



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

# A morphosyntactic description of Northern Sotho as a basis for an automated translation from Northern Sotho into English

Gertrud Faaß

A thesis submitted in accordance with the requirements  
for the degree of Ph.D. in African Languages at  
the University of Pretoria.  
June 2010 (final version).

**Supervisor:** Prof. D.J. Prinsloo  
**Co-Supervisor:** Prof. U. Heid

## Summary

This PhD thesis provides a morpho-syntactic description of Northern Sotho from a computational perspective. While a number of publications describe morphological and syntactical aspects of this language, many of them are in the form of prescriptive study books (inter alia Lombard (1985); Van Wyk et al. (1992); Poulos and Louwrens (1994)) or of descriptive articles in linguistic journals or conference proceedings (inter alia Anderson and Kotzé (2006); Kosch (2006); De Schryver and Taljard (2006)), so far no comprehensive description is available that would provide a basis for developing a rule-based parser to analyse Northern Sotho on sentence level. This study attempts to fill the gap by describing a substantial grammar fragment. Therefore, Northern Sotho morpho-syntactic phenomena are explored which results in the following descriptions:

- language units of Northern Sotho are identified, i.e. the tokens and words that form the language. These are sorted into word class categories (parts of speech), using the descriptions of Taljard et al. (2008) as a basis;
- the formal relationships between these units, wherever possible on the level of parts of speech, are described in the form of productive morpho-syntactic phrase grammar rules. These rules are defined within the framework of generative grammar.

Additionally, an attempt is made to find generalisations on the contextual distribution of the many items contained in verbs which are polysemous in terms of their parts of speech. The grammar rules described in the preceding chapter are now explored in order to find patterns in the co-occurrence of parts of speech leading towards a future, more general linguistic modelling of Northern Sotho verbs. It is also shown how a parser could work his way step-by-step doing an analysis of a complete sentence making use of a lexicon and the rules developed here.

We have also implemented some relevant phrase grammar rules as a constraint-based grammar fragment, in line with the theory of *Lexical-Functional Grammar* (Kaplan and Bresnan, 1982). Here, we utilized the Xerox Linguistic Environment (XLE) with the friendly permission of the Xerox Palo Alto Research Centre (PARC).

Lastly, the study contains some basic definitions for a proposed machine translation (MT) into English attempting to support the development of MT-rules. An introduction to MT and a first contrastive description of phenomena of both languages is provided.

## Key terms

Northern Sotho, English, language units, word classes, electronic grammars, morpho-syntactic analysis, Northern Sotho Verbal Phrase features, word class distribution, automated translation, machine translation, grammar implementation.



## Declaration

I declare that **A morphosyntactic description of Northern Sotho as a basis for an automated translation from Northern Sotho into English** is to the best of my knowledge and belief, my original work. All the sources that I have used or quoted have been indicated and acknowledged by means of complete references. The material has not been submitted, either in whole or part, for a degree at this or any other university.

---

Dipl. Ling. Gertrud Faaß

## Acknowledgements

I should like to express my gratitude towards my promoters, Prof. Daan J. Prinsloo and Prof. Ulrich Heid. I freely admit that without both of their support, this work would have never been finished.

Many thanks to the other members of the department of African Languages at the University of Pretoria, especially Prof. Elsabé Taljard.

I am indebted to the DAAD, the German Academic Student Exchange Service, as their financial contribution made the project possible.

Thanks to the National Research Foundation of South Africa, who supported the funding of my working environment.

My thanks also go to friends and family, who all supported me throughout the time it took to complete this book.

There were a number of other people, for whom the space here is not enough to list them all. They helped me with translations and contributed to my general understanding of the languages and cultures of South Africa. *Ke a leboga, baie dankie* and thank you to all of you, as you offered me a grip of what is called *Ubuntu*.

# List of Abbreviations

1st	First person (inflectional element)
2nd	Second person (inflectional element)
3rd	Third person (inflectional element)
ADVP	Adverb Phrase
AP	adjective phrase
c-structure	constituent structure
CALL	Computer Aided Language Learning
CFG	Context-Free (Phrase Structure) Grammar
CONJ	conjunction
DM	Distributed Morphology
f-structure	functional structure
HLT	Human Language Technology
LFG	Lexical Functional Grammar
MT	Machine Translation
NP	Nominal Phrase
NPs	Nominal Phrases
OBJ-TH	thematic object
PARC	Palo Alto Research Centre
past	Subscript: past tense (inflectional element)
PF	Phonological Form
pl	Subscript: plural (inflectional element)
POS	Part of Speech
PP	Particle Phrase
PSC	University of Pretoria Sepedi Corpus
refl	reflexive
sg	Subscript: singular (inflectional element)



SMT	Statistical Machine Translation
TNS-ASP	TENSE-ASPECT
TTS	Text to Speech
VBP	Basic Verbal Phrase
Vend	The suffix(es) a verb stem ends in
VIE	Verbal Inflectional Element
VIEimp	imperative verbal inflectional phrase
VP	Verbal Phrase
VPimp	imperative VP
VPpred	predicative VP
VPs	Verbal Phrases
XLE	Xerox Linguistic Environment

