UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Lexicographic Perspectives on the use of Sepedi as a High Function Language

by

Diapo Nelson Lekganyane

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Literature

in the Faculty of Humanities
University of Pretoria
Pretoria
May 2001

Promoter :Prof.D.J.Prinsloo
Co-Promoter :Prof.A.Carstens

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TABLE OF CONTENTS

		PAGES
CHAPTER 1	**INTRODUCTION**	
1.1	Introduction	1
1.2	Research questions	1
1.3	Objectives of the study	2
1.4	Methodology	4
1.4.1	Theoretical framework	4
1.4.2	Data collection	4
1.5	Preview	5
CHAPTER 2	**LANGUAGE POLICY AND THE REVALORIZATION OF THE AUTOCHTHONOUS LANGUAGES**	
2.1	Introduction	9
2.2	Language policy and practices	10
2.3	Language principles in the constitution	11
2.4	Language stipulations	13
2.5	Language planning	15
2.5.1	Status planning	19
2.5.1.1	Statutory and governmental measures	19
2.5.1.2	An increase in the economic value of a language	20

2.5.1.3 Educational value 21

2.5.1.4 The Socio-cultural meaning of language 22

2.5.1 Corpus planning 23

2.5.2 Acquisition planning 29

2.6 Conclusion 30

CHAPTER 3 LEXICOGRAPHIC APPROACHES TO CORPUS BUILDING

3.1 Introduction 33

3.2 The Cobuild Corpus 34

3.2.1 Aim 34

3.2.2 The composition of the text categories 34

3.2.3 Frequency counting 36

3.2.4 Concordances 36

3.3 The Lancaster-Oslo/Bergen (LOB) Corpus 37

3.3.1 Aim 37

3.3.2 The composition of the text categories 38

3.3.3 Processing of the LOB Corpus 39

3.3.3.1 Frequency counting 39

3.3.3.2 Concordance 40

3.4 Longman Lancaster English Language Corpus 41

3.4.1 Aim 41

3.4.2	The composition of the text categories	42
3.4.3	Processing of the corpus	45
3.4.3.1	Frequency counting	45
3.4.3.2	Concordance	45
3.5	Comparison: schematic representation of the Cobuild Corpus, LOB and Longman English Language Corpus	47
3.6	The compilation of the English high function corpus	48
3.6.1	Introduction	48
3.6.2	Selection of the text categories	48
3.6.3	Scanning	49
3.6.4	Organising the data	49
3.6.5	Alphabetic ordering	50
3.6.6	Word frequency counts	50
3.6.7	Concordance	51
3.6.8	Frequently used academic words	52
3.6.9	Low frequency academic words	53
3.6.10	Conclusion	54
3.7	Aim	54
3.8	Organising principles	55
3.9	Data collection and information processing	56

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

CHAPTER 4 THE TREATMENT OF HIGH FUNCTION WORDS IN SEPEDI BILINGUAL DICTIONARIES

4.1.	Introduction	57
4.2.	Equivalence	58
4.2.1.	Absolute equivalence	58
4.2.2.	Partial equivalence	59
4.2.2.1	Divergence	60
4.2.2.2	Convergence	65
4.2.2.2.1	Lexical convergence	65
4.2.3	Zero equivalence	66
4.2.3.1	Linguistic gaps	66
4.2.3.2	Referential gaps	67
4.3	A sample of English high function words	71
4.3.1	English high function words and the translation equivalents given in the existing Sepedi bilingual dictionaries	72
4.3.1.1	*Area*	74
4.3.1.2	*Assume*	77
4.3.1.3	*Assign*	79
4.3.1.4	*Expert*	82
4.3.1.5	*Policy*	84
4.4.	Conclusion	87

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

CHAPTER 5 **LEXICAL GAPS IN SEPEDI CONCERNING HIGH FUNCTION CONCEPTS**

5.1 Introduction 89

5.2 English high function words 90

5.2.1 *Creditor* 92

5.2.2 *Equilibrium* 94

5.2.3 *Hierarchy* 96

5.2.4 *Innovation* 98

5.2.5 *Relevance* 100

5.3. Conclusion 101

CHAPTER 6 **CONCLUSION** 103

BIBLIOGRAPHY 108

APPENDICES i

LEXICOGRAPHIC PERSPECTIVES ON THE USE OF SEPEDI AS A HIGH-FUNCTION LANGUAGE

SUMMARY OF STUDY

The study was aimed at establishing in a scientific way the extent to which Sepedi is able to fulfill its role as one of the official languages in South Africa. The primary research question was whether **Sepedi could function as a medium of communication in all the higher domains of life.**

The research question was addressed by firstly making a study of the language clauses and the language stipulations of the South African Constitution. The second step was to situate the research within the theoretical paradigm of language development and language planning and to take cognizance of both international and local perspectives in this field.

