

Lexikos 19 Supplement

Lexikos 19 Supplement

Redakteur

Editor

J.C.M.D. du Plessis



African Association for Lexicography

AFRILEX-REEKS 19:2009

AFRILEX SERIES 19:2009



BURO VAN DIE WAT

STELLENBOSCH

Die publikasie van hierdie boek is moontlik gemaak deur 'n ruim skenking van die L.W. Hiemstra-Trust – Opgerig deur Riekie Hiemstra ter herinnering aan Ludwig Wybren (Louis) Hiemstra.

The publication of this book was made possible by a generous donation from the L.W. Hiemstra Trust – Established by Riekie Hiemstra in memory of Ludwig Wybren (Louis) Hiemstra.

Uitgewer Publisher

BURO VAN DIE WAT
Posbus 245
7599 STELLENBOSCH

Kopiereg © 2009 deur die uitgewer
Alle regte streng voorbehou
Eerste uitgawe 2009

Tipografie en uitleg
deur Tanja Harteveld en Riette Ruthven
Bandontwerp deur Piet Grobler
Geset in 10 op 12 pt Palatino

Gedruk en gebind deur Onyx Press
Stewartstraat 61 Goodwood

ISBN: 978-0-9814434-0-9
ISSN 1684-4904

Geen gedeelte van hierdie publikasie mag sonder skriftelike verlof van die uitgewer gereproduseer of in enige vorm of deur enige elektroniese of meganiese middel weergegee word nie, hetsy deur fotokopiëring, plaat- of bandopname, mikroverfilming of enige ander stelsel van inligtingsbewaring

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, including electronic, mechanical, photographic, magnetic or other means, without the prior written permission of the publisher

Menings wat in artikels en resensies uitgespreek word, is nie noodwendig dié van AFRILEX of die Buro van die WAT nie

Opinions expressed in the articles and reviews are not necessarily those of AFRILEX or of the Bureau of the WAT

Lexikos is elektronies beskikbaar by Sabinet Online en African Journals OnLine (AJOL)
Lexikos is electronically available from Sabinet Online and African Journals OnLine (AJOL)

Indekse Indexes

Arts and Humanities Citation Index®, Current Contents®/Arts & Humanities, Current Contents®/Social and Behavioral Sciences, Journal Citation Reports/Social Sciences Edition, Social Sciences Citation Index®, and Social Scisearch®; Linguistic Bibliography Online; Linguistics Abstracts Online; Linguistics and Language Behavior Abstracts; MLA International Bibliography; R.R.K. Hartmann's Bibliography of Lexicography; Scopus

Adviesraad / Advisory Board

Prof. A. Adamska-Salaciak (Pole/Poland)
Prof. H. Béjoint (Frankryk/France)
Prof. H. Chimhundu (Zimbabwe)
Prof. F. Dolezal (VSA/USA)
Prof. R.H. Gouws (RSA)
Prof. R.R.K. Hartmann (Groot-Brittanje/Great Britain)
Prof. M.H. Heliel (Egipte/Egypt)
Dr. V. Kukanda (Gaboen/Gabon)
Prof. W. Martin (België en Nederland/Belgium and The Netherlands)
Prof. I.A. Mel'čuk (Kanada/Canada)
Prof. A.M.F.J. Moerdijk (Nederland/The Netherlands)
Prof. M. Schlaefer (Duitsland/Germany)
Prof. J. Taeldeman (België/Belgium)
Dr. J. Tent (Australië/Australia)
Prof. P.G.J. van Sterkenburg (Nederland/The Netherlands)
Prof. L.S. Vikør (Noorweë/Norway)
Prof. H.E. Wiegand (Duitsland/Germany)

Redaksiekomitee / Editorial Committee

Prof. H.M. Batibo (Botswana)
Dr. H.L. Beyer (Namibië/Namibia)
Prof. W.A.M. Carstens (RSA)
Dr. E. Chabata (Zimbabwe)
Dr. A.R. Chuwa (Tanzanië/Tanzania)
Prof. C.J. Conradie (RSA)
Prof. J.D. Emejulu (Gaboen/Gabon)
Prof. A.E. Feinauer (RSA)
Prof. R. Finlayson (RSA)
Dr. S. Hadebe (Zimbabwe)
Prof. I.M. Kosch (RSA)
Dr. P.A. Louw (RSA)
Mnr. K.J. Mashamaite (RSA)
Dr. P.A. Mavoungou (Gaboen/Gabon)
Dr. J.S. Mdee (Tanzanië/Tanzania)
Prof. B.M. Mini (RSA)
Dr. V.M. Mojela (RSA)
Mnr. M.C. Mphahlele (RSA)
Mnr. M.H. Mpungose (RSA)
Dr. H.S. Ndinga-Koumba-Binza (Gaboen/Gabon)
Prof. D.J. Prinsloo (RSA)
Prof. P.H. Swanepoel (RSA)

Inhoud / Contents

Voorwoord	viii
Foreword	x
<i>J.C.M.D. du Plessis</i>	
Redaksionele doelstellings	xii
Editorial Objectives	xiii
Redaktionelle Ziele	xiv
Politique éditoriale	xv

Artikels / Articles

Linguistic Theory in the Practical Lexicography of the African Languages	1
<i>Emmanuel Chabata</i>	
Users, User-Friendliness and Projected Uses of <i>Isichazamazwi SesiNdebele</i> : An Analysis	16
<i>Langa Khumalo</i>	
Divergent Approaches to Corpus Processing: The Need for Standardisation	26
<i>Esau Mangoya</i>	
The Microstructure of <i>Isichazamazwi SesiNdebele</i>	38
<i>Mandlenkosi Maphosa and Dion Nkomo</i>	
Not Mere Lexicographic Cosmetics: The Compilation and Structural Features of <i>Isichazamazwi SezoMculo</i>	51
<i>Nobuhle Moyo and Dion Nkomo</i>	
Sense Relations in the Treatment of Meaning in <i>Isichazamazwi SesiNdebele</i>	71
<i>Eventhough Ndlovu</i>	
The Treatment of Borrowed Nouns in <i>Isichazamazwi SesiNdebele</i> and <i>Isichazamazwi SezoMculo</i>	86
<i>Eventhough Ndlovu</i>	

Projekte / Projects

Looking beyond Meaning in the <i>Advanced Ndebele Dictionary</i> <i>Langa Khumalo</i>	102
The Compilation of a Shona Children's Dictionary: Challenges and Solutions <i>Peniah Mabaso</i>	112
Target Users' Expectations versus the Actual Compilation of a Shona Children's Dictionary <i>Esau Mangoya</i>	120
<i>Duramazwi reDudziramutauro neUvaranomwe</i> against the Background of Shona Lexicography <i>Gift Mheta and Maxwell Kadenge</i>	134

Leksikonotas / Lexiconotes

The Treatment of Encyclopedic Information in <i>Duramazwi</i> <i>reMimhanzi</i> <i>Gift Mheta</i>	144
The Use of Loan Translation as a Term-Creation Strategy in <i>Duramazwi reMimhanzi</i> <i>Gift Mheta and Itai Muhwati</i>	150
Phonological Adaptation of Borrowed Terms in <i>Duramazwi</i> <i>reMimhanzi</i> <i>Gift Mheta and William Zivenge</i>	157
Affirming Verb Lexemes in <i>A Practical Ndebele Dictionary</i> and in <i>Isichazamazwi SesiNdebele: The Case of -wa Verbs</i> <i>Jerry Zondo</i>	166
Voorskrifte aan Skrywers	175
Instructions to Authors	176
Richtlinien für Autoren	177
Instructions aux Auteurs	178

Voorwoord

In *Lexikos* 17 verskyn 'n spesiale afdeling "Leksikofokus" bestaande uit tien artikels wat handel oor publikasies van die African Languages Research Institute (ALRI) by die Universiteit van Zimbabwe, Harare. Hierdie *Supplement* tot *Lexikos* 19 is 'n voortsetting van daardie afdeling waarin huidige en vroeëre ALRI-personeellede die probleme bespreek wat hulle teëgekóm en die oplossings wat hulle hiervoor gevind het tydens die samestelling van baanbrekerswoordeboeke in die twee hoof- inheemse Zimbabwiese tale, Sjona en Ndebele.

Alhoewel die skrywers vanuit 'n praktiese gesigspunt skryf, hou hulle altyd die teoretiese benaderingswyses in gedagte wat die basis vorm vir die voortbrenging van algemene en gespesialiseerde woordeboeke. Hierdie artikels vul mekaar aan deur die verskillende aspekte van elke woordeboek te belig. Die skrywers huiwer nie om die waarde van hulle eie en hul kollegas se produkte te beoordeel deur hul werkswyse en werksmetodes te herbeskou en te herbeoordeel nie. Lesers word hier vyftien artikels van verskillende lengte en omvang aangebied wat in baie gevalle beskou kan word as werksdokumente wat waardevolle inligting bevat vir ander leksikograwe wat beoog om die samestelling van soortgelyke woordeboeke in ander Afrikatale te onderneem.

ALRI het twee algemene eentalige Sjonawoordeboeke gepubliseer: *Duramazwi reChiShona* (1996) en *Duramazwi Guru reChiShona* (2001). Esau Mangoya skryf oor die behoefte aan standaardisering by korpusprosessering, terwyl Emmanuel Chabata linguïstiese teorie in die praktiese leksikografie bespreek deur voorbeelde van artikels in hierdie woordeboeke te gebruik.

Behalwe hierdie algemene eentalige woordeboeke is ook 'n Sjonakinderwoordeboek voltooi. Peniah Mabaso gee 'n kort oorsig van die uitdagings waarvoor die span samestellers te staan gekóm het, terwyl Esau Mangoya uitvoerig beskryf hoe die voorstelle van die teikengebruikers benut is by die samestelling van die woordeboek.

Van die drie gespesialiseerde Sjonawoordeboeke wat sover verskyn het, word verskillende aspekte van twee van hulle bespreek, naamlik *Duramazwi reMimhanzi* (2005) en *Duramazwi reDudziramutauro neUvaranomwe* (2007). Drie notas oor *Duramazwi reMimhanzi* word ingesluit: Gift Mheta skryf oor die behandeling van ensiklopediese inligting, terwyl hy saam met Itai Muhwati die gebruik van leenvertaling as 'n termskeppende strategie en saam met William Zivenge die fonologiese aanpassing van ontleende terme behandel. *Duramazwi reDudziramutauro neUvaranomwe* word deur Gift Mheta en Maxwell Kadenge teen die agtergrond van Sjonaleksikografie gehandel.

Tot sover is een algemene eentalige Ndebelewoordeboek *Isichazamazwi SesiNdebele* (2001) deur ALRI gepubliseer. Drie artikels word aan hierdie woordeboek gewy: Langa Khumalo behandel sy gebruikers, gebruikersvriendelikheid en beoogde gebruike, Mandlenkosi Maphosa en Dion Nkomo bespreek sy

mikrostruktuur en Eventhough Ndlovu behandel betekenisverwantskappe by die behandeling van betekenis daarin. In 'n gespesialiseerde artikel bespreek Jerry Zondo die behandeling van **-wa**-werkwoorde in hierdie woordeboek in vergelyking met *A Practical Ndebele Dictionary* (1971).

Die enigste gespesialiseerde Ndebele woordeboek wat tot sover verskyn het, is *Isichazamazwi SezoMculo* (2006). Nobuhle Moyo en Dion Nkomo skryf uitvoerig oor sy samestelling en strukturele eienskappe. Eventhough Ndlovu bespreek die lemmatisering van geleende substantiewe in *Isichazamazwi Sesi-Ndebele* en *Isichazamazwi SezoMculo*.

Laastens verduidelik Langa Khumalo hoe ander aspekte as betekenis in die beplande *Gevorderde Ndebelewoordeboek* behandel sal word.

J.C.M.D. du Plessis

Buro van die Woordeboek van die Afrikaanse Taal

Foreword

In *Lexikos* 17 appears a special section "Lexicofocus" consisting of ten articles which deal with publications by the African Languages Research Institute (ALRI) at the University of Zimbabwe, Harare. This *Supplement* to *Lexikos* 19 is a continuation of that section, in which present and former ALRI staff members discuss the problems they encountered and the solutions they found for these while compiling pioneer dictionaries in the two major indigenous Zimbabwean languages, Shona and Ndebele.

Although the authors write from a practical viewpoint, they always keep in mind the theoretical methods of approach which form the basis for producing general and specialized dictionaries. These articles complement each other, elucidating different aspects of each dictionary. The authors do not hesitate to decide the merits of their own and their colleagues' products, reconsidering and re-evaluating their procedures and techniques. Readers are here presented with fifteen articles of varying length and comprehensiveness which, in many cases, can be seen as work documents containing valuable information for other lexicographers who aim to undertake the compilation of similar dictionaries in other African languages.