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Language introduction . . . . .	1
1.2	Aims . . . . .	3
1.3	Methods . . . . .	6
1.4	A general introduction to grammar . . . . .	7
1.4.1	Introduction . . . . .	7
1.4.2	What is a grammar? . . . . .	8
1.4.2.1	Word versus token . . . . .	9
1.4.2.2	Semantics versus syntax . . . . .	10
1.4.3	The computational perspective . . . . .	12
1.4.4	A brief introduction to electronic grammars . . . . .	13
	Context-Free Grammar (CFG) . . . . .	13
1.5	Layout of the study . . . . .	18
<b>2</b>	<b>The word classes of Northern Sotho</b>	<b>20</b>
2.1	Introduction . . . . .	20
2.2	The noun ( $N_{\text{categ}}$ ) . . . . .	22
2.2.1	The noun class system . . . . .	22
2.2.2	Noun class prefixes - an overview . . . . .	24
2.2.2.1	Noun classes 1 to 4 . . . . .	24
2.2.2.2	Noun classes 5, 14, and 6 . . . . .	26
2.2.2.3	Noun classes 7 and 8 . . . . .	27
2.2.2.4	Noun classes 9 and 10 . . . . .	27
2.2.2.5	Noun class 15 - The infinitive . . . . .	28
2.2.2.6	The locative classes 16 – 18, <i>N</i> - and <i>ga</i> -classes . . . . .	32



2.2.2.7	Notes on the (semi-)automated identification of noun classes	33
2.3	The pronoun	33
2.3.1	The emphatic (or absolute) pronoun (PROEMP <sub>categ</sub> )	35
2.3.2	The possessive pronoun (PROPOSS <sub>categ</sub> )	36
2.3.3	The quantitative pronoun (PROQUANT <sub>categ</sub> )	36
2.4	The concords	39
2.4.1	Introduction	39
2.4.2	The subject concord (CS <sub>categ</sub> )	39
2.4.3	The object concord (CO <sub>categ</sub> )	42
2.4.4	The possessive concord (CPOSS <sub>categ</sub> )	42
2.4.5	The demonstrative concord (CDEM <sub>categ</sub> )	43
2.4.6	The demonstrative copulative (CDEMCOP <sub>categ</sub> )	44
2.5	The adjective (ADJ <sub>categ</sub> )	45
2.6	The enumerative (ENUM)	47
2.7	The verb stem (V)	49
2.7.1	Introduction	49
2.7.2	Notes on some verbal suffix clusters	49
2.7.3	The auxiliary verb (V <sub>aux</sub> )	51
2.7.4	The copulative (VCOP)	52
2.8	Adverbs (ADV)	57
2.9	The morphemes (MORPH)	57
2.9.1	The imperfect or present tense morpheme <i>a</i> (MORPH <sub>pres</sub> )	57
2.9.2	The perfect or past tense morpheme <i>a</i> (MORPH <sub>past</sub> )	58
2.9.3	The future tense morphemes (MORPH <sub>fut</sub> )	58
2.9.4	The potential morpheme <i>ka</i> (MORPH <sub>pot</sub> )	58
2.9.5	The negation morphemes (MORPH <sub>neg</sub> )	59
2.9.6	The infinite morpheme <i>go</i> (MORPH <sub>cp15</sub> )	59
2.9.7	The deficient morphemes (MORPH <sub>def</sub> ) and the progressive morpheme <i>sa</i> (MORPH <sub>prog</sub> )	59
2.10	Particles (PART)	59
2.10.1	The agentive particle <i>ke</i> (PART <sub>agen</sub> )	60
2.10.2	The connective particles (PART <sub>con</sub> )	60
2.10.3	The copulative particle <i>ke</i> (PART <sub>cop</sub> )	61

2.10.4	The hortative particles (PART_hort)	62
2.10.5	The instrumental particle <i>ka</i> (PART_ins)	62
2.10.6	The locative particles (PART_loc)	62
2.10.7	The question particles (PART_que)	62
2.10.8	The temporal particle <i>ka</i> (PART_temp)	62
2.10.9	The question words (QUE <sub>categ</sub> )	63
2.10.10	Miscellaneous	65
2.11	Summary	65
<b>3</b>	<b>A fragment of the grammar of Northern Sotho</b>	<b>69</b>
3.1	Introduction	69
3.2	The Verbal Phrase (VP)	71
3.2.1	Introduction	71
3.2.1.1	Basic Verbal Phrase (VBP) versus Verbal Inflectional Element (VIE)	71
3.2.1.2	Terminology used in this chapter	73
3.2.1.3	Introduction to the modal system	74
3.2.1.4	The slot system	78
3.2.1.5	Labels used on nodes	79
3.2.1.6	Labelling information on the transitivity of verbs	80
3.2.1.7	Saturated verb forms	83
3.2.1.8	The object concord as part of the verb	85
3.2.2	The Basic Verbal Phase (VBP)	89
3.2.3	The imperative	89
3.2.4	The infinitive	92
3.2.5	The indicative	94
3.2.5.1	The imperfect tense	97
3.2.5.2	The perfect tense	100
3.2.5.3	The future tense/aspect	102
3.2.5.4	A summary of the indicative mood	104
3.2.6	The situative	105
3.2.6.1	The imperfect tense	109
3.2.6.2	The perfect tense	110
3.2.6.3	The future tense/aspect	112

