nouns which denote locality, language or culture are found only with a specific group of nouns.

- (207) a. Class 7: language, culture:
'Tshivenda'

‘Tshivhuru’

b. Class 11: language:

‘Luisimane’

‘Luvhuru’

c. Class 14: place:

‘vhulaudzi’

‘vhuphani’

This information (207a + b + c) is not possible with all nouns. Not all derived nouns can denote a language, a specific culture or a locality. This type of information is realised in limited nouns. The inclusion of these nouns is not restricted as that of augmentative, derogatory and diminutive nouns.

4.2.2.2.2 Derivations from Verbs

In the African languages, the noun class prefix has a certain contribution to the derivation. Thus both the noun class prefix and the suffixes [-I] and [-o] have to be taken into account when dealing with deverbatives. Such derivations may then be classified into major categories with the feature [\pm human] where [+ human] is spelled out as]-I] and [-human] as [-o].

There are nouns derived from verbs with the suffix [-I].

(208) a. Class 1: person names

Muimbi ‘singer’

Mutambi ‘player’

b. Class 7: Exceptional

tshibiki ‘good cook’

tshivhadi ‘good sculptor’

The nouns derived from verbs with the suffix [-i] are possible with all verbs.

Their inclusion in a dictionary is very difficult. It may as well depends on the high usage frequency of the noun.

In Tshivenda all verbs which indicate an action done by a human being may form nouns.

(209)	Class 1:	person names
	mu-rung-i	'sewer'
	mu- bik-i	'cook'

These (209) derivations may appear with all nouns.

(210)	Class 7:	exceptional
	tshi-limi	'plougher'
	tshi-shumi	'worker'

These (210) derivations may as well appear with all nouns. Therefore it will be very difficult to include all of them in a dictionary. Treating them as lemmas will depend on how they are frequently used.

Other nouns are derived from verbs with the suffix [-o].

(211)	a. noun class 3:	event 1 manner
	Murungo	'sewing'
	Muvhuso	'government'

These derivations are not possible with all verbs. A verb such as [vhasa] [to light wood] is not possible in this group.

- b. noun class 7: theme/concrete
tshivhaso 'fire place'
tshirendo 'poem'

Not with every verb can we get derivations belonging to this group. We only have a limited number of verbs from which we can get derivations indicated above.

- c. noun class 9: Result/Resultative (abstract)
pfunzo 'education'
phindulo 'answer'

The above derivations are only possible with a specific group of verbs. Not all verbs can form these derivations.

- d. noun class 11: abstract
lufuno 'love'
lutendo 'faith'

There are a group of verbs which can form these derivations. Other verbs form other derivations. There are verbs which only form these derivations.

- e. noun class 11: event
lulindo 'guarding'
lupenyu 'lightning'

These derivations can be formed by a great number of verbs. So, their inclusion in a dictionary must be principled.

- f. noun class 20: manner
kushumele 'way of working'
kutshimbilele 'way of walking'

These derivations are possible with all verbs. Most of the verbs in Tshivenda can form the nouns belonging to this group. But not all of these nouns could be included in a dictionary.

Most of the derived nouns will not be included in a dictionary of Tshivenda.

Derived nouns such as those formed by noun class 20 [ku] and the verb are too many to include in a single Venda dictionary.

These type of derivations cannot be included in a single dictionary. Moreover these derivations provide meaning which is apparent. Their meaning has already been explained in the articles of the verbs.

There are derivations from class 15. These derivations may appear with all verbs.

- (212) a. u- tshimbila ‘to walk, walking’
 b. u- gidima ‘to run, running’
 c. u- shuma ‘to work, working’

In Tshivenda these are infinitives which are derived from class 15 [u]. These infinitives appear with all the verbs in Tshivenda. Derived infinitives are not included in a Tshivenda dictionary. However in languages like Xhosa, there are those infinitives which are derived and those which are no longer derived.

- (213) a. ukutya ‘food’, ‘the process of eating’

These infinitives which are no longer derived are entered in a dictionary. Unfortunately these infinitives which are no longer derived do not exist in Tshivenda.