ALRI has published two general monolingual Shona dictionaries: *Duramazwi reChiShona* (1996) and *Duramazwi Guru reChiShona* (2001). Esau Mangoya writes about the need for standardization in corpus processing, while Emmanuel Chabata discusses linguistic theory in practical lexicography using examples of articles from these dictionaries.

Except for these general monolingual dictionaries a Shona Children's Dictionary has been completed. Peniah Mabaso gives a short overview of the challenges with which the compiling team was confronted, while Esau Mangoya describes comprehensively how the suggestions of the target users were utilized in the compilation of the dictionary.

Of the three specialized Shona dictionaries that have appeared so far, different aspects of two of these are discussed, namely *Duramazwi reMimhanzi* (2005) and *Duramazwi reDudziramutauo neUvaranomwe* (2007). Three notes about *Duramazwi reMimhanzi* are included: Gift Mheta writes about the treatment of encyclopedic information, while together with Itai Muhwati the use of loan translation as a term-creation strategy and together with William Zivenge the phonological adaptation of borrowed terms is discussed. *Duramazwi reDudziramutauo neUvaranomwe* is treated by Gift Mheta and Maxwell Kadenge against the background of Shona lexicography.

So far one general monolingual Ndebele dictionary *Isichazamazwi SesiNdebele* (2001) has been published by ALRI. Three articles are devoted to this dictionary: Langa Khumalo deals with its users, user-friendliness and projected uses, Mandlenkosi Maphosa and Dion Nkomo discuss its microstructure and

Eventhough Ndlovu deals with sense relations in its treatment of meaning. In a specialized article, Jerry Zondo discusses the handling of **-wa** verbs in this dictionary in comparison with *A Practical Ndebele Dictionary* (1971).

The only specialized Ndebele dictionary that has appeared so far is *Isichazamazwi SezoMculo* (2006). Nobuhle Moyo and Dion Nkomo write extensively about its compilation and structural features. Eventhough Ndlovu discusses the lemmatization of borrowed nouns in *Isichazamazwi SesiNdebele* and *Isichazamazwi SezoMculo*.

Lastly, Langa Khumalo explains how other aspects than meaning would be handled in the planned *Advanced Ndebele Dictionary*.

J.C.M.D. du Plessis

Bureau of the Woordeboek van die Afrikaanse Taal

Redaksionele doelstellings

Lexikos is 'n tydskrif vir die leksikografiese vakspesialis en word in die AFRILEX-reeks uitgegee. "AFRILEX" is 'n akroniem vir "leksikografie in en vir Afrika". Van die sesde uitgawe af dien *Lexikos* as die amptelike mondstuk van die *African Association for Lexicography* (AFRILEX), onder meer omdat die Buro van die WAT juis die uitgesproke doel met die uitgawe van die AFRILEX-reeks gehad het om die stigting van so 'n leksikografiese vereniging vir Afrika te bevorder.

Die strewe van die AFRILEX-reeks is:

- (1) om 'n kommunikasiekanaal vir die nasionale en internasionale leksikografiese gesprek te skep, en in die besonder die leksikografie in Afrika met sy ryk taleverskeidenheid te dien;
- (2) om die gesprek tussen leksikograwe onderling en tussen leksikograwe en taalkundiges te stimuleer;
- (3) om kontak met plaaslike en buitelandse leksikografiese projekte te bewerkstellig en te bevorder;
- (4) om die interdisiplinêre aard van die leksikografie, wat ook terreine soos die taalkunde, algemene taalwetenskap, leksikologie, rekenaarwetenskap, bestuurskunde, e.d. betrek, onder die algemene aandag te bring;
- (5) om beter samewerking op alle terreine van die leksikografie moontlik te maak en te koördineer, en
- (6) om die doelstellings van die *African Association for Lexicography* (AFRILEX) te bevorder.

Hierdie strewe van die AFRILEX-reeks sal deur die volgende gedien word:

- (1) Bydraes tot die leksikografiese gesprek word in die vaktydskrif *Lexikos* in die AFRILEX-reeks gepubliseer.
- (2) Monografiese en ander studies op hierdie terrein verskyn as afsonderlike publikasies in die AFRILEX-reeks.
- (3) Slegs bydraes wat streng vakgerig is en wat oor die suiwer leksikografie of die raakvlak tussen die leksikografie en ander verwante terreine handel, sal vir opname in die AFRILEX-reeks kwalifiseer.
- (4) Die wetenskaplike standaard van die bydraes sal gewaarborg word deur hulle aan 'n komitee van vakspesialiste van hoë akademiese aansien voor te lê vir anonieme keuring.

Lexikos sal jaarliks verskyn, terwyl verdienstelike monografiese studies sporadies en onder hulle eie titels in die AFRILEX-reeks uitgegee sal word.

Editorial Objectives

Lexikos is a journal for the lexicographic specialist and is published in the AFRILEX Series. "AFRILEX" is an acronym for "lexicography in and for Africa". From the sixth issue, *Lexikos* serves as the official mouthpiece of the *African Association for Lexicography* (AFRILEX), amongst other reasons because the Bureau of the WAT had the express aim of promoting the establishment of such a lexicographic association for Africa with the publication of the AFRILEX Series.

The objectives of the AFRILEX Series are:

- (1) to create a vehicle for national and international discussion of lexicography, and in particular to serve lexicography in Africa with its rich variety of languages;
- (2) to stimulate discourse between lexicographers as well as between lexicographers and linguists;
- (3) to establish and promote contact with local and foreign lexicographic projects;
- (4) to focus general attention on the interdisciplinary nature of lexicography, which also involves fields such as linguistics, general linguistics, lexicology, computer science, management, etc.;
- (5) to further and coordinate cooperation in all fields of lexicography; and
- (6) to promote the aims of the *African Association for Lexicography* (AFRILEX).

These objectives of the AFRILEX Series will be served by the following:

- (1) Contributions to the lexicographic discussion will be published in the specialist journal *Lexikos* in the AFRILEX Series.
- (2) Monographic and other studies in this field will appear as separate publications in the AFRILEX Series.
- (3) Only subject-related contributions will qualify for publication in the AFRILEX Series. They can deal with pure lexicography or with the intersection between lexicography and other related fields.
- (4) Contributions are judged anonymously by a panel of highly-rated experts to guarantee their academic standard.

Lexikos will be published annually, but meritorious monographic studies will appear as separate publications in the AFRILEX Series.

Redaktionelle Ziele

Lexikos ist eine Zeitschrift für Fachleute der Lexikographie, die in der AFRILEX-Serie erscheint. "AFRILEX" ist ein Akronym für "Lexikographie in und für Afrika". Von der sechsten Ausgabe an dient *Lexikos* als amtliches Sprachrohr der *African Association for Lexicography* (AFRILEX), u.a. weil das Büro des WAT das gerade angesprochene Ziel mit der Ausgabe der AFRILEX-Serie verfolgt, die Gründungsziele eines solchen lexikographischen Vereins für Afrika zu fördern.

Die folgenden Ziele werden mit den Publikationen der AFRILEX-Serie verfolgt: Man möchte:

- (1) ein Medium schaffen für die nationale und internationale Diskussion, besonders aber der Lexikographie in Afrika mit seinen zahlreichen Sprachen dienen;
- (2) die Diskussion fördern, unter Lexikographen als auch zwischen Lexikographen und Linguisten;
- (3) Kontakt herstellen und fördern zwischen südafrikanischen und ausländischen lexikographischen Projekten;
- (4) die Aufmerksamkeit lenken auf die interdisziplinäre wissenschaftliche Praxis der Lexikographie, die Beziehung aufweist zur Linguistik, allgemeinen Sprachwissenschaft, Lexikologie, Computerwissenschaft, zum Management und zu anderen Bereichen;
- (5) die Zusammenarbeit auf allen Gebieten der Lexikographie fördern und koordinieren;
- (6) die Ziele der *African Association for Lexicography* (AFRILEX) fördern.

Gemäß den Zielsetzungen der AFRILEX-Serie werden:

- (1) Beiträge zur lexikographischen Diskussion in der Fachzeitschrift *Lexikos* veröffentlicht;
- (2) monographische und andere Studien auf diesem Gebiet als getrennte Publikationen in der AFRILEX-Serie erscheinen;
- (3) nur einschlägige Beiträge, die sich ausschließlich mit Lexikographie oder mit fachverwandten Gebieten befassen, für Aufnahme in der AFRILEX-Serie in Betracht gezogen;
- (4) Beiträge anonym von einem aus Spezialisten des Faches von hohem akademischen Ansehen bestehenden Ausschuss beurteilt.

Lexikos erscheint jährlich. Ausgewählte monographische Studien dagegen erscheinen gelegentlich als getrennte Publikationen in der AFRILEX-Serie.

Politique éditoriale

La revue *Lexikos*, destinée aux spécialistes de lexicographie, est publiée dans la collection AFRILEX (acronyme de "lexicographie en Afrique et pour l'Afrique"). Depuis son sixième numéro, *Lexikos* est l'organe officiel de l'*African Association for Lexicography* (AFRILEX), entre autres parce que le Bureau du WAT s'est donné pour objectif de promouvoir le développement d'une telle association lexicographique en Afrique par la publication de la collection AFRILEX.

Les objectifs de la collection AFRILEX sont de :

- (1) créer un forum de discussion national et international sur la lexicographie, particulièrement au service de la lexicographie en Afrique, qui représente une grande diversité de langues;
- (2) stimuler le débat entre lexicographes, ainsi qu'entre lexicographes et linguistes;
- (3) établir et promouvoir le contact avec des projets lexicographiques locaux ou étrangers;
- (4) attirer l'attention générale sur la nature interdisciplinaire de la lexicographie, qui touche des domaines comme la linguistique générale, la lexicologie, l'informatique, le management, etc.;
- (5) favoriser et coordonner la coopération dans tous les domaines de la lexicographie; et
- (6) promouvoir les orientations de l'*African Association for Lexicography* (AFRILEX).

Pour atteindre ces objectifs, la collection AFRILEX

- (1) publiera les contributions aux discussions sur la lexicographie dans la revue *Lexikos*, dans la collection AFRILEX;
- (2) publiera sous forme de publications séparées dans la collection AFRILEX des monographies et autres travaux dans le domaine de la lexicographie;
- (3) ne publiera dans la série AFRILEX que des travaux dans le domaine de la lexicographie, qu'ils traitent de lexicographie pure ou des rapports entre la lexicographie et d'autres disciplines voisines; et
- (4) soumettra de manière anonyme toutes les propositions à des experts hautement qualifiés, pour en garantir le niveau académique.

Lexikos est publié annuellement, mais les travaux de qualité exceptionnelle seront publiés sous forme de publications séparées dans la collection AFRILEX.

Linguistic Theory in the Practical Lexicography of the African Languages

Emmanuel Chabata, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe; and Siswati National Lexicography Unit, Tshwane University of Technology, Nelspruit Campus, Nelspruit, Republic of South Africa (echabata@arts.uz.ac.zw)*

Abstract: In this article, we look at the relationship between linguistics and lexicography. We specifically look at the relevance of data derived from theoretical linguistic investigations to the compilation of dictionaries in African languages. Our point of departure is that since it is language description that lies at the core of both lexicography and linguistic theory, lexicographers can improve their work by using insights from theoretically-guided linguistic investigations. Our view is that as long as lexicographers focus on words and their existence in the linguistic system, they cannot work effectively without referring to linguistic theory, consciously or unconsciously. Lexicography is not only concerned with dictionary creation, that is, with the collection of lexical units and their proper description in dictionary entries, but also with the theoretical aspects concerning the lexicon. It is necessary for dictionaries to capture all lexical interrelationships of a phonetic, morphological, syntactic or semantic nature. Drawing examples from a few dictionaries on African languages, we try to show how dictionary compilers have benefited from specific theoretical investigations in general linguistics. We look at how the different linguistic theories have contributed to the improvement in the quality of the contents of some dictionaries of African languages. Our conclusion is that there is a stronger bond between linguistic theory and lexicographic practice than is generally assumed. Ways must therefore be found to understand the various links between the two disciplines. There should be a deliberate move from mutual neglect to collaboration between the two disciplines.