3.2.7	The relative . . . . .	112
3.2.7.1	The imperfect tense . . . . .	116
3.2.7.2	The perfect tense . . . . .	117
3.2.7.3	The future tense . . . . .	119
3.2.8	A summary of the modifying moods . . . . .	119
3.2.9	The consecutive . . . . .	122
3.2.10	The subjunctive and the habitual . . . . .	123
3.2.11	A summary of the dependent moods . . . . .	125
3.3	The copulative verbal phrase (VP <sub>cop</sub> ) . . . . .	125
3.3.1	Introduction . . . . .	125
3.3.2	The identifying copulative . . . . .	129
3.3.2.1	The stative . . . . .	132
3.3.2.2	The dynamic . . . . .	135
3.3.3	The descriptive copulative . . . . .	142
3.3.3.1	The stative . . . . .	144
3.3.3.2	The dynamic . . . . .	147
3.3.4	The associative copulative . . . . .	154
3.3.4.1	The stative . . . . .	157
3.3.4.2	The dynamic . . . . .	160
3.3.5	A summary of the copulative constellations . . . . .	168
3.4	Auxiliary verbs . . . . .	169
3.5	Other verbal structures . . . . .	171
3.5.1	The hortative constellation . . . . .	171
3.5.2	Potential forms . . . . .	171
3.6	Adverbial phrases (ADVP) . . . . .	172
3.7	Summary of the verbal phrases . . . . .	174
3.8	Constellations of the Noun Phrase (NP) . . . . .	175
3.8.1	Introduction . . . . .	175
3.8.2	An overview of some nominal phrases . . . . .	177
3.8.3	The Pronominal Noun Phrase ( <sup>pro</sup> NP) . . . . .	180
3.8.4	Nominal phrases headed by demonstrative concords . . . . .	181
3.8.4.1	The relative Noun Phrase ( <sup>rel</sup> NP) . . . . .	181
3.8.4.2	The Adjectival Phrase (AP) . . . . .	182

3.8.5	The Possessive Noun Phrase ( <sup>pos</sup> NP)	183
3.9	The Particle Phrase (PP)	185
3.10	A sentence of Northern Sotho	186
3.10.1	The basic proposition	186
3.10.2	The question	188
3.11	A brief summary of our grammar fragment	188
<b>4</b>	<b>Features of verbal phrases</b>	<b>190</b>
4.1	Introduction	190
4.2	Parsers: approaches to describe natural languages	192
4.2.1	Vertical parsing directions	194
4.2.1.1	Horizontal parsing directions	194
4.2.1.2	Right-corner parsing	194
4.3	A sample analysis	195
4.4	Data categories of main verbs	205
4.5	Data categories of copulative verbs	213
4.5.1	A right-to-left perspective	213
4.5.2	A left-to-right perspective	219
4.6	Conclusions	225
<b>5</b>	<b>Implementation of a grammar fragment</b>	<b>226</b>
5.1	Introduction	226
5.1.1	Lexical-Functional Grammar (LFG)	227
5.1.2	The LFG formalism	227
5.1.2.1	Representations: constituent structure and functional structure	227
5.1.2.2	Predicate argument structure versus syntactic structure	227
5.1.2.3	An example analysis	229
Lexicon entries.		229
The rules section.		231
Constructing c- and f-structure		231
5.1.2.4	The Pargram project	238
5.2	Implementation	238

5.2.1	The lexicon . . . . .	238
5.2.2	The basic verbal phrase (VBP) . . . . .	246
5.2.3	The imperative . . . . .	249
5.2.4	The predicative independent indicating mood:	
	the indicative . . . . .	256
5.2.4.1	General rules for the indicative . . . . .	256
5.2.4.2	The imperfect indicative . . . . .	256
5.2.4.3	The perfect indicative . . . . .	257
5.2.4.4	The predicative independent indicating mood: The future indicative . . . . .	266
5.2.5	Summary . . . . .	266
<b>6</b>	<b>A basis for an automated translation</b>	<b>269</b>
6.1	Introduction . . . . .	269
6.2	An Introduction to Machine Translation . . . . .	270
6.2.1	System architecture and interfaces between modules . . . . .	272
6.2.2	Reversibility of resources and processes . . . . .	274
6.2.3	Developments in MT . . . . .	275
6.3	MT from Northern Sotho to English:	
	general lexical and structural issues . . . . .	275
6.3.1	Lexical ambiguities in the source language . . . . .	276
6.3.2	Structural ambiguities in the source language . . . . .	278
6.3.3	Differences in argument structure . . . . .	281
6.3.4	Translating the verbal relative / possessive . . . . .	283
6.3.5	Differences in word order . . . . .	284
6.3.6	Lack of determiners . . . . .	284
6.4	MT from Northern Sotho to English: Transfer phenomena of parts of speech	285
6.4.1	Nouns . . . . .	285
6.4.2	Concords . . . . .	286
6.4.3	Morphemes . . . . .	288
6.4.4	Particles . . . . .	289
6.4.5	Adjectives . . . . .	289
6.4.6	Adverbs . . . . .	289
6.4.7	Question words . . . . .	289