4.2.3 SEMANTICS OF THE NOUN

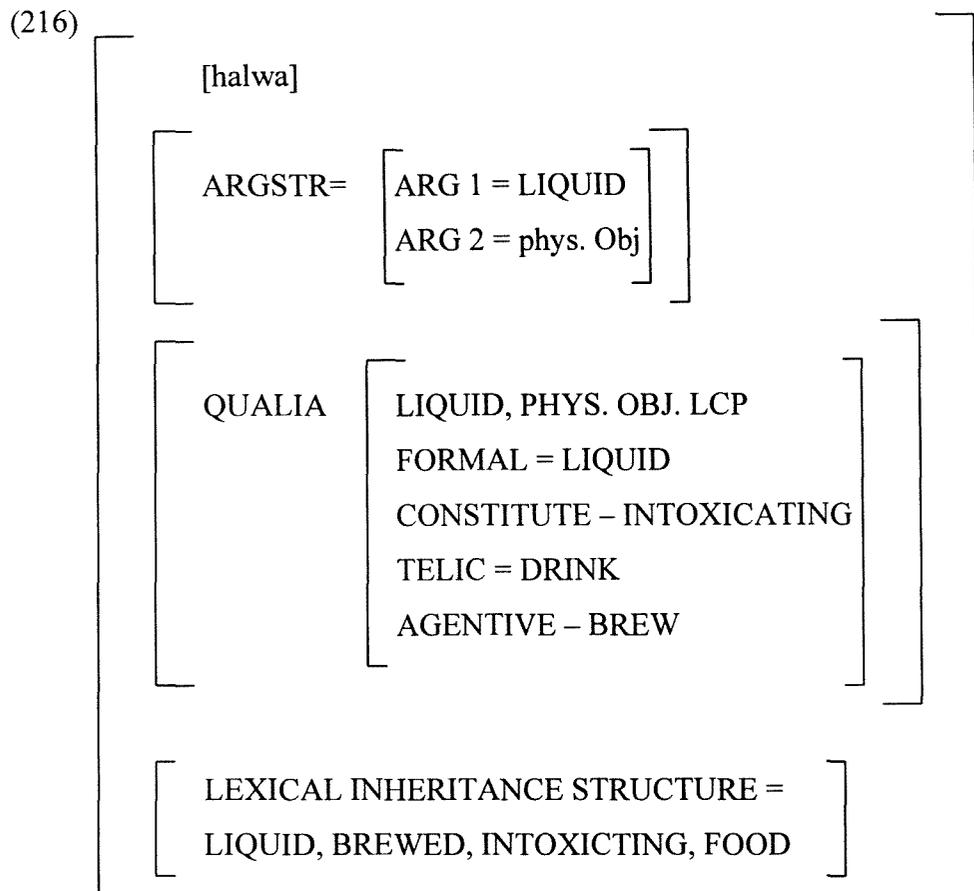
It has already been discussed in chapter 3 that in a generative lexicon, there are 4 levels of representations, namely:

- (214) a. argument structure,
b. event structure,
c. qualia structure, and
d. lexical inheritance structure.

The semantics of the noun is studied through four types of arguments, namely:

- (215) a. true arguments,
b. default arguments,
c. shadow arguments, and
d. true adjuncts.

The argument structures of noun depend on the number of different senses that a specific noun may have. Nouns may appear with two and/or more than two arguments, i.e. they may have one, two and/or more than two senses. Semantically, the noun [halwa] could have the following representation:



Now the question with regard to a dictionary is how much of this information will be necessary. This will depend on two things; namely:

- (217) a. the dictionary's target group, and
- b. the dictionary typology.

In a comprehensive monolingual dictionary of Tshivenda, the information about the semantics of the noun [halwa] could be represented as follows:

- (218) Halwa, n. lushaka lwa tshinwiwa (telic)
 Tsha tshiludi (formal) tshi no
 Kamba (constitute) tso itwaho
 (agentive) nga luvhele.