Keywords: AFRICAN LANGUAGES, DICTIONARY, GRAMMATICAL CATEGORISATION, HEADWORD SELECTION, LEXICAL MEANING, LEXICOGRAPHER, LEXICOGRAPHY, LINGUIST, LINGUISTICS, LINGUISTIC THEORY, PRACTICAL LEXICOGRAPHY, TONE MARKING

Opsomming: **Linguistiese teorie in die praktiese leksikografie van die Afrikatale.** In hierdie artikel kyk ons na die verhouding tussen die linguistiek en die leksikografie. Ons kyk veral na die tersaaklikheid van gegewens ontleen aan teoretiese linguistiese ondersoeke vir die samestelling van woordeboeke in die Afrikatale. Ons uitgangspunt is dat, aangesien dit taalbeskrywing is wat die kern van sowel die leksikografie as die linguistiese teorie vorm, leksikograwe hulle werk kan verbeter deur die insigte van teoreties-geleide linguistiese ondersoeke te gebruik. Ons siening is dat solank leksikograwe op woorde en hul bestaan in die linguistiese stelsel fokus, hulle nie doeltreffend kan werk sonder om bewustelik of onbewustelik na linguistiese teorie

te verwys nie. Dit is noodsaaklik vir woordeboeke om alle onderlinge leksikale verhoudinge van 'n fonetiese, morfologiese, sintaktiese en semantiese aard te ondervang. Deur voorbeelde uit 'n paar woordeboeke van Afrikatale te neem, probeer ons toon hoe woordeboeksamestellers voordeel getrek het uit spesifieke teoretiese ondersoeke in die algemene linguistiek. Ons kyk hoe die verskillende linguistiese teorieë bygedra het tot die verbetering van die gehalte van die inhoud van sommige woordeboeke van Afrikatale. Ons gevolgtrekking is dat daar 'n sterker verband tussen linguistiese teorie en leksikografiese praktyk is as wat algemeen aanvaar word. Maniere moet daarom gevind word om die verskillende skakels tussen die twee dissiplines te verstaan. Daar behoort 'n doelbewuste beweging te wees van wedersydse verwaarlosing tot samewerking tussen die twee dissiplines.

Sleutelwoorde: AFRIKATALE, GRAMMATIESE KATEGORISERING, LEKSIKALE BETENENIS, LEKSIKOGRAAF, LEKSIKOGRAFIE, LINGUIS, LINGUISTIEK, LINGUISTIESE TEORIE, PRAKTIESE TEORIE, TOONMARKERING, TREFWOORDKEUSE, WOORDEBOEK

1. Introduction

The debate on the relationship between lexicography and linguistics, or more specifically, on the relevance or usefulness of theory-based linguistic descriptions to dictionary making, is not new. In fact, the debate could be as old as the establishment of lexicography as a discipline. On the one hand, there is scholarship in favour of the separation of theoretical linguistics and lexicography. This scholarship does not recognise lexicography as a branch of linguistics. Because of this, the view is that there should be no link between theoretical linguistics and dictionary making. The linguists in this category considered lexicographers as non-linguists. In linguistic circles, dictionaries were therefore regarded as lacking linguistic interest since they were said to be produced without the help of any linguistic theory. The dictionary was seen as too unscientific to be worthy of any serious academic interest. It may be because of such scholarship that lexicography has sometimes even been referred to as an art (see, for example, Landau 2001), and sometimes as a descriptive venture that should avoid any link with theory whatsoever (see, for example, Geeraerts 1987). In defence of their field of study, lexicographers did not see the relevance of linguistic research in the compilation of dictionaries. They did not recognise how theoretical linguists could contribute to the practical work of dictionary making. As noted by Béjoint (2000: 170), practical lexicographers felt that academics would be of little use in lexicographic work with all its practical and social constraints to which linguistic theory was said to be ill-adapted. The same is also echoed in Urdang (1963: 594) who warns that although more theoreticians will be a welcome addition to the field of lexicography, they must remember that their theories should be interpretable above all in terms of practicality.

It is during the second half of the twentieth century that the interest of linguists in lexicography began to be noticed. Lexicographers also started to see

linguistic theory as an important tool that they could use in improving their dictionaries. The new philosophy on both sides was that the two fields are interrelated, making it possible to successfully contribute to and supplement each other. This scholarship was concretised by the publication of Zgusta's 1971 *Manual of Lexicography*, which heralded a new approach towards lexicography, linking lexicography to linguistic theories (and we should note in passing that this work of Zgusta has had much influence on African lexicography). By devoting a great deal of space to issues not primarily concerned with lexicography but rather with linguistics and linguistic theory (for example, focusing on topics such as lexical meaning, formal variation of words, and combinations of words), Zgusta gives a clear indication that linguistic influence does not only, or even primarily, run along the lines of formal grammar but that the dictionary needs to reflect real language usage and not only the language of the ideal speaker-hearer (Gouws and Prinsloo 2005: 2). Given the then scepticism prevalent among some linguists regarding the position of lexicography as a sub-domain of linguistics, Zgusta's work signals that a sound lexicographic practice utilises a sound linguistic analysis. Thus, since 1971, varying degrees of proximity have prevailed between lexicography and linguistics with different theories in linguistics having a bigger or lesser influence on lexicographic practice. Commenting on the current relationship between the two fields, Béjoint (1994: 177) notes that "lexicography and linguistics are now inextricably mixed" because precision and quality of work in lexicography often has to rely on evidence provided by linguistics. He (Béjoint 1994: 172) observes that, to date, the relations between linguists and lexicographers are strong, going in both directions. To illustrate his observation, Béjoint (2000) notes that linguists such as David Crystal and John Sinclair are involved in the compilation of dictionaries and lexicographers like Patrick Hanks are also linguists. Similarly, Hartmann (1983: 4) argues that in order to compile a dictionary, one must have a notion of the "word" and an understanding of how words are used in interpersonal discourse. He (Hartmann 1983: 4) further argues that the discipline that has contributed most to an understanding of the use of words in communication is linguistics, that is, much of the literature on lexicography has consciously and explicitly related itself to linguistic theories in general and to theories of lexical semantics in particular.

In this article, we try to show how lexicographers seek the intervention of linguistic theory in many aspects of their work. We argue that dictionary compilers consciously or unconsciously depend on previous linguistic studies, for example, when deciding on what kinds of semantic information to present in a dictionary (see Hartmann 1983 and Svensén 1993 among others). Our position is that lexicography could be made more systematic, consistent and comprehensive with the help of a rigorous use of data derived from theoretical linguistic analyses. In line with this, we propose a stronger theoretical linguistic basis to enhance the linguistic authority of dictionaries. In fact, we will note that, knowingly or unknowingly, lexicographers in some African languages

have been influenced to a great extent by insights derived from linguistic theory. This observation is supported by the fact that most of the current lexicographers working with African languages are linguists by training, and that they are even attached to departments of linguistics or African languages at different African universities. Because of this, we believe that linguistic theory has provided part of the much needed training for doing good lexicographic work. We therefore agree with Yong and Peng (2007: 41) who argue that the link between lexicography and theoretical linguistics explains why both practical and theoretical lexicography have undergone significant advances in the past three decades, a period during which there was a marked development in theoretical linguistics.

Our belief is also that a theory, in general, both serves as a guide to discovery and as a tool bringing insights in a revealing way. It helps us to deal with both the known and the unknown. It is in this light that Hyman (2006: 13) argues that with linguistic theories, we obtain a better description and understanding of language issues in question. He (Hyman 2006: 13) further holds the opinion that linguistic theories might lead researchers to ask questions they may not otherwise have raised about language structure and to seek data they may not otherwise have considered. In addition, theoretical awareness helps researchers see connections that might otherwise not be made, as well as pinpointing problems that might have been overlooked. Experience in compiling dictionaries in African languages has shown that by using data from theoretically guided investigations into linguistic structure, lexicographers increase their chances of producing a more accurate and comprehensive language description. It is actually because of the great influence that theoretical investigations into linguistic structure have had on African lexicography that Gouws and Prinsloo (2005: 2) conclude that the advent of modern lexicography is positioned within the broader framework of linguistics. It is also because of this that they go a step further and propose that lexicographers should always take cognisance of developments in linguistic theory and that data presented in dictionaries should result from a sound linguistic analysis.

For us to appreciate the importance or relevance of data derived from theoretical linguistics for lexicography, we have to accept that language is at the centre of both theoretical linguistics and lexicography. Good (1988: 81) notes that both practical dictionary writers and linguists are involved, strictly speaking, in the same task: that of providing a description of the lexical stock of a language, or alternatively, of modelling the linguistic knowledge of native speakers. For example, in lexicography the most common tasks include the provision of meaning as well as the description of grammatical behaviour and pronunciation of words, their stylistic characteristics and other aspects of their usage, all of which are at the core of the development and application of linguistic theories. In both disciplines and/or practices, the goal is to describe language. The only difference may be that whilst in theoretical linguistics each theoretical approach is a description of language from some narrow and spe-

cific perspective, in practical lexicography the goal is to try and describe a language or languages in their totality — the aim in practical lexicography being to fully describe linguistic concepts. If this assumption is accepted, therefore, we would proceed to argue that it is only data from various perspectives or theoretical investigations that can or should be corroborated to help in the provision of the much needed basis for fuller descriptions of lexical items in dictionaries. The processes of compiling, editing and extending dictionaries as well as research on dictionaries should thus profit from progress and developments in linguistics and its various branches, that is, among others, sociolinguistics, psycholinguistics, pragmatics and semantics. We have already argued above that since lexicographers specifically deal with words, they should rely on advances in linguistic theory for understanding what a word is before they can be able to provide words with definitions, pronunciation and other details. They should, for example, have a fuller understanding of whether a word is a sequence of sounds (that is, a phonological word), a sequence of letters (a graphic word), a formal unit composed of several components (a morphological word) or a unit expressing meaning (semantic word). Given the interdisciplinary nature of dictionaries, lexicographers should not only follow developments in linguistic theory but should also benefit from many other fields such as terminology, information technology, language teaching, translation, psychology, philosophy, history and science.

However, whilst we acknowledge the importance of linguistic theory for dictionary compilation, we should not forget that lexicographers are doing scientific work but publishing it for users whose pursuits are always more practical, and who are usually not trained linguists or lexicographers. In the light of this, therefore, practical lexicographers should negotiate their way well between linguistic data derived from serious linguistic theories and the parallel needs of dictionary users. In a way, our considered position is that although linguistic theories are useful to practical lexicography by providing relevant and precise data about linguistic concepts, lexicographers should not be obsessed with linguistic theories; that is the field for linguists. Instead, these theories should only be used to enhance the efforts of practical lexicographers in their task of dictionary making, a task with its own specific needs and goals. Lexicographers should not be led by the demands of different linguistic theories to the extent that the principles or goals of lexicography as a discipline are compromised. For example, they should not fail lexicographic user-perspective ideals such as user-friendliness and historical and cultural sensibilities of the target audience. The reason for this is aptly captured by Gouws and Prinsloo (2005) who describe a dictionary as "the display-window of linguistics". They note that the people who look at these display-windows are not trained linguists but rather average members of the relevant speech community. Similarly, Zgusta (1971) argues that lexicographers need to be familiar with linguistics in a much broader sense, taking into consideration not only the whole structure of the language in question but also the culture of the respective lin-

guistic community. Thus, our position is that lexicographers should not aim at producing linguistically sophisticated products, but that their linguistic sophistication should help them produce informative and useful products for their intended users.

Having declared our position regarding the importance of data derived from theory-based linguistic investigations as well as other lexicographic aspects to which lexicographers have to pay attention, we will try to show how some linguistic theories might be taken into account to the benefit of lexicographic practice. In our presentation, we will consider different parts of a typical monolingual dictionary entry, and then try to show how the treatment of the respective parts has benefited from a relevant linguistic theory or theories. To appreciate the different aspects on which we will focus, let us consider the way some lexicographic data has been presented in the following dictionary entry from Chimhundu (1996), a general monolingual Shona dictionary.

- (1) **muswe** [musve] DK z 3. 1 *Muswe* inhengo yemuviri wemhuka inomera kumagumo emusana, kana kurembera, inopinimidza kumashure. (A *tail* is a part of an animal's body found hanging at the end of its back.) FAN *muchira* 3. 2 *Muswe* munhu anosaririra shure pakuita zvinhu. (*Muswe* is a person who lags behind in doing something.)

In this example, the word given in bold, **muswe** (tail), is the headword, that is, the word being described, and in square brackets is shown its variant form. Following this variant, the letters D and K represent tone marking. Whilst D is an abbreviation for **-dzikisa** (low tone), K is an abbreviation for **-kwidza** (high tone). Thus, **muswe** is pronounced as *low*, *high*. On the tone marking follows the abbreviation z, the short form for **zita** (noun) indicating that **muswe** belongs to the grammatical category of nouns. This is also followed by a number 3, which shows the class to which the noun belongs in the Shona noun class system. If it is a verb, for example, tone marking would be followed by abbreviations indicating whether the verb is transitive or intransitive. After the information on grammatical categorisation, the two meanings of the word numbered as 1 and 2 follow. At the end of meaning 1, FAN *muchira* 3 is found, where FAN is an abbreviation for **fanana**, loosely meaning "same as", and implying that **muchira** is synonymous with **muswe** and also belongs to class 3.