6.4.8	Verb stems . . . . .	290
6.4.9	Auxiliaries . . . . .	290
6.4.10	Copulatives . . . . .	291
6.4.11	Adjectives . . . . .	291
6.5	XLE in machine translation . . . . .	291
6.6	Summary . . . . .	293
<b>7</b>	<b>Summary and conclusions</b>	<b>295</b>
7.1	Aims of this study . . . . .	295
7.2	Summary of results . . . . .	296
7.2.1	Chapter 2: The word classes of Northern Sotho . . . . .	296
7.2.2	Chapter 3: A fragment of the grammar of Northern Sotho . . . . .	300
7.2.3	Chapter 4: Features of verbal phrases . . . . .	307
7.2.4	Chapter 5: Implementation of a grammar fragment . . . . .	309
7.2.5	Chapter 6: A basis for an automated translation . . . . .	309
7.3	Conclusions and future work . . . . .	310
	<b>Bibliography</b>	<b>312</b>

# List of Tables

2.1	Overview of the prefixes of the noun classes 1 to 4 and their referring annotations . . . . .	25
2.2	Overview of the prefixes of the noun classes 5, 14 and 6 and their referring annotations . . . . .	26
2.3	Overview of the prefixes of the noun classes 7 and 8 and their respective annotations . . . . .	27
2.4	Overview of the prefixes of the noun classes 9 and 10 and their referring annotations . . . . .	28
2.5	The emphatic pronouns . . . . .	37
2.6	The possessive pronouns . . . . .	38
2.7	The quantitative pronouns . . . . .	38
2.8	The three sets of subject concords . . . . .	41
2.9	The object concords . . . . .	43
2.10	The possessive concords . . . . .	44
2.11	The demonstrative concords (pronouns) . . . . .	45
2.12	The demonstrative copulatives and their variants . . . . .	46
2.13	Some examples of frequently used adjectives . . . . .	48
2.14	Derivations of the verb <i>gadika</i> ‘roast, thrash’ in GNSW . . . . .	54
2.15	Examples of auxiliary verbs . . . . .	55
2.16	Examples of verbs that may be used as auxiliaries . . . . .	56
2.17	Some copula of Northern Sotho . . . . .	56
2.18	The connective particle <i>na</i> fused with pronouns . . . . .	61
2.19	Question words of Northern Sotho . . . . .	64
2.20	Excerpt of the tagset: Miscellaneous part of speech . . . . .	65
2.21	The tagset of Northern Sotho 1 / 2 . . . . .	67

2.22	The tagset of Northern Sotho 2 / 2 . . . . .	68
3.1	Lombard's modal system . . . . .	75
3.2	Lombard's definition of independent moods compared with the respective constellations described by Poulos and Louwrens . . . . .	76
3.3	Lombard's definition of dependent moods compared with the respective constellations described by Poulos and Louwrens . . . . .	77
3.4	A schematic representation of the slot system . . . . .	79
3.5	The intransitive VBP (imperative VP) . . . . .	89
3.6	A transitive VBP (imperative VP) . . . . .	90
3.7	A double transitive VBP (imperative VP) . . . . .	90
3.8	A transitive VBP containing an object concord (imperative VP) . . . . .	90
3.9	A double transitive VBP containing an object concord (imperative VP) . . . . .	91
3.10	The non-predicative negative imperative VP . . . . .	91
3.11	The imperative mood (table contains all VBPs) . . . . .	92
3.12	The infinitive . . . . .	94
3.13	The three sets of subject concords . . . . .	96
3.14	Update on names of the constellations forming the VBP . . . . .	98
3.15	The long form of the predicative imperfect indicative . . . . .	98
3.16	The short and the negated form of the predicative imperfect indicative . . . . .	99
3.17	The perfect indicative . . . . .	101
3.18	The future indicative . . . . .	103
3.19	A summary of the independent indicative forms . . . . .	104
3.20	The present tense of the situative . . . . .	109
3.21	The perfect tense of the situative . . . . .	111
3.22	The future situative . . . . .	112
3.23	The present tense of the relative . . . . .	116
3.24	The perfect tense of the relative . . . . .	118
3.25	The future relative . . . . .	120
3.26	A summary of the modifying moods . . . . .	121
3.27	The consecutive . . . . .	122
3.28	The subjunctive/habitual . . . . .	124
3.29	Summary of the dependent moods . . . . .	125
3.30	Overview of copulative constellations . . . . .	128