But the English representation inserted above will not be visible in a Venda monolingual dictionary. All the representations will be in Tshivenda. The type of definitions such as the one above are only included in comprehensive monolingual dictionaries. It is because these dictionaries have a capacity of extra-linguistic materials, such as those represented above at (218).

Other dictionaries such as a standard dictionary of Tshivenda will have a limited representation of the above information at (218) above. These dictionaries will have both Venda and English representations of the noun [halwa].

(219) halwa, n. beer, made from maize (VE, 1989)

The standard dictionary of Tshivenda will only indicate one essential aspect of the meaning of the noun halwa. The noun [halwa] is defined as “made from maize”. The other essential aspects of the noun [halwa] are not indicated. This is what limits standard dictionaries as compared to comprehensive monolingual dictionaries.

Standard dictionaries are good for a certain group of people, especially those that are learning the language as their second language. But for first language speakers we need a full representation of the meaning of the noun [halwa]. This information could be indicated in a comprehensive monolingual dictionary of Tshivenda.

The noun [halwa] has two arguments. In a Tshivenda dictionary this could be indicated by two senses. These two senses of the noun [halwa] could be represented either polysemously or monosemously.

(220) Halwa n. 1. Tshiludi tso no kamba

Bodelo lo no fara halwa

As a polysemous item the noun halwa could have the above meaning. But the two senses, with development of time could be separated.

- (221) Halwa¹ n. tshiludi tsho no kamba.
Halwa² n. bodelo lo no fara halwa

The above (221) is an indication of the noun [halwa]. This is due to its two arguments. The noun [halwa] could be taken as ‘an intoxicating drink’; others could take it as ‘a bottle for storing beer’. Here the noun [halwa] could be placed under two generic concepts.

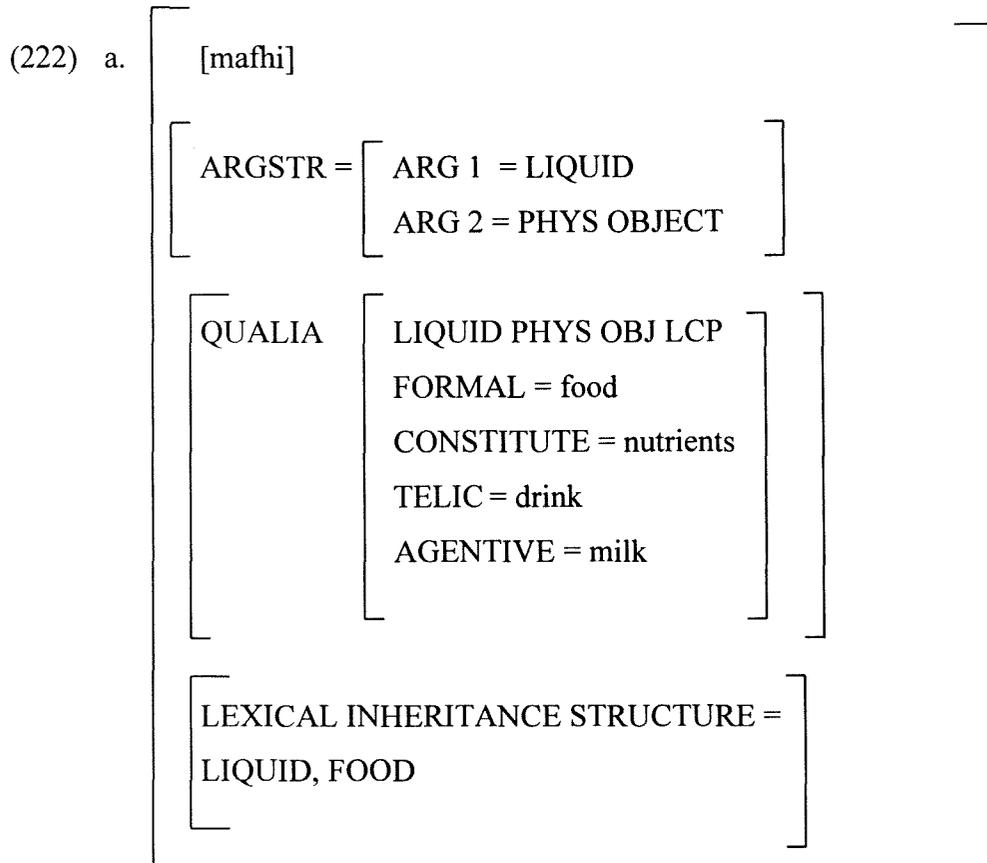
With other co-hyponyms the noun [halwa] could be placed under the genus [liquids]. Here the noun [halwa] is treated as belonging to the same genus as other [liquids] such as water, milk and juice. The meaning of [halwa] could as well be explained in the context of [SOLIDS or PHYSICAL OBJECTS] such as stones, wood, bricks and porridge.