Now, also drawing our examples mainly from Shona, we will try to show how the lexicographer makes his/her decisions on each of these aspects, and how linguistic theory can help in the decision-making process. Aspects that will be focused on are headword selection, tone marking, grammatical categorisation and meaning.

2. Headword selection

It has become common knowledge in lexicographic practice that a dictionary cannot contain all the words and phrases of a language, no matter its size or

scope. Because of this, headword selection forms an important stage in the compilation of any dictionary, for it determines the major contents of a particular dictionary. As noted by Landau (1984: 185), the decision to include or exclude a particular category of words relates directly to the purpose of the dictionary. Whatever the lexicographer eventually includes in or omits from the dictionary, must be the result of convincing reasons; the selection principles should be well-considered. The important question then is, "What are the lexicographer's selection criteria and where does linguistic theory come in?" To illustrate lexicographers' usual dilemma, let us consider the following sets of Shona derivations.

- (2) (a) **-famba** (walk) *vs* **mufambi** ((some)one who is walking)
 (b) **-tenga** (buy) *vs* **mutengesi** (literally: (some)one who is selling; figuratively: (some)one who is selling out)
 (c) **-bika** (cook) *vs* **-bikwa** (be cooked) *vs* **-bikisa** (literally: cause someone to cook; figuratively: acquire/come into the possession of an own household as a newly married woman)

Derivation is commonly referred to as a process of word formation by which new lexical items are created (see, for example, Bolinger 1968 and Katamba 1993). If such a definition of derivation is maintained, then both **mufambi** and **mutengesi** are forms of different lexemes (derived from the verbs **-famba** and **-tenga** respectively) and should both be listed in the dictionary as distinct entries, each with its own definition. However, a look at Chimhundu's (2001) Shona dictionary, for example, shows that whilst the noun **mufambi** was excluded, **mutengesi** was included. The question one would ask is "Why?", since the two look quite similar in form and in their manner of derivation; both are nouns derived from verb bases by a similar process of nominalisation. We can only infer that **mufambi** was omitted because its meaning can easily be traced from that of the verb **-famba**, from which it is derived. On the other hand, **mutengesi** was included because, in addition to the easily predictable sense of "(some)one who is selling", the process of nominalisation has also introduced a figurative sense, which cannot be deduced from the way the noun is derived from the verb. The same can be said of the extended verbs **-bikwa** and **-bikisa** where **-bikwa** is excluded from the dictionary and **-bikisa** is included. As we will argue later when we look at lexical meaning, such reasoning and decisions are informed by insights from theoretical research in cognitive linguistics dealing with derivations of various kinds.

3. Tone Marking

The correct pronunciation of words is one reason why the speakers of a language consult dictionaries. To show the importance of tone marking, let us look at the following entries from Chimhundu (1996).

- (3) (a) **guru** K- z 5. *Guru* imhuri yomurume mumwe chete, inenge iine vakadzi vakawanda navana vazhinjiwo. (*Guru* is a family owned by one man, who has many wives and many children.)
- (b) **guru** D- z 5. Iri zimwena rinoita pachuru, rakakura, munogara nyoka kana zvimwewo zvikara zvesango. (This is a big hole usually on an anthill, where snakes and other dangerous wild animals can reside.)
- (c) **guru** DK z 5. *Guru* inhengo yedumbu remhuka, munogaiwa chikafu chose chinenge chadyiwa, rakafanana nematauro okugezesa mukati maro. (*Guru* is a part of an animal's stomach, where food is digested and whose inside looks like a bathing towel.)

Although these three words are spelt the same, they refer to different things solely because they are pronounced differently. For this reason all of them are listed in the dictionary as different entries.

Tone marking is a historical problem in African lexicography. In fact, the general history of phonological studies has undergone many changes, with phonologists proposing different ways of the phonological representation of sounds in dictionaries. Because of the confusion surrounding this area of lexicographic description, in most traditional dictionaries of African languages tone marking occurred in a haphazard manner, with most dictionary compilers not even knowing where and how this should be done. Even in some recent dictionaries (see, for example, TUKI 2000, Hadebe 2001, Nkomo and Moyo 2006, and Mbatha 2006), tone marking was actually omitted for fear of mishandling it. However, the insights of phonological theories, particularly Goldsmith's Autosegmental Phonological Theory, have shown that unlike languages having an abundance of contour tones, in African tone languages the primary opposition is between level tones. It is research in phonological theory by Goldsmith and others that has established that many African languages have two level tones only, *high* and *low*, which is marked at syllable level. In fact, it is research on phonological theory that made it possible for tone to be marked with some measure of consistency in some dictionaries of African languages (see, for example, Dale 1981 and Chimhundu 1996, 2001, among others, where tone is actually contrasted and marked as either *high* or *low* on each syllable). It is also research in phonological theory that has shown that words that differ only in tone are as distinct from each other as words with different vowels, hence the need to treat them as distinct in dictionaries. Furthermore it is phonological theory that has provided the notation allowing African lexicographers to mark tone in their dictionaries.

4. Grammatical Categorisation

We have already intimated in the introductory section that one of the lexicographer's chief tasks is to deal with grammatical categorisation and labelling of lexical items, that is, placing words or parts of words into word classes tradi-

tionally identified, among others, as verb, noun, pronoun, ideophone or adjective. In other words, the lexicographer has to show the characteristic syntactic functions of words selected for inclusion in the dictionary. For example, if a word is a noun, information about its class membership as well as its singular/plural categorisation would be provided. If it is a verb, information about its transitivity would be presented. However, the provision of grammatical information about words is sometimes a big challenge to lexicographers, especially those without proper linguistic training. Experience in the process of compiling dictionaries has shown that word class categorisation is difficult to state explicitly in some cases. Faced with a situation where they could not classify some words, the editors of Chimhundu (1996) eventually created a **kanu** (interjective) category to which all the words causing categorisation problems were assigned. They agreed that further research was needed on the relevant words. However, a look at a theory like Lexical Functional Grammar (LFG) shows that linguistic theory has much to offer in trying to solve problems similar to this one. Since it deals with the lexical representation of words, LFG helps in shedding light on such issues as word class categorisation as well as the morphosyntactic characteristics of lexical items. To illustrate the different kinds of information to which we are referring here, let us look at the LFG representation of the following Shona sentence.

- (4) John akanwa mvura
 John \emptyset -aka-nw-a mvura
 John CL1-RM.PST.PRF-drink-FV water
 NP SC-TAM-Root-TAMP NP
 'John drank water'.

In this sentence, there are three lexical entries, that is, (a) **John**, (b) **akanwa** and (c) **mvura**, which can be analysed using LFG principles as follows:

John , N	(PRED) = John (NUM) = SING (PERSON) = CL1
akanwa , V	(PRED) = -nwa (\uparrow SUBJ)(\uparrow OBJ) [(SUBJ NUM) = SING & (SUBJ PERS) = 3] (TENSE) = PAST
mvura , N	(PRED) = mvura (NUM) = SING (PERSON) = CL5

This analysis shows that much useful lexicographic information is revealed. For example, we notice from its representation that **akanwa** belongs to the verbal category. We also notice that the verb, **-nwa**, requires two semantic arguments which are associated with two grammatical functions, that is, 'subject' and 'object', making it a transitive verb. Transitivity, as we have already suggested, is one aspect that is usually indicated for all verbs in most dictionaries.

Let us consider how this verb is treated in Chimhundu (1996).

- (5) -**nwa** κ it. *Kunwa* kumedza chinhu chakaita sedoro, mvura kana tii. (Kunwa is to swallow something like beer, water or tea.)

In the above example, *it.*, an abbreviation of *itika*, shows that -**nwa** is a transitive verb that takes a subject and an object in a sentence. Therefore, the theoretical principles that guide lexicalist frameworks such as LFG and the analyses that result from the application of these principles should be very useful in providing relevant lexicographic information.

5. Lexical Meaning

Although it has become the norm in modern lexicography also to include information about, for example, the spelling, pronunciation and etymology of words in dictionaries, Béjoint (2000: 6) notes that the main objective of dictionaries is to define words and terms. In other words, the most central aspect is word meaning, especially in monolingual dictionaries. In his description of monolingual dictionaries, Landau (2001: 8-9) says:

A monolingual dictionary [...] provides many kinds of information about its entry words but most importantly gives definitions [...]. The chief purpose of a monolingual dictionary is to explain, in words likely to be understood, what other words mean.

The importance attached to word meaning in monolingual dictionaries means that the defining task should be taken seriously. However, as again noted in Zgusta (1971), Hartmann (1983), Good (1988), Jackson (1988), Svensén (1993) and Landau (2001), among others, the description of word meaning is one of the most difficult tasks lexicographers must perform. In providing the designative and connotative meanings of words, lexicographers are usually faced with many difficult choices. For example, they have to answer the following basic questions, which should help them to create definitions that are precise, comprehensive, exhaustive and all-inclusive and that are also given in both a linguistic and cultural context: (a) Should the focus be on the senses of individual words?, (b) Which readings of a word should be considered relevant?, (c) Which type(s) of meaning should be defined?, (d) Which linguistic perspective(s) should be taken?, and (e) Which defining format(s) should be used? In answering these questions, lexicographers have many other variables to consider. For example, they have to think about their audience and the purpose of the dictionary. In handling all these, lexicographers cannot solely depend on their intuition; instead, they have to rely on research results from other fields. Good (1988) argues that in order to understand this complex situation and to make informed decisions and choices, but at the same time to avoid being too casual in the way they display the various senses they attribute to a word, lexi-

cographers can benefit from insights from the linguistic or philosophical debate on the semantic structure of words. Linguistic theories regarding semantics or meaning as the centre of grammatical description such as those of the conceptualist or cognitive tradition[†] (for example, Cognitive Grammar, Construction Grammar and Frame Semantics) can be exploited in supporting lexicographic decisions. Central to the conceptualist approach to meaning is the assumption that meaning is a cognitive phenomenon and thus it should be analysed as such. Following this assumption, the description of lexical meaning should, therefore, be understood with reference to a structured background of experience, beliefs or practices constituting a kind of conceptual prerequisite for understanding meaning (Fillmore and Atkins 1992: 76). Linked to this is also the hypothesis that knowledge of language emerges from language use, that is, that semantic structure is built up from cognition of specific utterances on specific occasions of use. Viewed this way, the assumption in the conceptualist approach is that word concepts may not be understood apart from the social and cultural institutions in which the action, state or thing is situated. Instead, the view is that to fully capture the meanings of words, both the language and the socio-cultural context in which they are used has to be understood. Thus, obtaining insights from the principles guiding the treatment of meaning in Cognitive Grammar, for example, lexicographers should be able to create definitions fulfilling one aspect of the meanings desired in dictionaries, especially monolingual dictionaries: socio-cultural relevance. Because meaning is too broad a concept, let us look at a few challenging aspects of it, and try to show how linguistic theory has helped or should help lexicographers.

Let us start with the treatment of polysemy, one of the lasting challenges in the provision of meaning in monolingual dictionaries. Polysemy is commonly described as a situation where a word is associated with two or more distinct but related senses. Faced with a word with multiple senses, lexicographers' first challenge is to decide which meaning or meanings to include in or exclude from the dictionary. Depending on the intended audience and the purpose of the dictionary, lexicographers may focus on the most common sense(s) only, or may include less common ones as well. They may restrict their defining efforts to general senses or may include marked readings of the words as well. Having selected the senses to include in the dictionary, lexicographers have to face the next challenge, that of sense ordering in cases where more than one sense has been selected. Generally speaking, haphazard and ad hoc decisions regarding sense ordering are found in most dictionaries of African languages. However, as noted in Inglis (2003), the principles guiding the Prototype Theory can contribute much to solving problems linked to polysemy by ensuring consistency in dealing with issues such as sense ordering. The Prototype Theory makes a crucial distinction between central and peripheral senses of a word and provides guidelines for determining the core-periphery distinction. The study of verbs using the Prototype Theory (see, for example, Chabata 2007) has shown that in each polysemous situation, there is a fairly specific

central sense of a word, which should be regarded as its prototypical or core sense. The other separate but related sense(s) become(s) the word's less typical sense(s). Whilst prototypical senses are classified as those that designate aspects basic to human experience, hence typical of the way we use language to express our conceptualisation of the physical world, the peripheral senses are those whose interpretation depends on similar or related elements of the more basic senses. Such principles of classifying word meanings, if followed, make it easier for the lexicographer in determining the core sense(s) of words, which should logically be listed first and also the peripheral sense(s) that should be listed later or last. To illustrate what is meant, let us compare the ordering of the meanings provided for the verb **-minyura** in Chimhundu 1996 and Chimhundu 2001, respectively:

- (6) (a) **-minyura** D it. 1 *Kuminyura* kutaura zvinhu zvisinganzwikwi, kana zvine svoto kune vakateerera. (*Kuminyura* is to talk in an inaudible manner, or in a way that provokes those who are listening.) 2 *Kuminyura* magodo kuapesanisa kana kuasvodogora panzvimbo paanosangana, sezvakaita pagokora kana pachitsitsinho. (*Kuminyura* is to dislocate bones at the places where they meet, such as at the elbow or the toes.)
- (b) **-minyura** D it. 1 *Kuminyura* magodo kuapesanisa kana kuasvodogora panzvimbo paanosangana, sezvakaita pagokora kana pachitsitsinho. 2 *Kuminyura* kutaura zvinhu zvisinganzwikwi, kana zvine svoto kune vakateerera.