3.31	groups of copulas . . . . .	131
3.32	The <b>stative</b> forms of the <b>identifying</b> copulative (COPSID) ( <b>present tense</b> , figure (3.22)) . . . . .	133
3.33	The <b>stative</b> forms of the <b>identifying</b> copulative (COPSID) ( <b>perfect tense</b> , figure (3.23)) . . . . .	134
3.34	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDID) ( <b>present tense</b> , Figure 3.24) . . . . .	136
3.35	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDID) ( <b>perfect tense</b> , Figure 3.25) . . . . .	137
3.36	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDID) ( <b>future tense</b> , Figure 3.26) . . . . .	138
3.37	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDIDD) ( <b>dependent constellations</b> part 1 of 2, Figure 3.27) . . . . .	139
3.38	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDIDD) ( <b>dependent constellations</b> part 2 of 2, Figure 3.27) . . . . .	141
3.39	The <b>dynamic</b> forms of the <b>identifying</b> copulative (COPDI) ( <b>non-predicative constellations</b> , Figure 3.28) . . . . .	141
3.40	The <b>stative</b> forms of the <b>descriptive</b> copulative (COPSDC) ( <b>present tense</b> , Figure 3.29) . . . . .	145
3.41	The <b>stative</b> forms of the <b>descriptive</b> copulative (COPSDC) ( <b>perfect tense</b> , Figure 3.30) . . . . .	146
3.42	The <b>dynamic</b> forms of the <b>descriptive</b> copulative (COPDDC) ( <b>present tense</b> , Figure 3.31) . . . . .	149
3.43	The <b>dynamic</b> forms of the <b>descriptive</b> copulative (COPDDC) ( <b>perfect tense</b> , Figure 3.32) . . . . .	150
3.44	The <b>dynamic</b> forms of the <b>descriptive</b> copulative (COPDID) ( <b>future tense</b> , Figure 3.33) . . . . .	151
3.45	The <b>dynamic</b> forms of the <b>descriptive</b> copulative (COPDDCD) ( <b>dependent constellations</b> , Figure 3.34) . . . . .	152
3.46	The <b>stative</b> forms of the <b>associative</b> copulative (COPSAS) ( <b>present tense</b> , Figure 3.39) . . . . .	158
3.47	The <b>stative</b> forms of the <b>associative</b> copulative (COPSAS) ( <b>perfect tense</b> , Figure 3.40) . . . . .	159

3.48	The <b>dynamic</b> forms of the <b>associative</b> copulative (COPDAS) ( <b>present tense</b> , Figure 3.41) . . . . .	162
3.49	The <b>dynamic</b> forms of the <b>associative</b> copulative (COPDAS) ( <b>perfect tense</b> , Figure 3.42) . . . . .	163
3.50	The <b>dynamic</b> forms of the <b>associative</b> copulative (COPDAS) ( <b>future tense</b> , Figure 3.43) . . . . .	164
3.51	The <b>dynamic</b> forms of the <b>associative</b> copulative (COPDACD) ( <b>dependent constellations</b> , Figures 3.44/3.45) . . . . .	166
3.52	The <b>dynamic</b> forms of the <b>associative</b> copulative (COPDA) ( <b>non-predicative constellations</b> , Figure 3.46) . . . . .	167
3.53	The auxiliary verbal phrase <sup>AUX</sup> VP . . . . .	170
3.54	The hortative constellation . . . . .	171
3.55	The potential forms . . . . .	173
3.56	Slots describing the noun phrase . . . . .	177
3.57	The basic noun phrase . . . . .	178
3.58	The extended noun phrase . . . . .	179
3.59	The basic noun phrase including <sup>pro</sup> NPs . . . . .	181
3.60	The possessive noun phrase <sup>pos</sup> NP . . . . .	184
4.1	The polysemy of <i>a</i> (taken from Faaß et al. (2009)) . . . . .	191
4.2	The remaining set of possible analyses when considering the VBP ending in <i>a</i> . . . . .	197
4.3	The remaining set of possible analyses when considering <i>a</i> . . . . .	198
4.4	The remaining analyses when considering <i>ba</i> – indicative . . . . .	199
4.5	Possible analyses when considering <i>ba</i> – auxiliary . . . . .	201
4.6	Verbal endings of Northern Sotho moods . . . . .	208
4.7	Feature selection of Northern Sotho moods: subject concords in slot zero-2 . . . . .	210
4.8	Feature selection of Northern Sotho moods: negation morphemes in slot zero-2 . . . . .	211
4.9	Feature selection of Northern Sotho moods: slot zero-1 . . . . .	212
4.10	Distribution of <i>ba</i> , <i>eba</i> , <i>bile</i> . . . . .	214
4.11	Distribution of <i>be</i> . . . . .	215
4.12	Distribution of <i>bago</i> , <i>bego</i> , <i>bilego</i> . . . . .	216
4.13	Distribution of <i>le</i> , <i>lego</i> . . . . .	216
4.14	Distribution of <i>na</i> , <i>(e)na</i> , <i>nago</i> . . . . .	217
4.15	Distribution of VCOP <sub>categ</sub> . . . . .	217