In order to differentiate the noun [halwa] from the other co-hyponyms, extra – linguistic or supplementary information is required. This information could only be indicated in a monolingual Tshivenda dictionary.

Componential analysis also plays an important role in explaining the different types of [halwa]. The noun [halwa] could have intoxicating or non-intoxicating effect. This information as well could be indicated by providing encyclopaedic information in a dictionary.

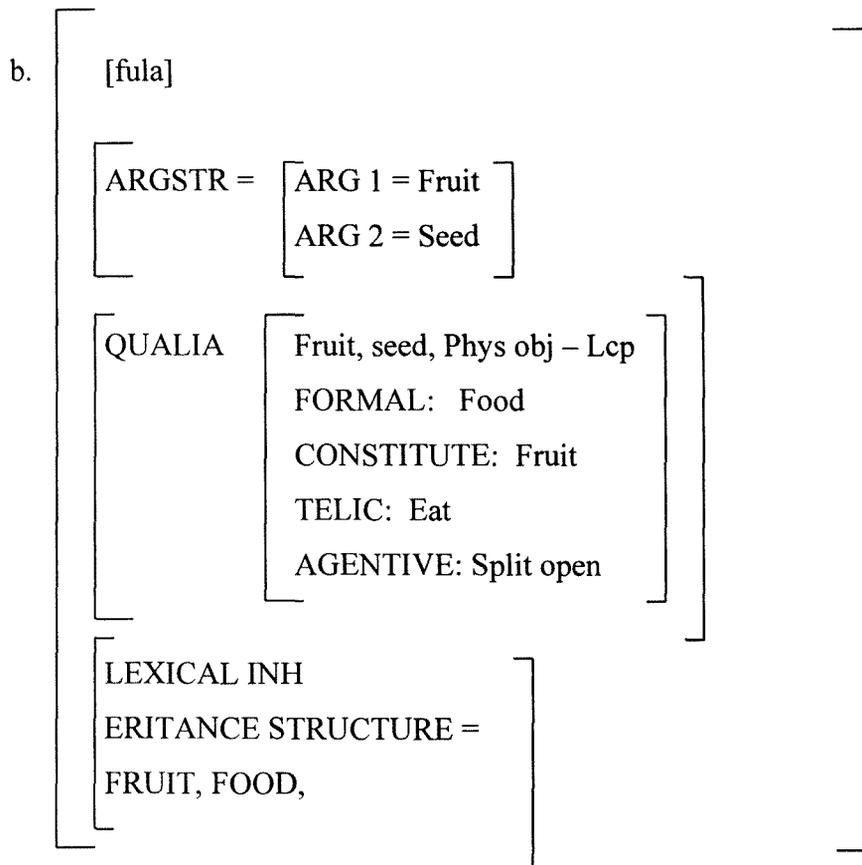
4.3 CONCLUSION

Lemmas in a Tshivenda monolingual dictionary could have the following representation:

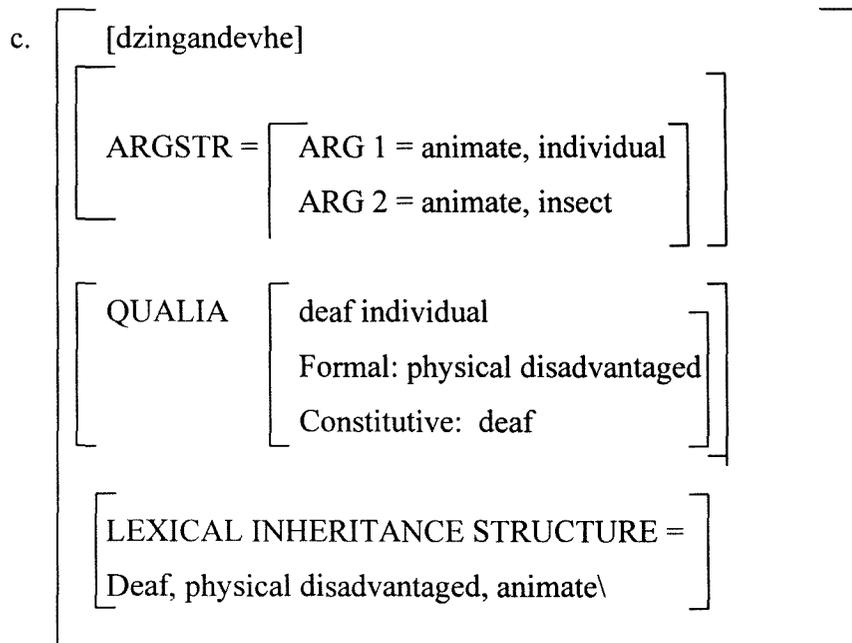


Mafhi n. (class 6, usually referred to as mafhi in both singular and plural forms)

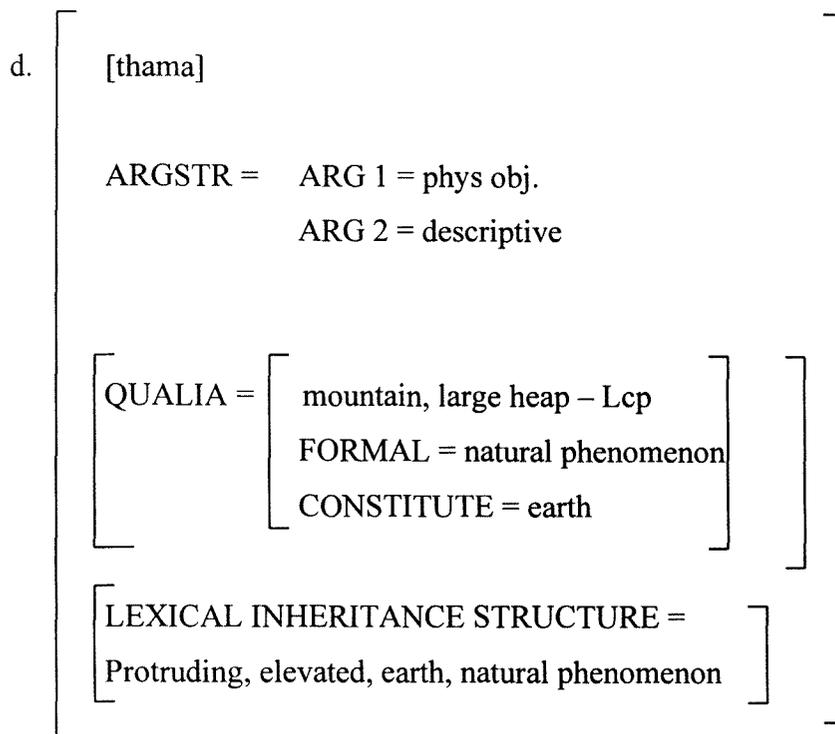
1. Zwiliwa zwi re kha tshiludi (liquid) kana luphuse (powder) zwine zwa fhatha (nourish) muvhili zwine zwa tou nwiwa. Mafhi a wanala nga u tou hamiwa kha kholomo kana bva kha mavhele.
2. Mudzio kana mufareledzo wa tshiomate, sa wa vumba, tsimbi kana pulasitiki une khawo nga ngomu ha shelwa mafhi.