In order to plan the way forward it was necessary to determine the exact position of Sepedi as far as the adequacy of its lexicon was concerned.

As a result of the fact that there are very few written sources in Sepedi regarding some of the high function domains it was decided to compile a corpus of South African English high function words to form the basis for determining the lexical capacity of Sepedi in higher function domains. The procedure for building a corpus was determined by studying the methodology and praxis of three well-known international corpora of English. Following the same basic strategies as the compilers of these corpora an English High Function Corpus was compiled. By making use of standard techniques in computational lexicography frequencies and spreading were determined, and keywords were studied in context.

In order to establish whether Sepedi does indeed have translation equivalents for the English high function words, and has the lexical capacity to act as a fully-fledged official language, the most appropriate sources to be used as diagnostic instruments were found to be the existing bilingual English-Sepedi dictionaries. The treatment of lexical items in the English-Sepedi bilingual dictionaries led to the conclusion that only 8 of the English high function words were not entered in any of the three dictionaries studied. The investigation also provided valuable insight into the quality and comprehensiveness of these dictionaries and indicated that they are in dire need of revision. Language planners and lexicographers of Sepedi are advised to make use of the heuristic and diagnostic tools available in modern-day metalexicography, but also to make sure that all lexicographical work reflects the linguistic reality within the speech community.

'N LEKSIKOGRAFIESE PERSPEKTIEF OP DIE GEBRUIK VAN SEPEDI AS 'N HOËR FUNKSIE TAAL

SAMEVATTING VAN DIE STUDIE

Die doel van die studie was om op 'n wetenskaplike wyse te bepaal tot watter mate Sepedi in staat is om haar rol as een van die amptelike landstale van Suid-Afrika te vervul. Die basiese navorsingvraag was of **Sepedi kan funksioneer as kommunikasiemedium in alle hoër domeine van kommunikasie**.

Ter beantwoording van die navorsingsvraag is in die eerste plek 'n studie gemaak van die taalstipulasies en –bepalings in die Suid-Afrikaanse Konstitusie. Die tweede stap was om die navorsing te posisioneer binne die teoretiese paradigma van taalontwikkeling en taalbeplanning en om deeglik kennis te neem van beide internasionale en nasionale perspektiewe op hierdie gebied.

Ten einde instaat te wees om die pad vorentoe uit te stippel was dit nodig om presies ter bepaal hoe toerenkend die leksikon van Sepedi is.

Aangesien daar baie min geskrewe bronne van hoër domeine van kommunikasie in Sepedi bestaan, is besluit om 'n korpus van hoër domeine van kommunikasie vir Suid-Afrikaanse Engels saam te stel. Hierdie woorde het die basis gevorm vir die bepaling van die leksikale kapasiteit van Sepedi in hoër domeine van kommunikasie. Die wyse waarop die korpus saamgestel is is bepaal deur bestudering van die metodologie en praksis van drie welbekende internasionale korpusse vir Engels. Die Suid-Afrikaanse korpus van hoër domeine van kommunikasie vir Engels is vervolgens in navolging van dieselfde basiese strategieë wat vir eersgenoemde korpusse gevolg is saamgestel.

Wat betref die vraag Sepedi werklik vertaalekwivalente vir Engelse hoër funksie woorde het, asook oor die leksikale kapasiteit beskik om as 'n volwaardige amptelike landstaal te funksioneer, is bevind dat die aangewese bronne wat as diagnostiese

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

instrumente geskik is, bestaande tweetalige Engels-Sepedi woordeboeke is. Uit die bewerking van leksikale items in die Engels-Sepedi tweetalige woordeboeke is gekonkludeer dat slegs 8 van die Engelse hoër funksie woorde nie in enigeen van die drie woordeboeke opgeneem is nie. Die ondersoek het ook belangrike insigte gebring ten opsigte van die kwaliteit en omvattendheid van hierdie woordeboeke en dat hersiening daarvan dringend nodig is.