4.16	Distribution of <i>se, sego/seng</i> . . . . .	218
4.17	Distribution of constellations solely containing copulas . . . . .	219
4.18	Distribution of CSPERS and CSNEUT (CSPCSN) . . . . .	220
4.19	Distribution of CSPCSN and future tense morphemes . . . . .	220
4.20	Distribution of CSPCSN followed by a negation morpheme . . . . .	220
4.21	Distribution of CSPCSN followed by <i>ka se</i> . . . . .	220
4.22	Distribution of CSPCSN followed by auxiliary constellations . . . . .	221
4.23	Distribution of constellations beginning with $1CS_{\text{categ}}$ . . . . .	221
4.24	Distribution of $2CS_{\text{categ}}$ . . . . .	221
4.25	Distribution of $2CS_{\text{categ}}$ followed by future tense morphemes . . . . .	222
4.26	Distribution of $2CS_{\text{categ}}$ followed by negation morphemes and negation clusters	223
4.27	Distribution of $2CS_{\text{categ}}$ followed by auxiliary constellations . . . . .	223
4.28	Distribution of constellations beginning with $3CS_{\text{categ}}$ . . . . .	223
4.29	Constellations beginning with (negation) morphemes . . . . .	224
5.1	Functional equations of <i>monna o reka apola</i> . . . . .	236
5.2	Abbreviated version of functional equations of <i>monna o reka apola</i> . . . . .	237
5.3	Constellations forming the VBP . . . . .	246
6.1	Translations of the Northern Sotho verb <i>ja</i> into English . . . . .	277
7.1	The tagset of Northern Sotho 1 / 2 . . . . .	298
7.2	The tagset of Northern Sotho 2 / 2 . . . . .	299
7.3	Lombard's modal system . . . . .	300
7.4	A schematic representation of the slot system . . . . .	301
7.5	A summary of the independent indicative forms . . . . .	303
7.6	A summary of the modifying moods . . . . .	304
7.7	Summary of the dependent moods . . . . .	305
7.8	Overview of copulative constellations . . . . .	306
7.9	The hortative constellation . . . . .	307
7.10	The potential forms . . . . .	308

## List of Figures

1.1	Geographical extension of the Northern Sotho dialects in South Africa . . . .	1
1.2	Northern Sotho as a part of the South-Eastern Bantu languages . . . . .	2
1.3	Example parse trees in CFG . . . . .	16
2.1	First analysis: <i>ba rata go bala dipuku</i> ‘they like to read books’ . . . . .	29
2.2	Second analysis: <i>ba rata go bala dipuku</i> ‘they like to read books’ . . . . .	30
2.3	<i>ba rata go di bala</i> ‘they like to read them’ . . . . .	30
2.4	<i>go botša monna fela</i> ‘to tell (the) man only’ . . . . .	31
2.5	<i>go se dirwe</i> ‘not to be done’ . . . . .	31
2.6	<i>go ya go nyala</i> ‘going to marry’ . . . . .	32
3.1	First analysis of <i>nna ke tlo apea dijo</i> ‘I (personally) will cook (the) food’ . .	72
3.2	Second analysis of <i>nna ke tlo apea dijo</i> ‘I (personally) will cook (the) food’ .	73
3.3	Analysis (a): <i>ke rata mošemane yo o tlile maabane</i> *‘I like, this boy arrived yesterday’ . . . . .	81
3.4	Analysis (b): <i>ke rata mošemane yo o tlile maabane</i> ‘I like (the) boy, this one arrived yesterday’ . . . . .	81
3.5	Analysis (c): <i>ke rata mošemane yo o tlile maabane</i> *‘I like this boy, (some)one arrived yesterday’ . . . . .	82
3.6	Analysis (d): <i>ke rata mošemane yo o tlile maabane</i> ‘I like this boy, (some)one arrived yesterday’ . . . . .	82
3.7	Analysis (e): <i>ke rata mošemane yo o tlile maabane</i> *‘I like this boy, (some)one arrived yesterday’ . . . . .	82
3.8	<i>Nthuše!</i> ‘Help me!’ . . . . .	84
3.9	<i>Mphe puku!</i> ‘Give me the book!’ . . . . .	84
3.10	A discontinuous verb in <i>monna o a bo nwa</i> ‘(a) man drinks it’ . . . . .	86