1. Fula, n. (class 5 derived from mufula in class 3). 1. Mutshelo wa muri wa mufula une wa liwa wo vhibva hune khawo ra fhonda mukumbi.
2. Thambo ine ya devhiwa ha bva migogo ine ya liwa. Thambo i a tavhiwa ha mela miri ya mifhula.



1. Dzingandevhe, n. (class 5, with prefix ma- in madzingandevhe) 1. Muthu ane a sa kone u pfa tshithu nga ndevhe dzawe.
2. Tshikhokhonono tshituku tsho no fhufha tshine tsha fana na nemeneme; tshine tshi sa liwi.



1. Thavha n. (class 9, with prefix dzi- in dzithavha, but usually referred to as thavha in plural, instead of dzithavha) 1. Fhethu kha shango hune ha vha nthanga matungo ha vha ho tuluwaho.
2. Thulwi ya zwithu zwinzhi zwo namelanaho nga ndila i songo dzudzanyeaho, kana nga ndila yo kalulaho.

In a monolingual dictionary extra-linguistic information helps the user to understand the lemma better. The comprehensive monolingual dictionary has an informative function. Its aim is to help in the transfer of meaning rather than being normative, it often employ the system of labelling to show deviation.

CHAPTER 5

In chapter 1 the main aim of the thesis was given. The problem statement was mentioned as the investigative analysis of the entry and treatment of nouns in a dictionary. The whole study was limited to the analysis of the entry and treatment of nouns with special reference to a Tshivenda dictionary.

The research centered mainly on the work of various lexicographers and linguists as it can be seen through out the whole study.

In order to achieve the aim of this study, the science which is concerned with the making of the dictionary and the finished product has been explored. This has been referred to as dictionary research or metalexigraphy and has the following fields as indicated by Hausmann (1986):

- (i) Lexicographic theory
- (ii) Criticism of dictionaries
- (iii) Research on dictionary use
- (iv) Research on dictionary status and marketing
- (v) History of lexicography

Lexicographic theory has been chosen as the one which is relevant to this study. Its areas of interest has been explicated as indicated below:

- a. The theory of Lexicographic description of language
- b. The theory of Lexicographic language research
- c. The theory on the relationship between lexicography and other disciplines
- d. The theory of organisation of lexicographic activities, and
- e. The theory on the purposes of dictionaries.

Chapter 1 also mentioned the various types of lemmata. Nouns were investigated in chapter 2 as lexical lemmata.

Chapter 2 explored the entry and treatment of nouns in Non-African and African language dictionaries, with specific reference to Tshivenda dictionaries. The main focus was on the two main lexical traditions which had influenced the entry and treatment of nouns in dictionaries.

It was found that dictionaries belonging to the Nguni languages and very few of the Sotho languages have treated nouns according to the stem tradition. The word tradition was used by dictionaries belonging to Tshivenda, Northern Sotho and Tsonga languages respectively.

The handling of exceptions was also explored.

Chapter 3 has explored nouns in syntax, morphology and semantics.

The main focus was on nouns in the X-bar theory. The discussion looked at the existence of the noun as the head of the Noun Phrase (NP). Different phrase structures were looked at. The discussion has looked also at theta role and Predicate Argument Structure where NPs are assigned theta role. The NP in a sentence has three relationships:

- a. Obligatory relation
- b. Unique relation, and
- c. Locally structured relation

The discussion also looked at the functional positions of the subject, the object and the complement of a preposition. The discussion has covered case theory where relations between nouns are expressed. The movement of the passives and subject inversion leave the subject position empty and the agreement and pro are then coindexed with the feature [existential].