A look at this entry shows that it is provided with exactly the same definitions in both dictionaries. The only difference lies in the ordering of the senses, which is reversed. The obvious question one is tempted to ask is "Why?" In trying to understand the reason for this, we should note that in the example from the 2001 dictionary, the sense ordering follows what we have noted as the principles guiding the ordering of multiple senses using the Prototype Theory. One could conclude that the editors of the later dictionary were influenced by insights from the Prototype Theory to meaning, hence the decision to re-order the definitions.

Another challenge concerns the meanings of complex structures such as complex nominal constructions and extended verbs that we have already touched upon in our discussion of headword selection. A look at some dictionaries published in certain African languages (see, for example, Moreno 1988 for Nambya; Chimhundu 1996, 2001 for Shona; Hadebe 2001 for Ndebele; and Mbatha 2006 for Zulu) shows that the meanings of complex structures are generally treated as predictable from their respective constructional patterns. In fact, it is because of this general treatment that most complex structures are excluded from these dictionaries. The assumption is that it should be fairly easy for dictionary users who want to know the meanings of complex structures to simply add up the meanings of their component parts. With regard to extended verbs, for example, the assumption is that it is easy to deduce the meanings of these verbs by the addition of the meanings of parts of the ex-

tended form, that is, the meanings of the verb base and the verbal extension. This explains why the majority of extended verbs and their senses were excluded from the above-mentioned dictionaries. However, the study of the causatively extended verbs in Nambya (see, for example, Chabata 2007), and the application of the principles guiding Cognitive Grammar in its treatment of the meanings of complex structures, have shown that verbal extensions add meanings to verb bases, which are more specialised than is usually assumed, and some of these meanings are not analysable from the sum total of the meanings of parts of the extended verb. More often than not, the meanings of extended verb forms cease to be compositional as is implied in the various treatments that extended verbs have received in most existing dictionaries of African languages. Instead, it was discovered that extended verbs normally have both compositional and non-compositional meanings. It was further discovered that the non-compositional meanings are a result of semantic extension through processes such as metaphor, metonymy and specialisation. Such theoretical insights should be useful when dealing with headword and sense selection; that is, if the meaning of an extended verb is predictable from its constructional pattern, then it should be excluded from the dictionary, and if the meaning is unpredictable or idiosyncratic, then it should be included in the dictionary. To illustrate this, let us take an example of one Nambya extended verb, *-lyisa*, which is derived by the addition of the causative extension *-is-* to the verb root, *-ly-* (eat). The sum total of the meanings of *-ly-* and *-is-* would give us the meaning, 'cause to eat; feed'. However, *-lyisa* has two other senses, that is, (a) 'brainwash' and (b) 'poison', which are not analysable from the meanings of parts of the verb. Whilst meaning (a) is a metaphorical extension from the basic sense of causing someone to eat something, meaning (b) is a form of specialisation. From a Cognitive Grammatical point of view, therefore, the creation or derivation of complex structures such as extended verbs is not a mere mathematical process of addition but a process yielding both predictable and unpredictable meanings. The principles of Cognitive Grammar are thus informative in understanding the different kinds of meaning of complex structures. Insights from Cognitive Grammar regarding the predictable/unpredictable distinction, therefore, should help lexicographers to decide which senses to include in or omit from a dictionary. In this case, predictable senses can be omitted, for they are mathematically derivable, but unpredictable senses may need to be included since they may not be recoverable by any other means.

6. Conclusion

In this article, we have tried to show that theoretically guided investigations into linguistic structure are useful to the lexicographer of African languages. Taking our examples mostly from Shona, we have also tried to show that every aspect of a dictionary entry benefits in some way from research in theoretical

linguistics. This does not however mean that African lexicographers, whom we have also identified as trained linguists, ask themselves continually when constructing entries which linguistic theory would be useful for handling a specific aspect. Instead, linguistic theory is part and parcel of their lexicographic equipment. All in all, we maintain that knowingly or unknowingly, consciously or unconsciously, lexicographers, not only of African languages but of all languages, benefit from insights derived from theoretical linguistic research. In fact, we believe that data from linguistic theories is relevant to dictionary making in any language, African or otherwise. Although this article mainly focused on the relevance of linguistic data to lexicography, we should note that the links between the two fields of study are often mutual. For example, as we have already tried to show, whilst lexicographers depend on data from linguistic research in almost all the stages of compiling dictionaries, linguists in turn utilise lexicographic information in answering many linguistic questions, for example checking such factors as, among many others, the evidence of the meaning range and the origin and semantic change of lexical items.

Note

- † For a fuller discussion of how meaning is viewed from a conceptualist or cognitive linguistic approach, refer to Fillmore (1982), Langacker (1987, 2000), Svorou (1994), Goldberg (1995, 2006), Ungerer and Schmid (1996), Sweetser (1999), Kövecses (2002) and Taylor (2002, 2003), among many others.

References

- Béjoint, H.** 1994. *Tradition and Innovation in Modern English Dictionaries*. Oxford: Clarendon Press.
- Béjoint, H.** 2000. *Modern Lexicography: An Introduction*. New York: Oxford University Press.
- Bolinger, D.L.** 1968. Entailment and the Meaning of Structures. *Glossa* 2: 119-127.
- Chabata, E.** 2007. *The Nambya Verb with Special Emphasis on the Causative*. Unpublished Ph.D. Thesis. Oslo: University of Oslo, Department of Linguistics and Scandinavian Studies.
- Chimhundu, H. (Ed.)**. 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. (Ed.)**. 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Dale, D.** 1981. *Duramazwi: A Shona-English Dictionary*. Gweru: Mambo Press.
- Fillmore, C.J.** 1982. Frame Semantics. The Linguistic Society of Korea (Ed.). 1982. *Linguistics in the Morning Calm*: 111-157. Seoul: Hanshin.
- Fillmore, C.J. and B.T. Atkins.** 1992. Toward a Frame-based Lexicon: The Semantics of RISK and its Neighbors. Lehrer, A. and E.F. Kittay (Eds.). 1992. *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organisation*: 75-102. London/Hillsdale: Lawrence Erlbaum Associates.
- Geeraerts, D.** 1987. Types of Semantic Information in Dictionaries. Ilson, R. (Ed.). 1987. *A Spectrum of Lexicography: Papers from AILA, Brussels 1984*: 1-10. Amsterdam: John Benjamins.
- Goldberg, A.E.** 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago/London: The University of Chicago Press.

- Goldberg, A.E.** 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Goldsmith, J.A.** 1976. *Autosegmental Phonology*. Unpublished Ph.D. Thesis. Boston: Massachusetts Institute of Technology, Thesis Department of Foreign Literatures and Linguistics.
- Good, C.** 1988. Lexicography and Linguistic Theory; with Special Reference to German. *New German Studies* 15(2): 81-110.
- Gouws, R.H. and D.J. Prinsloo.** 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PReSS.
- Hadebe, S. et al. (Eds.)**. 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hartmann, R.R.K.** 1983. On Theory and Practice: Theory and Practice in Dictionary-making. Hartmann, R.R.K. (Ed.). 1983. *Lexicography: Principles and Practice*: 3-11. London: Academic Press.
- Hyman, L.M.** 2006. *African Languages and Phonological Theory*. Berkeley: University of California. http://www.linguistics.berkeley.edu/~hyman/GLOT_Phonology_African_Lgs.pdf
- Inglis, D.** 2003. *Cognitive Grammar and Lexicography*. Payap University Graduate School.
- Jackson, H.** 1988. *Words and their Meaning*. London/New York: Longman.
- Katamba, F.** 1993. *Morphology*. London: The Macmillan Press.
- Kövecses, Z.** 2002. *Metaphor: A Practical Introduction*. Oxford/New York: Oxford University Press.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Langacker, R.W.** 1987. *Foundations of Cognitive Grammar. Volume 1: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R.W.** 2000. *Grammar and Conceptualization*. Berlin/New York: Mouton de Gruyter.
- Mbatha, M.O. (Ed.)**. 2006. *Isichazamazwi sesiZulu*. Pietermaritzburg: New Dawn Publishers.
- Moreno, A.** 1988. *Nambya Dictionary*. Gweru: Mambo Press.
- Nkomo, D. and N. Moyo (Eds.)**. 2006. *Isichazamazwi SezoMculo*. Gweru: Mambo Press.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford/New York: Oxford University Press.
- Svorou, S.** 1994. *The Grammar of Space*. Amsterdam/Philadelphia: John Benjamins.
- Sweetser, E.** 1999. Compositionality and Blending: Semantic Composition in a Cognitive Realistic Framework. Janssen, T. and G. Redeker (Eds.). 1999. *Cognitive Linguistics: Foundations, Scope, and Methodology*: 129-162. Berlin: Mouton de Gruyter.
- Taylor, J.R.** 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Taylor, J.R.** 2003. *Linguistic Categorization*. Third Edition. Oxford: Oxford University Press.
- TUKI.** 2000. *English-Swahili Dictionary*. Second Edition. Dar es Salaam: Institute of KiSwahili Research.
- Ungerer, F. and H. Schmid.** 1996. *An Introduction to Cognitive Linguistics*. London/New York: Longman.
- Urdang, L.** 1963. Review of Problems in Lexicography. *Language* 39: 586-594.
- Yong, H. and J. Peng.** 2007. *Bilingual Lexicography from a Communicative Perspective*. Amsterdam/Philadelphia: John Benjamins.
- Zgusta, L.** 1971. *Manual of Lexicography*. The Hague: Mouton.

Users, User-Friendliness and Projected Uses of *Isichazamazwi SesiNdebele*: An Analysis*

Langa Khumalo, *Department of Linguistics, University of KwaZulu-Natal, Durban, Republic of South Africa (la_nga@yahoo.co.uk)*

Abstract: This article discusses the first-ever monolingual general Ndebele dictionary, *Isichazamazwi SesiNdebele* (henceforth the ISN) within the context of the history of lexicography and the compilation of dictionaries in Ndebele. It further assesses the scope of the dictionary with regard to its structure. It also discusses decisions taken by the editors during the writing of the ISN in an attempt to compile a user-friendly dictionary primarily aimed at secondary schools and the general Ndebele population. The assumptions of the editors, coming from a non-existing lexicographic tradition, were mainly based on their experiences and intuitions as Ndebele mother-tongue speakers, educators and language practitioners. The article also accounts for four major uses of the dictionary.

Keywords: TARGET USER, USER-FRIENDLY, CORPUS, FRONT MATTER, BACK MATTER

Opsomming: **Gebruikers, gebruikersvriendelikheid en die beplande gebruike van *Isichazamazwi SesiNdebele*: 'n Ontleding.** Hierdie artikel bespreek die heel eerste eentalige algemene Ndebelewoordeboek *Isichazamazwi SesiNdebele* (voortaan die ISN) binne die konteks van die geskiedenis van die leksikografie en die samestelling van woordeboeke in Ndebele. Dit beoordeel verder die omvang van die woordeboek met betrekking tot sy struktuur. Dit bespreek ook besluite wat deur die redakteurs van die ISN geneem is tydens die skryf van die ISN in 'n poging om 'n gebruikersvriendelike woordeboek saam te stel wat primêr gerig is op sekondêre skole en die algemene Ndebelebevolking. Die aannames van die redakteurs wat van 'n niebestaande leksikografiese tradisie kom, is hoofsaaklik gebaseer op hul ervarings en intuïesies as moedertaalsprekers, opvoeders en taalpraktisyns. Die artikel verantwoord ook vier hoofgebruike van die woordeboek.