3.11	<i>monna o nwa bjalwa</i> ‘(a) man drinks (a) beer’ . . . . .	87
3.12	<i>monna o a bo nwa</i> ‘(a) man drinks it’ . . . . .	88
3.13	<i>go sepela go re bontšha mafase</i> ‘travelling lets us see the world’ . . . . .	93
3.14	<i>go sepela go a lapiša</i> ‘walking is exhausting’ . . . . .	93
3.15	<i>ke ba bone ge ba tsena</i> ‘I saw them when they entered’ . . . . .	106
3.16	<i>ge ba boile ba tlo apea dijo</i> ‘when they have returned they will cook the food’	106
3.17	<i>Jesu o tlo boa ka wona mokgwa wo le mmonego a eya legodimong</i> ‘Jesus will return (the) very same way as you saw him when going to heaven’ . . . . .	108
3.18	<i>kgoši ye e bušago</i> ‘(a) chief who reigns’ . . . . .	114
3.19	Analysis of the English relative: ‘a chief who reigns’ . . . . .	114
3.20	<i>Thuthuthu ye monna o e ratago</i> ‘(a) motorbike that (a) man likes’ . . . . .	115
3.21	<i>kgoši ye e bušitšego</i> ‘(a) chief who reigned’ . . . . .	117
3.22	Identifying stative present tense (cf. Table 3.32) . . . . .	132
3.23	Identifying stative perfect tense (cf. Table 3.33) . . . . .	132
3.24	Identifying dynamic present tense (cf. Table 3.34) . . . . .	135
3.25	Identifying dynamic perfect tense (cf. Table 3.35) . . . . .	135
3.26	Identifying dynamic future tense (cf. Table 3.36) . . . . .	140
3.27	Identifying dynamic dependent clauses (cf. Tables 3.37 and 3.38) . . . . .	140
3.28	Identifying dynamic non-predicative clauses (cf. Table 3.39) . . . . .	140
3.29	Descriptive stative present tense (cf. Table 3.40) . . . . .	144
3.30	Descriptive stative perfect tense (cf. Table 3.41) . . . . .	144
3.31	Descriptive dynamic present tense (cf. Table 3.42) . . . . .	147
3.32	Descriptive dynamic perfect tense (cf. Table 3.43) . . . . .	148
3.33	Descriptive dynamic future tense (cf. Table 3.44) . . . . .	148
3.34	Identifying dynamic dependent clauses (cf. Tables 3.45) . . . . .	148
3.35	Identifying dynamic non-predicative clauses (cf. Table 3.39) . . . . .	153
3.36	<i>mo seswantšhong Mna Ramokgopa o na le mosadi wa gagwe</i> ‘on (the) pho- tograph, Mr. Ramokgopa is with his wife’ . . . . .	155
3.37	<i>bona ba na le polase</i> ‘these ones own (a) farm’ . . . . .	156
3.38	<i>ga ke na thuthuthu</i> ‘I don’t have (a) motorbike’ . . . . .	156
3.39	Associative stative present tense (cf. Table 3.46) . . . . .	157
3.40	Associative stative perfect tense (cf. Table 3.47) . . . . .	157
3.41	Associative dynamic present tense (cf. Table 3.48) . . . . .	160
3.42	Associative dynamic perfect tense (cf. Table 3.49) . . . . .	161

3.43	Associative dynamic future tense (cf. Table 3.50) . . . . .	161
3.44	Associative dynamic dependent clauses 1 of 2 (cf. Table 3.51) . . . . .	161
3.45	Associative dynamic dependent clauses 2 of 2 (cf. Tables 3.51) . . . . .	165
3.46	Associative dynamic non-predicative clauses (cf. Table 3.52) . . . . .	165
3.47	Example of a particle phrase (PP) . . . . .	185
3.48	An example analysis of two sentences connected with a conjunction . . . . .	187
3.49	An example analysis of a sentence containing a question . . . . .	188
3.50	A Northern Sotho question making use of the “wh question word” <i>eng</i> . . . . .	189
4.1	A partial analysis of ( <i>ge ba bona noga ba a</i> ) <i>e bolaya</i> ‘obj-3rd-cl14:9 kill’ . . . . .	196
4.2	A partial analysis of ( <i>ge ba bona noga ba</i> ) <i>a e bolaya</i> ‘subj-3rd(-cl1/-cl6)/pres/past obj-3rd-cl14:9 kill’ . . . . .	199
4.3	A first partial analysis of ( <i>ge ba bona noga</i> ) <i>ba a e bolaya</i> ‘subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	200
4.4	Hypothetical partial analysis 1 of ( <i>ge ba bona</i> ) <i>noga ba a e bolaya</i> ‘snake subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	202
4.5	Hypothetical partial analysis 2 of ( <i>ge ba bona</i> ) <i>noga ba a e bolaya</i> ‘snake subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	202
4.6	Hypothetical partial analysis 3 of ( <i>ge ba bona</i> ) <i>noga ba a e bolaya</i> ‘snake subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	203
4.7	Hypothetical partial analysis 1 of ( <i>ge</i> ) <i>ba bona noga ba a e bolaya</i> ‘obj-3rd-cl12 see (;) snake subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	204
4.8	Hypothetical partial analysis 2 of ( <i>ge ba</i> ) <i>bona noga ba a e bolaya</i> ‘see snake (;) subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	204
4.9	Hypothetical partial analysis 3 of ( <i>ge ba</i> ) <i>bona noga ba a e bolaya</i> ‘see snake subj-3rd-cl12 pres obj-3rd-cl14:9 kill’ . . . . .	205
4.10	Resulting analysis of <i>ge ba bona noga ba a e bolaya</i> ‘when <i>subj-3rd-cl12</i> see snake subj-3rd-cl12 pres obj-3rd-cl19 kill’ . . . . .	206
5.1	Parallel structures of LFG . . . . .	228
5.2	Basic sentence, consisting of a NP and a VP . . . . .	232
5.3	The c-structure of <i>monna o reka apola</i> . . . . .	233
5.4	Numbered nodes at c-structure of <i>monna o reka apola</i> . . . . .	234
5.5	Instantiation of metavariables ( $\uparrow$ and $\downarrow$ ) of f-structures in the c-structure of <i>monna o reka apola</i> . . . . .	235