In the morphology of the noun, the discussion looked at nominal inflection and the discussion looked at four things, namely:

- a. Case,
- b. Noun class,
- c. Agreement, and
- d. Number

The discussion involving lexical derivation looked at four things, namely:

- a. Feature value switches,
- b. Expressive derivations,
- c. Nominal affixes in Tshivenda, and
- d. Nominal function derivation

In feature value switches, the discussion looked at marking of nouns, by gender and locative affixes. The discussion also looked at how nouns are expressed in Tshivenda.

Expressive derivation reflects five functions, namely:

- a. Diminutive,
- b. Augmentative,
- c. Pejorative,
- d. Affectionate, and
- e. Honorific

The discussion looked at nominal affixes such as gender affixes: -kadzi'-ana/-nyana; diminutive affixes: -ana/-nyana and locative affix -ni. In nominal function derivation the discussion looked at the functions of suffixes [-I] and [-o].

The discussion involving the semantics of the noun looked at four level of representations in a generative lexicon, namely:

- a. Argument structure

- b. Qualia structure
- c. Lexical inheritance structure, and
- d. Interaction of semantic levels

The discussion involving argument structure has looked at types of arguments which a noun could have. A noun could have the following arguments:

- (i) True arguments,
- (ii) Default arguments,
- (iii) Shadow arguments, and
- (iv) True arguments.

Qualia structure specified four essential aspects of a word meaning such as:

- (i) Constitute,
- (ii) Formal,
- (iii) Telic, and
- (iv) Agentive.

Nouns were discussed as having a small number of generic concepts. The chapter concluded by discussing the interaction of four levels of representation into one system of lexical semantic representations and the semantics of nominals.

Chapter 4 investigated the lemmatization of nouns in a Tshivenda dictionary. The study looked at the practical representation of nouns in a dictionary. The discussion looked at the nouns under syntax, morphology and semantics. Under syntax the discussion looked at the information which could be included in a lemma of a noun. It was established that citations and examples, as well as labels of syntactic nature could be included within the lemma of the noun.

In morphology the discussion looked at noun class which is regarded as an inherent feature of nouns. The discussion looked at two traditions. Stem tradition which is used by the Nguni languages and the word tradition which is used by Tshivenda, Sesotho and Tsonga languages respectively, in lemmatizing nouns.

Lastly the discussion looked at the semantic representation of nouns in a monolingual dictionary. The information discussed at chapter 3 about the semantics of nominal were applied here. Each lexical entry of a noun was discussed with all its relevant information.

Chapter 5 provided a brief summary of what has been discussed in the whole thesis.

6. BIBLIOGRAPHY

1. Bosman D.B. et al. 1979. **Tweetalige Woordeboek**. Cape Town: Tafelberg.
2. Butterfield, J. et al. 1997. **Collins Spanish Dictionary plus Grammar**. Great Britain: Harper Collins Publisher.
3. Chandor, A. 1985. **The Penguin Dictionary of Computer**. England: Penguin Books.
4. Cuenod, R. 1967. **Tsonga-English Dictionary**. Johannesburg: Swiss Mission.
5. De Villiers, M. 1988. **Nasionale Woordeboek**. Pretoria: Nasou Beperk.
6. Doke, C.M. et al. 1990. **English-Zulu/Zulu-English Dictionary**. Johannesburg: Witwatersrand University Press.
7. Du Plessis, J.A. 1996. **Venda Syntax**. Stellenbosch. University of Stellenbosch SCAL.
8. Du Plessis, J.A. 1996. **The Categories and Phrase Structures of the African Languages**. Stellenbosch. University of Stellenbosch. SCAL.
9. Du Plessis, J.A. 1997. **Morphology of the African Languages**. Stellenbosch. University of Stellenbosch.
10. Du Plessis, J.A. 1999. **Lexical Semantics and The African Languages**. Stellenbosch. University of Stellenbosch SCAL.
11. Eksteen, L.C. 1997. **Pharos' Groot Woordeboek/Major Dictionary**. Cape Town: Pharos.
12. Encyclopedia Americana, Volume 20. 1992. United States of America: Grolier Incorporated.
13. Feinauer, A.E. 1997. **The Treatment of Semantic Information in Dictionaries**. Stellenbosch: University of Stellenbosch Press.
14. Geeraerts, D. 1984. **Dictionary Classification and the Foundation of Lexicography**. I.T.L. Review 63: 37-63.
15. Geeraerts, D. 1991. **Principles of Monolingual Lexicography in Hausmann F.J. et al.** 1989 – 1991. 287 – 296.