Sleutelwoorde: TEIKENGEBRUIKER, GEBRUIKERSVRIENDELIK, KORPUS, VOORWERK, AGTERWERK

Historical Background

Hadebe (2002) has given a detailed history of lexicographic work in the Ndebele language. From his account, it is clear that the Ndebele language has a

* This article is a revised and updated version of a paper presented at the Sixth International Conference of the Linguistic Association of SADC Universities, held at the University of Botswana, Gaborone, Botswana, 6–8 August 2001.

short lexicographic history. Until the publication of the ISN there was no monolingual Ndebele dictionary. There was only one bilingual English–Ndebele dictionary, *A Practical Ndebele Dictionary* by Pelling (1971), which has about 4 500 entries. However, one of the earliest Ndebele lexicographic works was compiled by a missionary of the London Missionary Society, W.A. Eliot. It appeared in 1897, with the title *Dictionary of the Tebele and Shona Languages*. Later he published what can be considered the first bilingual Ndebele–English dictionary which used the early Ndebele orthography. Changes in the Ndebele orthography would take place in the 1930s. Another early work of that period was a word list compiled by a British South Africa Police trooper, which was published in Cape Town in 1903 under the title of *Matebele and Makalaka Vocabulary: Intended for the Use of Prospectors and Farmers in Mashonaland* (see Chimhundu 1992: 18).

Pelling (1971) has been the only dictionary in the current Ndebele orthography until the publication of the ISN in 2001. There is not much information either in the front or the back matter of the dictionary to indicate the principles followed in the compilation of this bilingual dictionary. It has been observed by Hadebe (2002) that Pelling's dictionary mainly updates the orthography while its contents are chiefly a revision of Eliot's dictionary. While the ISN is based on different principles, Hadebe admits that the background set by Pelling's dictionary cannot be ignored, especially in so far as the standardization of the Ndebele language is concerned.

Hadebe (2002) further observes that the absence of the publication of dictionaries does not mean the absence of lexicographic work of one kind or another. Some dictionary projects were initiated but never finished while others were finished, but for various reasons, the manuscripts were never published. There are two cases in point, the dictionary project by S.J. Mhlabi, then Senior Language Assistant in the African Languages and Literature Department of the University of Zimbabwe, and that of G. Sibanda also from the University of Zimbabwe. Mhlabi's work is a bilingual Ndebele–English dictionary in the form of a complete manuscript, with its cards catalogued. This project was supervised by the then Professor of African Languages, George Fortune, who retired before the manuscript could be sent for publication. Sibanda's work was a privately funded project initiated by the Ndebele Language Committee. According to Hadebe's account, the completed work was sent to the publishers and, after having received some comments for further editing, the manuscript disappeared without being published.

The current lexicographic work on Ndebele by the African Languages Research Institute (ALRI), previously the ALLEX Project, at the University of Zimbabwe, Harare, is being documented through annual reports, so that it is easier to gather information on the progress of this work than it was for other previous projects. The ALLEX Project, the brainchild of Professor Herbert Chimhundu, was generously funded by NUFU (Norwegian Universities' Committee for Development Research and Education) through the development of a North–South cooperation. This project was aimed at producing monolingual

dictionaries and glossaries in the African languages of Zimbabwe. This was a large collaborative research project involving more than thirty academics, students, technical staff and other support staff at the University of Zimbabwe and computer and lexicography specialists from the University of Oslo in Norway and the University of Gothenburg in Sweden.

The research that led to the publication of the ISN resulted in the questioning of a number of tenets that had been taken for granted about the Ndebele language. For instance, it has been generally assumed that the Ndebele language does not have varieties (Chimhundu 1993: 59) or that the variation was insignificant. Yet, the corpus of oral material has revealed some variations in speech behaviour according to the geographical sources of this material (Hadebe 2002). It is, however, clear from the foregoing that Ndebele lexicography has a short history.

The General Scope of the ISN

The ISN is the first-ever monolingual dictionary in the Ndebele language. The title of the dictionary (*Isichazamazwi SesiNdebele*, literally, the explainer of words in Ndebele) already indicates that it is a dictionary of Ndebele in Ndebele, in which the resources of the language are used for the first time lexicographically to analyze and describe itself. As stated above, until the publication of the ISN, the Ndebele people were using Pelling's bilingual Ndebele–English dictionary.

The ISN is a medium-sized, general-purpose dictionary designed to be inexpensive and easy to handle so as to be accessible to secondary school pupils and the general Ndebele population. According to commonly accepted modern standards, it is well printed, with the utilization of different kinds of typographical conventions, which demonstrates the movement from headword to synonyms in the treatment of each major entry in the dictionary. The components of a typical major entry are: the headword, in large bold lower-case roman letters; the type of word in question (i.e. grammatical information), described by an abbreviated label in standard lower-case roman letters; a number, in the case of a noun, indicating the class to which it belongs; an explanation of the meaning or meanings in lower-case roman letters, with the headword, when referred to, printed in italics; an illustrative example or examples given in italics showing the headword in context; and finally, synonyms, printed in italics under the rubric FAN, an abbreviation for *amagama alengcazelo ezifanayo*, which means 'similar forms'. The following is an example of an entry in the ISN:

- indwangu** bz 9. *Indwangu yinyamazana elesiphongo, ifana lenkawu kodwa yona inkulu. Indwangu ziqede umumbu emasimini. FAN imfene.*
(**baboon** n. 9. A *baboon* is an animal that has a forehead, it is like a monkey but is bigger. *Baboons ate all the maize in the fields. SYN imfene.*)

Tone markings of each headword were not provided throughout the diction-

ary. This was a deliberate decision taken by the editors after considering both the merits and the demerits of omitting it. Whereas Ndebele like most Bantu languages is tonal, the omission of tone marking therefore excludes the complete grammatical representation of the language. It was argued that, since the ISN is a monolingual dictionary, a mother-tongue speaker would hardly require tone marking of two or more homonyms, for the context can easily disambiguate them. Further, it was considered that the main target users, i.e. secondary school pupils, are not at a level where they would have been introduced to such grammatical detail, so that it would have caused complications. As a result, it was decided to include tone marking, together with the phonetic transcription of each headword, in the planned *Advanced Ndebele Dictionary*, which will be a successor to the ISN with target users anticipated to be mainly college and university students and language practitioners.

The ISN has a total of 20 008 headwords. In order to make the selection of headword entries as widely representative of the culture of the Ndebele people as possible within the compass of a general-purpose, medium-sized dictionary, the editors ensured that the collection of material was done as widely as possible. The four provinces where Ndebele is predominantly spoken were identified, that is, Matebeleland South, Bulawayo, Matebeleland North and the Midlands provinces. Areas where research was carried out were in secondary schools, colleges and universities, especially for terms covering language and literature, the media, sports and games, religion, trade and commerce, law and administration. Research in rural areas covered terms dealing with the traditional life and the rural livelihood of the Ndebele people, including aspects such as hunting, farming, marriage ceremonies, traditional dances and chief's courts. Oral interviews were carried out and books (covering different genres like novels, drama, poems and text books) were collected to create a corpus from which to draw headwords, definitions from contextual usage and examples. Thus, the ISN is a corpus-aided dictionary, using "language that has actually come from the mouths or pens of ... speakers themselves", as Fortune (2001: 122) puts it.

Target Users of the ISN

As has been stated above, the ISN is a general dictionary. This means that it is a general-purpose reference work, appealing to a wide spectrum of users. The work has a four-fold objective. The first is to provide a medium-sized dictionary aimed at teachers and students at secondary school. This is consistently emphasized because it determines the size of the dictionary. A medium-sized dictionary, easy to handle, was thought to be appropriate, for it would be cheaper, making it affordable to the target users. The dictionary was projected to assist teachers and students to understand and teach the structure of their language through the provision, for the first time, of a technical terminology in Ndebele, describing its linguistic features. Teachers and students of Ndebele

are more likely to need to consult a dictionary than other users and to use its contents in the course of their daily lives as well as to convey its contents to others (Fortune 2001: 122).

The second objective is to facilitate the use and understanding of Ndebele for mother-tongue speakers. For young children, the mother tongue is the first means of communication and the chief instrument for interpreting and understanding the world in which they live. Without mastery of language, there can be no mastery of life. For school children in particular, the mother tongue is important. Children acquire knowledge very rapidly if they are allowed to build on what they already know. All children bring their mother tongue to school. The ISN hopes to impact positively on schoolchildren facing the difficulty of receiving instructions in English, currently the only official language in Zimbabwe.

For the ordinary reader, such a reference work can provide, in an easily understandable manner, the meaning, use, and function of words which would not be so fully grasped if conveyed in, and then translated from, a foreign language as has been the case hitherto. The dictionary contains the language these users utilize every day; it reflects their traditional, rural and even contemporary use of language. "Now sifted and sorted, linguistically analysed and self-defining, it is handed back to them to put to use," says Fortune (2001: 122). It mirrors, as it were, their lives and reflects it.

The third objective is to provide a unitary language, which is the property and inheritance of all. Although the Ndebele people believe that Ndebele does not have regional variants, evidence from the corpus suggests otherwise. Thus contributions from all Ndebele-speaking regions were sought to add to the vocabulary of the language, particularly the vocabulary that has come as a direct influence from the "dominated so-called minority groups" like Venda, Sotho and Kalanga. This resulted in a good deal of synonymity. As far as synonyms were concerned, widely-used forms were treated fully with definition and exemplification, while forms with a narrower distribution were entered as headwords, receiving a shorter treatment by referring users to the main entry. This procedure is a straightforward and economical way of promoting the unification of the language, while reflecting its inner differentiation of various regional uses.

The fourth objective, which Fortune (2001) refers to as the "animating heart of the ALLEX project" is to promote the status and use of the language. The dictionary, it is hoped, will help to make people use it fittingly in widening areas of life, and to value it as conferring self-respect and a means towards an improved and developed standard of life. This will be achieved by a greater participation in affairs at all levels through the medium of the mother tongue. It is envisaged that the Ndebele language will be capable of developing into a medium of expression and communication which enables the Ndebele people to carry out their affairs in all spheres of life in their mother tongue and to communicate at all levels of their concerns, aims and techniques for national development. It therefore wants to make a contribution to a change of policy

for Ndebele to be recognized as an official language in Zimbabwe. Hitherto it has remained inferior to English, which is the only official language.

Related to the fourth aim is the fact that the ISN will contribute towards language raising. According to Chimhundu and Gronvik (2007: 1), language raising is a term coined and used in Norway around 1850, describing an effort to "create a standard language (from the vernacular), cultivate it and have it accepted in the language community where it belongs". Language raising refers to all struggles and attempts to develop a literary language from the basis of a vernacular. Once the language has been sufficiently developed, it assumes wider functions in education, industry, commerce and law. It is therefore envisaged that the ISN is to contribute to the conducting of all kinds of education in the mother tongue. This is the medium through which education is most effective and personalized, and through which it can best permeate through society, raising the level of skills and abilities with a resultant higher standard of life. The ISN is a firm basis upon which to build.

User-Friendliness of the ISN

User-friendliness can be viewed in two ways. It can pertain either to the size and cover format of the dictionary or to the presentation of the definitions and their examples. The ISN is user-friendly because, being not too bulky, it is easy to handle. It is medium-sized and therefore big enough to cover a wide variety of terms. It has a large cover format. The fonts and spacing of the text fits more elegantly and more legibly within its wider margins. This clear layout and presentation simplify it for the user to access information in the dictionary.

Even to general users, it is basic knowledge that a dictionary lists words and their meanings. Consequently, the importance of a definition which is an intrinsic component of a dictionary cannot be overemphasized. In fact, defining is probably one of the most important exercises in the dictionary-making process. It is because of this rigorous process that users can interact with a particular word, discovering its meaning as explained in each definition.

This makes defining formats, a format being "the way in which something is arranged and presented" (Sinclair 1990: 218), crucial in composing definitions. Defining formats are pertinent insofar as they help the editors to write clear, concise, specific, systematic and culturally relevant definitions that make dictionaries user-friendly (Khumalo 2002: 266). These formats are geared towards providing editors with certain guidelines to follow in constructing a definition, hence contributing to consistency in the treatment of related sets of headwords. This view is accentuated by Gellerstam's (1993: 47) observation that "the development of definition formats is advantageous in that once a particular format has been agreed upon, a number of words can be defined with it". The guiding principle of the editors throughout the compilation of the ISN was, therefore, to produce a user-friendly dictionary.

The editors of the ISN used the Collins Birmingham University Interna-

tional Language Database (COBUILD), that is, the COBUILD-type format for their definitions. Defining is one of the most important processes in dictionary making because it is important for users to know the meanings of words so that they can use the words in a *prescribed* manner in both spoken and written language. A dictionary is prescriptive insofar as it distinguishes the formal and colloquial uses of any particular word.

The following are COBUILD defining formats as outlined by Hanks (1987) and Moon (1987):

- (a) The headword should appear in the definition.
- (b) The headword should be shown in a typical context of common usage because "context disambiguates" (Hanks 1987: 125).
- (c) Definitions should consist of full sentences.
- (d) Explanations should consist of two parts, with the first part showing the word in use and the second one explaining the meaning.
- (e) The format should suggest a preference rather than a restriction.