5.6	The f-structures of <i>monna o reka apola</i> . . . . .	237
5.7	F-structure containing a saturated transitive verb <i>monna o a ipshina</i> . ‘(a) man enjoys himself.’ . . . . .	240
5.8	C-structure of a positive imperative intransitive <i>Bolela!</i> ‘Speak!’ . . . . .	251
5.9	F-structure of a positive imperative intransitive <i>Bolela!</i> ‘Speak!’ . . . . .	252
5.10	C-structure of a positive imperative transitive <i>Bulang lemati!</i> ‘Close (the) door!’ . . . . .	252
5.11	F-structure of a positive imperative transitive <i>Bulang lemati!</i> ‘Close (the) door!’ . . . . .	253
5.12	C-structure of a positive imperative double transitive <i>Efa monna puku!</i> ‘Give (a) man (a) book!’ . . . . .	253
5.13	F-structure of a positive imperative double transitive <i>Efa monna puku!</i> ‘Give (a) man (a) book!’ . . . . .	254
5.14	C-structure of a negated imperative double transitive <i>Se fe monna puku!</i> ‘Do not give (a) man (a) book!’ . . . . .	254
5.15	F-structure of a negated imperative double transitive <i>Se fe monna puku!</i> ‘Do not give (a) man (a) book!’ . . . . .	255
5.16	Packed f-structure of a positive indicative, present tense, subject not present <i>o a bolela</i> . ‘(s)he/it speaks / you speak.’ . . . . .	258
5.17	C-structure of a positive indicative, present tense, subject not present <i>o a bolela</i> . ‘(s)he/it speaks / you speak.’ . . . . .	259
5.18	C-structure of a positive indicative, perfect tense <i>lesogana le e rekile</i> . ‘(a) young man bought it/them.’ . . . . .	260
5.19	Packed f-structure of a positive indicative, perfect tense <i>lesogana le e rekile</i> . ‘(a) young man bought it/them.’ . . . . .	261
5.20	C-Structure of a negated indicative, perfect tense <i>lesogana ga se la bolela</i> . ‘(a) young man did not speak.’ . . . . .	262
5.21	F-Structure of a negated indicative, perfect tense <i>lesogana ga se la bolela</i> . ‘(a) young man did not speak.’ . . . . .	262
5.22	C-Structure of a negated indicative, perfect tense <i>lesogana ga se le bolele</i> . ‘(a) young man did not speak.’ . . . . .	263
5.23	F-Structure of a negated indicative, perfect tense <i>lesogana ga se le bolele</i> . ‘(a) young man did not speak.’ . . . . .	263

5.24	C-Structure of a negated indicative, perfect tense <i>lesogana ga la bolela</i> . ‘(a) young man did not speak.’ . . . . .	264
5.25	F-Structure of a negated indicative, perfect tense <i>lesogana ga la bolela</i> . ‘(a) young man did not speak.’ . . . . .	264
5.26	C-Structure of a negated indicative, perfect tense <i>lesogana ga le a bolela</i> . ‘(a) young man did not speak.’ . . . . .	265
5.27	F-Structure of a negated indicative, perfect tense <i>lesogana ga le a bolela</i> . ‘(a) young man did not speak.’ . . . . .	265
5.28	C-structure of a positive indicative, future tense <i>mmutla o tlo tšhaba</i> . ‘(a) hare will flee.’ . . . . .	267
5.29	F-structure of a positive indicative, future tense <i>mmutla o tlo tšhaba</i> . ‘(a) hare will flee.’ . . . . .	267
5.30	C-structure of a negated indicative, future tense <i>mmutla o ka se tšhabe</i> . ‘(a) hare will not flee.’ . . . . .	268
5.31	F-structure of a negated indicative, future tense <i>mmutla o ka se tšhabe</i> . ‘(a) hare will not flee.’ . . . . .	268
6.1	The MT triangle . . . . .	271
6.2	Translating with the transfer system EUROTRA . . . . .	273
6.3	Example analysis: PP attachment (top node) . . . . .	279
6.4	Example analysis: second PP attached to first PP . . . . .	279
6.5	Example analysis: pp attachment (top node) . . . . .	280
6.6	Example analysis: pp attachment (second NP of coordination) . . . . .	280
6.7	Example analysis: ‘VP attachment’ (top node) . . . . .	280
6.8	Example analysis: ‘VP attachment’ (second NP of coordination) . . . . .	281
6.9	Simplified f-structure of <i>Tate o nyakela bana malekere</i> ). . . . .	283
6.10	Simplified f-structure of ‘Father looks for sweets for the children’). . . . .	283
6.11	Simplified f-structure of <i>ba go bala dipuku</i> ). . . . .	286
6.12	Simplified f-structure of <i>ba bego ba bolela</i> . . . . .	291
6.13	f-structure of <i>tate</i> . . . . .	293
6.14	f-structure of ‘father’ . . . . .	293