16. Gouws, R.H. 1991. **Review.** Hausmann, F.J. et al. Wörterbücher Dictionaries. Dictionaries. An International Encyclopedia of Lexicography. Berlin: Walter de Gruyter.
17. Gouws, R.H. et al. 1991a. **Toward a Lexicon-based Lexicography.** Dictionaries 13: 75 – 90.
18. Gouws, R.H. et al. 1994. **Basiswoordeboek Van Afrikaans.** Pretoria: J.L. Van Schaik.
19. Gouws, R.H. 1997. **The General Theory of Lexicography.** Stellenbosch: University of Stellenbosch Study Guide.
20. Hartmann, R.R.K. 1983. **Lexicography Principles and Practice.** London: Academica Press.
21. Hausmann, F.J. 1986. **The Training and Professional Development of Lexicographers in Germany in Lexicography.** AN Emerging Internal Profession Ilson. (ed) London: Manchester University Press.
22. Kriel, T.J. et al. 1989. **Pukuntsu Woordeboek N. Sotho-Afrikaans – N.Sotho.** Pretoria J.L. Schaik.
23. Kriel, T.J. et al. 1997. **Pharos' Popular Northern Sotho Dictionary.** Cape Town: Pharos.
24. Macmillan, J.H. and S. Schumacher. 1993. **Research in Education: A Conceptual Introduction.** United State of America: Harper Collins College Publisher.
25. Musehane, N.M. 1995. **The Nominal Compound in Tshivenda.** Stellenbosch. University of Stellenbosch Press.
26. Odental, F.F. et al. 1994. **Verklarende Handwoordeboek Van die Afrikaanse Taal.** Johannesburg: Perskor.
27. Odental, F.F. et al 1998. **Verklarende Handwoordeboek Van die Afrikaanse Taal.** Johannesburg: Perskor.
28. O'Grady, W. et al. 1996. **Contemporary Linguistics: An Introduction.** New York st. Martin's Press.
29. Pahl, H.E. et al. 1989. **The Greater Dictionary of Xhosa, Q to Z, Volume 3.** FortHare: University of Fort Hare Press.

30. Paroz, R.A. 1988. **A Southern Sotho**. English Dictionary. Monja: Sesotho Book Depot.
31. Procter, P. 1978. **Longman Dictionary of Contemporary English**. Great Britain: Longman.
32. Singh, R.A. 1982. **An Introduction to Lexicography**. Manasanganti: Institute of Indiana Languages.
33. Snyman, J.W. et al. 1990. **The Dictionary of Setswana-English-Afrikaans**. Pretoria: J.L. Van Schalk Limited.
34. Van Wyk, E.R. 1995. **Linguistics Assumptions and Lexicographical Traditions in the African Language Lexikos 5**: 82-96.
35. Wentzel, P.J. et al. 1982. **Improved Trilingual Dictionary of Venda – Afrikaans-English**. Pretoria: University of South Africa Press.
36. Wiegand, H.E. 1983. **On the Structure and Contents of a General Theory of Lexicographic Series 1**. Lexeter'83 Proceedings. Papers from the International Conference on Lexicography. 9-12 Sept. 1983. R.R.K. Harmann. (ed). Exeter: Max Nlemeyer Verlang.
37. The World Book Encyclopedia, N to Z, volume 14. 1987. London: World Book Inc.
38. The World Book Encyclopedia, N – to Z volume 14.1991. London: World Book Inc.
39. Zgusta, L. 1971. **Manual of Lexicography**. The Hague. Mouton.
40. Ziervogel, D. 1965. **Die Probleme Van Leksikografie in die Suid-Afrikaanse Bantoetale**. Taalfasette 1 : 456-53.