For COBUILD, a good definition should be short, culturally relevant and consistent with others of its type, using superordinate terms (Hanks 1987: 118). When a definition has been made simple and familiar, it thus helps the dictionary as a whole to become user-friendly. The COBUILD system demonstrates a marked departure from traditional lexicography which, according to (Hanks 1987: 117), "heavily relies on the use of parentheses whose function is not based on the standard usage of ordinary written English". The idea was to create a "dictionary as prose" (Hanks 1987: 118). COBUILD has therefore abandoned the use of parentheses because they are not user-friendly. Considerable sophistication is required on the part of dictionary users to work out what precisely is meant by each set of parentheses. This was an important view to consider, because the ISN was a first monolingual dictionary in Ndebele, which had no previous lexicographic tradition for the users to have developed the requisite sophistication. Therefore the COBUILD formats proved very useful in achieving user-friendliness, which was the editorial team's goal.

Evident characteristics of the COBUILD system of defining formats are: definitions should be as natural as possible, should project typical usage of a headword, and should be easy to understand. These guidelines were instrumental for the editors in organizing, crafting and presenting user-friendly definitions. In the ISN, each definition consists of a complete sentence, so that the user is shown the word as it appears in natural language use. Such a definition illustrates both the typical grammatical context and the typical use of the word. From the user's point of view, this method of definition is particularly useful for an inflecting language like Ndebele where headwords are not necessarily orthographic words. A good example is the following:

-thenga sz mwa. Nxa *uthenga* ulutho uyabe ulukhokhela imali ukuze uluthole.
(**buy** transitive verb. If you are *buying* something, you will be taking out money for it so that you can have it.)

The definitions present "the typical case" in such a way that the meaning is made clear. This enables dictionary users to become familiar with its use and further applications, so that they can employ it in sentences of their own. When the definition is perceived to be insufficient in itself, then an illustrative example is provided to clarify it further, always by way of a complete sentence extracted from the corpus which provides natural language in use.

Examples in the ISN

As a source of data for linguistic analysis, corpus work has played an important role in linguistic research, and an even more important role in lexicography. Kennedy (1998: 88) observes that a linguistic corpus, in whatever form, is important as a basis for more accurate and reliable descriptions of how languages are structured and used. Sinclair (1991: 171) defines a corpus as "a collection of naturally-occurring language text, chosen to characterize a state or variety of a language". According to Renouf (1987: 1), it is a body of collected texts, of the written or spoken word, which is stored and processed on computer for the purposes of linguistic research. Because of the importance of the corpus in any dictionary-making process, i.e. for purposes of headword and sense selection and obtaining examples of natural language usage, the ISN editors had collected a corpus of about a million running words before starting to compile the dictionary. The Ndebele corpus currently stands at around two and half million words. The editors greatly benefited from the Ndebele corpus while writing definitions for the ISN.

Each word occurs in its own special environment. This is referred to as the cotext. By studying the cotext, the editors managed to work out the various senses of each word in these different environments. This made a seemingly difficult task, which may take time when intuition is used, simple by employing the corpus to show up various instances of a particular word using a software program. Senses such as those listed below were identified through this quick method:

umuzi bz 3. *Umu*zi yindawo kumbe ikhaya lapho okuhlala kona abantu.
(**home** n. 3. A *home* is a place or a dwelling where people live.)

umuzi bz 3. *Umu*zi lichatha elimnyama emzimbeni umuntu azalwa elalo.
(**birthmark** n. 3. A *birthmark* is a coloured mark on the skin, which is there from birth.)

umuzi bz 3. *Umu*zi butshani obude obukhula ikakhulu emaxhaphozini.
(**umu**zi n. 3. *Umu*zi is a type of grass which grows on marshy ground.)

Concordances provide lexicographers with real text examples. It demonstrates a particular word in a typical sentence in agreement with other naturally co-occurring words. According to Sinclair (1991), examples from concordances have a status which made-up examples do not have, because they are typical.

The ISN editors obtained the examples for most of their headwords from the corpus, which provided the dictionary with more "natural" examples. These examples were viewed as a very important component of the definition. The ISN editors resolved that

- (a) since examples are part of the definition, they should be used only when they serve to complement the definition by clarifying the headword, that is, to illustrate meanings or uses only when it is necessary,
- (b) only one example should be given for reasons of saving space, and
- (c) they should also be short like the definitions.

These decisions were meant to benefit the users by making the dictionary more accessible and user-friendly. In many cases, the definition of a headword is so informative that there is no need for a further sentence to use as an illustrative example. However, when the definition may not cover all the senses inherent in a headword, providing an illustrative example clarifies the meaning of the word, enabling users to learn its use and further applications so that they can employ it in sentences of their own.

The Front Matter and the Back Matter

The introduction to the ISN, the front matter, gives, among others, the following important background information: (a) the history of the Ndebele people, their identity, the sociolinguistic factors that have influenced their language and the history of Ndebele orthography, (b) the state of the Ndebele corpus, (c) a brief description, i.e. a mini-grammar, of the structure of the Ndebele language, (d) the defining formats used for different parts of speech such as nouns, verbs, ideophones and adjectives, and (e) the metalanguage list, or the technical language the editors employ, partly derived from current use and partly specifically devised. The latter is presented in three columns giving (a) the abbreviations, (b) the Ndebele terms in full, and (c) the English equivalents. For example:

bz	ibizo	noun
lw	ulwangeni	palatal

The presentation of the alphabetically arranged terms in three columns provides an index for cross-reference. This, together with a comprehensive mini-grammar, is expected to help towards the teaching of Ndebele in Ndebele in secondary schools. Hitherto Ndebele has been without a grammar book describing the structure of Ndebele in Ndebele. It is in the field of linguistics and literature that the first creation of technical terminology has been made.

The back matter completes the structure of the dictionary. It is additional information inserted at the back of the dictionary, which the editors felt was useful and enriching to the target users. The back matter covers technical terms

in mathematics and government institutions, including judiciary information. It also covers a variety of cultural information. Most of the terminology is currently in use in Ndebele, but has not been formally presented, this being an attempt by the editors to standardize it.

Conclusion

The following points can be noted in conclusion: (a) the ISN is the first-ever monolingual Ndebele dictionary, (b) it is a medium-size, general-purpose dictionary drawn from language used in the everyday lives of the Ndebele people, (c) it covers the history of the Ndebele people and their culture, past and present (in the historical and contemporary sense), both in the front and the back matter, and (d) the main target users, learners and educators, will benefit not only from the definitions in the word list but also from the mini-grammar in the front matter. Finally, it can be said that a serious attempt was made to produce a user-sensitive, user-friendly and user-driven work for its envisaged target users.

References

- Chimhundu, H.** 1992. Standard Shona: Myth and Reality. Crawhall, N.T. (Ed.). 1992. *Democratically Speaking: International Perspectives in Language Planning*: 77-88. Cape Town: National Language Project.
- Chimhundu, H.** 1993. The Status of African Languages in Zimbabwe. *Southern African Political and Economic Monthly* 7(1): 57-59, October.
- Chimhundu, H. and O. Gronvik.** 2007. *The Alex Project and Language Raising in Zimbabwe*. A paper presented at the NUFU Dissemination Conference, Bergen, Norway, 5 June 2007.
- Eliot, W.A.** 1897. *Dictionary of the Tebele and Shona Languages*. Cape Town.
- Fortune, G.** 2001. Book Review: *Duramazwi reChishona*. *Zambezia* 28(1): 121-127.
- Gellerstam, M.** 1993. The Use of the Corpus. Chimhundu, H. (Ed.). 1993. *Report on the African Languages Lexical (ALLEX) Project Planning and Training Workshop*: 45-56. Harare: University of Zimbabwe.
- Hadebe, S.** 2002. *The Standardisation of the Ndebele Language through Dictionary-Making*. D.Phil. Thesis. Harare/Oslo: University of Zimbabwe/University of Oslo.
- Hadebe, S. et al. (Eds.).** 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hanks, P.** 1987. Definitions and Explanations. Sinclair, J.M. (Ed.). 1987: 115-136.
- Kennedy, G.** 1998. *An Introduction to Corpus Linguistics*. London/New York: Longman.
- Khumalo, L.** 2002. Defining Formats and Corpus-based Examples in the General Ndebele Dictionary, *Isichazamazwi SesiNdebele*. *Lexikos* 12: 264-274.
- Moon, R.** 1987. The Analysis of Meaning. Sinclair, J.M. (Ed.). 1987: 86-103.
- Pelling, J.N.** 1971. *A Practical Ndebele Dictionary*. Revised Edition. Harare: Longman Zimbabwe.
- Renouf, A.** 1987. *Corpus Development*. Sinclair, J.M. (Ed.). 1987: 1-40.
- Sinclair, J.M.** 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J.M. (Ed.).** 1987. *Looking Up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London/Glasgow: Collins ELT.

Divergent Approaches to Corpus Processing: The Need for Standardisation*

Esau Mangoya, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe (emangoya@arts.uz.ac.zw)*

Abstract: This article discusses some problems encountered in the processing of the Shona corpus. Most of the problems deal with the handling of adoptives, punctuation and individuals' idiolects. It also discusses the problem ensuing from an attempt to standardise the formats used in the handling of the corpus. The way a corpus is processed is critical in determining its quality. This article aims to show how the different linguistic backgrounds of the processors affect the appreciation of some vital aspects of the corpus. One of the acclaimed advantages of a corpus is that it allows research to be done on natural language. An ideal corpus should be a body of texts combined in a principled way to become a reliable language bank from which researchers retrieve data for various research purposes. With a good corpus, data can be provided giving an authoritative body of linguistic evidence which can support generalisations and against which hypotheses can be tested. As this proves the invaluable status of a corpus, the article assesses the processing of the Shona corpus and discusses how some aspects of the processing may impact negatively on its quality.

Keywords: CORPUS, STANDARDISATION, DATA, HEADWORD, LEXICOGRAPHY, RETRIEVE, TRANSCRIBE, TAGGING, ENCODING, PARSING, DIVERGENT

Opsomming: Uiteenlopende benaderings tot korpusverwerking: Die behoefte aan standaardisasie. Hierdie artikel bespreek 'n aantal probleme wat teëgekomp is by die verwerking van die Sjonakorpus. Die meeste van die probleme handel oor die hantering van leenwoorde, punktuering en die idiolekte van individue. Dit bespreek ook die probleem wat voortvloei uit 'n poging om die formate gebruik vir die hantering van 'n korpus te standaardiseer. Die manier waarop 'n korpus verwerk word, is krities vir die bepaling van sy gehalte. Hierdie artikel wil toon hoe die verskillende taalkundige agtergronde van die verwerkers die beoordeling van sommige van die wesenlike aspekte van die korpus beïnvloed. Een van die geloofde voordele van 'n korpus is dat dit toelaat dat navorsing oor die natuurlike taal gedoen word. 'n Ideale korpus behoort 'n geheel van tekste te wees wat op 'n geordende manier saamgestel is om 'n betroubare taalbank te wees waaruit navorsers inligting vir verskillende navorsingsdoeleindes kan verkry. Met 'n goeie korpus kan gegewens verskaf word om 'n betroubare geheel van taalkundige bewyse te gee wat veralgemenings kan bevestig en waarteen hipotesisse getoets kan word. Deurdat dit die

* This article is a revised version of a paper presented at the Seventh International Conference of the African Association for Lexicography, organised by the Dictionary Unit for South African English, Rhodes University, Grahamstown, Republic of South Africa, 8–10 July 2002.

waardevolle status van 'n korpus bewys, beoordeel die artikel die verwerking van die Sjonakorpus en bespreek hoe sommige aspekte van die verwerking die gehalte negatief mag beïnvloed.

Sleutelwoorde: KORPUS, STANDAARDISERING, DATA, TREFWOORD, LEKSIKOGRAFIE, VERKRY, TRANSKRIBEER, ETIKETTERING, ENKODERING, WOORDBENOEMING, AFWYKEND

Introduction

According to McEnery and Wilson (1996: 29), the Latin term *corpus* means 'body'. In linguistic studies, the meaning has been extended to indicate any 'body of texts'. It may basically be referred to as a collection of texts combined as a databank which can be used as a tool for any linguistic related research. Because the possible utterances in a language variety are infinite, the corpus can be considered as a sample. This sample is meant to be maximally representative of that language variety, presenting a clear picture of possible tendencies and proportions of linguistic elements (McEnery and Wilson 1996: 30). As a result, different varieties of texts and genres are to be included in the corpus, for instance novels, poetry, drama, newspapers and recorded interviews. For serious and comprehensive long-term use, the corpus has to be continuously developed with more and new texts being added. Through the continuous addition of new data, the changing store of texts can always reflect previous and current linguistic behaviour (Aarts and Meijs 1984: 4). The overall purpose of a corpus is to have raw data for use in various kinds of research. This is one of the newer approaches which has helped in the differentiation of linguistic studies. Hence, today, in linguistic study disciplines, corpus-based syntax and semantics are contrasted with non-corpus-based syntax and semantics. Different corpora may be compiled for various purposes and interests. Over the years, the processing of the corpus has evolved in such a way that today the mentioning of the term *corpus* automatically implies machine-readable data. In the past, the corpus could be in printed form. A corpus remains a basic data reference collection on which various forms of research can be carried out. McEnery and Wilson (1996: 32) stress the importance of a corpus: 'As a standard corpus also means that a continuous base of data is being used and thus variation between studies may be less likely to be attributed to differences to the data being used and more to the adequacy of the assumptions and methodologies in the study.'

In corpus-based lexicography, the corpus is used for headword selection, defining and providing examples. The use of a corpus in dictionary compilation is of particular interest to this article. Six corpora are currently being compiled by the African Languages Research Institute (ALRI) at the University of Zimbabwe, Harare, with the aim of using them in the compilation of monolingual and bilingual dictionaries. The corpora being processed are Shona, Ndebele, Kalanga, Nambya, Tonga and Shangaan.

While the Shona and Ndebele corpora are at an advanced stage of processing, the other four are in their initial stages. The Shona corpus stands at close to five million running words, and the Ndebele corpus at around three million running words. In this article, the focus is on the processing of the Shona corpus. The texts being included in the corpus consist of written material and oral interviews. The written material constitute about a quarter of the total corpus. This article aims to discuss some of the problems encountered during the processing of oral material.

To be a research tool, the corpus has to be qualitative. When there are inconsistencies in the handling of the corpus, some of the poorly processed materials will find their way into the language bank. A false impression on the size and quality will then be created. The poorly processed corpus may have retrieval limitations. Therefore, whatever is to constitute the language bank has to be processed adequately if it is to remain useful, accessible and relevant to research.

Except that the processing of a corpus takes a considerable time, in most cases it usually involves a chain of different individuals for the different stages, as shown here:

text creation → transcribing → encoding → tagging → parsing

Text creation focuses on the interviewee in the case of oral material or the author in the case of written material. However, the focus of this article is the processing of oral material. Transcribing involves the scripting of information from a cassette or recorder onto paper. Encoding is the keying in of the text from the paper into the computer for electronic storage. Tagging is the marking of the documents for purposes of retrieval from the corpus. Parsing is the application of the proof-reading programme to check for consistency of the tagging. All these stages may involve different individuals who have a different conceptualisation of certain aspects of the words in a language.

At times, this led to different approaches by individuals working on the corpus. The differences emanated from the nature of the treated oral texts. Some of the problematic material came from language contact areas.

Influence of language contact

One of the problems that lead to inconsistencies in the processing of the Shona corpus is caused by borrowed words, particularly from the Nguni languages and English with which Shona is in contact. The problematic words from Nguni mainly come from three fronts. There are words from Ndebele, a language spoken in the south-western parts of Zimbabwe. Ndebele is one of the Nguni group of languages also including Xhosa, Zulu and Swati which are spoken in the Republic of South Africa. The Shona and Ndebele groups came in contact during the second quarter of the 19th century in the time which has

come to be known as the Mfecane (Sibanda 1989: 25). This was a period of political instability among the Nguni which caused Mzilikazi and his followers to cross the Limpopo into the present-day Zimbabwe. A language contact zone was created in the areas stretching from Kwekwe down to Chirumhanzu south of Chivi to Mberengwa and across the Lowveld to Matibi.

These language contact areas have played a big role in the borrowing of words from Ndebele into Shona. Also at a later stage, around the 1890s during the time of the British colonisation, the African groups accompanying the British as guides were mainly from the Nguni group. Since these people spoke different languages, they communicated mainly through a pidgin which has come to be known as Fanagalo. This Fanagalo was composed mainly of words from Afrikaans, English and the Nguni languages. So, as the pioneer column moved up into Mashonaland, they created a second form of contact, thereby reinforcing the initial language contact as more words came into Shona through this new encounter. As a result many words from Nguni were introduced and accepted into Shona. Examples of such words are given under (1).

- (1) *-funda* (learn)
mufundisi (reverend/teacher)
-zama (try/attempt)
-bopa (inspan)

These words, originally from Nguni, have been introduced and accepted into the Shona corpus as borrowed words. Their adoption is no longer questionable since their introduction and reinforcement through the second form of contact also coincided with the introduction of institutions in which they have been extensively used. The words *-funda* and *mufundisi* are widely used in educational and religious circles, and *-zama* and *-bopa* are quite prevalent in agricultural and industrial sectors.

Adoption is a vital linguistic phenomenon which cannot be ignored, particularly when the corpus has to be used for the compilation of dictionaries. The reason why it was not crucial to consider the etymology of these words at this stage was that, until then, the focus had been on the compilation of general dictionaries, *Duramazwi reChiShona* and *Duramazwi Guru reChiShona*, published in 1996 and 2001 respectively, the first being a general medium-sized dictionary and the second an advanced dictionary.

However, there is a second group of words about whose status there were divergent views. These are words of common language usage, having a high frequency in Shona. They are words that feature much in social conversations. Some team members believed that they had to be considered foreign while others felt the opposite. The reason is that some of these words were more acceptable as already adopted compared to others. Examples of such words about which there were divergent views are given under (2).

- (2) *mnandi* (delicious/sweet)
mgane (friend)

- sakaza* (speak)
- saba* (be afraid of)

Part of the reason why team members felt that these could be acceptable Shona adoptives is that Shona orthography can easily make provision for them, representing them as *munandi*, *mugane*, *-sakaza* and *-saba* respectively. A lack of clear guidelines leaves the processors making subjective decisions about the categorisation of words in the corpus. Resultant discussions led to the conclusion that *mnandi* and *-sakaza*, when their frequent featuring in various literary genres is considered, could be regarded as adoptives as long as they were made to conform to Shona orthography.

There is a high level of subjectivity about which words should be considered adoptives. This subjectivity is also influenced by the orthographic closeness between the source language of the word and the adopting language. Therefore, some Nguni words found in the oral texts were unanimously considered foreign, basically owing to the difference in orthography between the two languages. Despite the long years of language contact, some of the words have not been accepted into Shona. The team was agreed that these were to be clearly indicated and rightly tagged as foreign in the corpus as shown by the examples under (3).

- (3) <foreign> *khombisa* </foreign> (seek for love)
<foreign> *thanda* </foreign> (love)
<foreign> *nkosi* </foreign> (king)

The guideline formats for corpus processing assumes that it is clear what constitutes a foreign and what an adopted word. It does not consider that there are different levels of acceptance by different individuals. Guideline formats should be designed in such a way that at different levels of processing, it becomes clear how to handle different words from other languages.

As pointed out, the main problem that led to contradictions was differences in spelling where cluster combinations in Ndebele such as *kh*, *th* and *nk* are unacceptable in Shona. There were situations where the words under (3) would be adapted to Shona orthography. These would appear as *kombisa*, *tanda* and *ngosi* respectively. Despite being written according to Shona orthography, however, they were tagged as foreign, seemingly under the influence of the way they are spelt in the source language.

There was no immediate solution to the problem of handling words from other languages. It was suggested that all the words individuals came across could be submitted to the team panel to consider their status. The question of subjectivity still remained central though. It depends on the individual to choose words that need to be discussed by the panel. It also requires knowledge to recognise when the status of a word is not clear, selecting it for consideration by the team. Because corpus building has to be a continuous process, panel meetings also have to become routine.

On the other hand, the Mfecane created another form of language contact in the eastern part of Zimbabwe. A further group, the Shangaan, which was linguistically closely related to the Nguni group, also crossed the Limpopo just about that time. The Shangaans did not settle as a bigger group like the Ndebeles. After having crossed the Limpopo, small groups remained behind while others proceeded into Zambia and finally into Malawi. Linguistically, an interesting scenario was created in the south-east of Zimbabwe stretching into Mozambique where Ndaou, one of the dialects of Shona, is spoken. The migrating groups actually fused with the Shona groups they found in those areas. Over the years, because of this language contact situation, the Shona spoken in these areas adopted some lexical items. This resulted in the creation of a peak dialect, Ndaou, spoken in and around the Chipinge district. There now exist linguistic elements that are problematic when they appear in the corpus. Ndaou has many salient linguistic elements when compared to the rest of the Shona dialects.

Inconsistencies occurred in the handling of texts from the rest of the Shona-speaking areas and those from the Ndaou dialect. In the texts from the rest of Mashonaland, some of the words from the Nguni group were clearly marked foreign while those very words are actually accepted as part of the vocabulary of Ndaou. Examples from the Ndaou vocabulary that would automatically be marked as foreign if they appear in texts from other dialect regions are given under (4).

- (4) -*tshaya* (beat)
 -*nqondo* (brain)
 -*gqoka* (put on)
 -*qonda* (go straight)

This historical background information on the language situation is not provided to the user of the corpus. When these words feature in any of the Shona dialects except Ndaou they are treated as foreign. At the same time, these have become natural Ndaou words which are recognised as such. Once it is realised that it is a Ndaou text, they are not treated and marked as foreign. However, the overall analysis of the whole corpus may give the impression that it has been poorly processed.

Part of the problem is that the Ndaou dialect was not greatly taken into consideration during the standardisation of the Shona dialects. Doke (1931), who played a pivotal role in the standardisation of the Shona dialects, recommended that the words from Ndaou should be used sparingly. There are new challenges now as the corpus has to reflect the living language of the people. If a representative corpus is to be produced, Ndaou should be considered as any other Shona dialect. What is evident here is that the problems arise as a result of neglecting Ndaou during standardisation. In some instances, the distinctive spelling of the Ndaou words in the corpus are not scripted but are deliberately removed and substituted with those allowed in general Shona orthography.

This makes it very difficult to recognise them in the new form. Examples of a group which involves the adapted words are shown under (5).

(5)	Original Ndau	Adapted form
	<i>nqondo</i>	<i>n'ondo</i> (brain)
	<i>-gqoka</i>	<i>-goga</i> (put on)
	<i>-qonda</i>	<i>-konda</i> (go straight)

The other adapted group consists of words which involve consonantal combination *hl*. The standard orthography does not allow for the spelling Ndau speakers would prefer. Some examples of these are given under (6).

(6)	Original Ndau	Adapted form
	<i>-hlaba</i>	<i>-shava</i> (pierce through)
	<i>-hlupa</i>	<i>-shupa</i> (trouble someone)
	<i>muhlobo</i>	<i>mutyovo / mutyowo</i> (way / method / type)

Once in the original form, they are marked as foreign, but are unmarked when adapted. In the case of *-hlaba* which has only one meaning, the adopted form becomes *-shava* which has three meanings in Shona. This again causes complications and more inconsistencies. All the stages of corpus building must represent what the creator of text really meant and intended. As a result, the tendency is to consider anything not conforming to the standard as foreign. If not marked foreign, there is forced adaptation.

When these words were brought to the team panel for consideration, it was agreed that as long as the text was Ndau they would not be marked as foreign. If they appeared in texts of other dialects they had to be given a foreign tag. However, this is a temporary solution implemented for the processing of the corpus leading to evident inconsistencies. It only suits the corpus processors but the solution does not address the major issue which causes these discrepancies. What is evident here is that the problems concern language planning and language policy. Corpus building is a grant project that should involve all language stakeholders including the government who has to act on matters of language planning and language policy. As for the corpus that has been produced, the problem of inconsistencies should be explained. The language situation should be outlined in order to inform the user about the existence of the inconsistencies and the reasons for it.

Another form of language contact also exists in Zimbabwe. Zimbabwe was a British colony in which English was declared the official language. English was raised above all the indigenous languages of the country. With the attainment of independence, Shona and Ndebele were also accorded official status alongside English. As a result, some English lexical items have found their way into the indigenous languages. In cases where these lexical items have been partially adapted into Shona, they were problematic in the processing.

Partially adapted adoptives from English

English, which enjoyed the monopoly of being the only official language for a long time before independence had much influence on the local languages. As a result there are many adoptives from English in Shona. The majority of people in Zimbabwe are bilingual, speaking their local indigenous languages and having English as a second language. It is this knowledge of the two languages which results in speakers adapting one part of English words and leaving the other part unaltered.

The handling of these partially adapted adoptives from English was problematic, because of this partial borrowing of lexical elements. Once words have been adopted, they become acceptable lexical items of the borrowing language. However, these are instances where some elements of the words have remained partly in their original form. Examples of these are given under (7), the unaltered elements being shown in bold italics.

- (7) *kusilaidha* (to slide)
*hazvisi ra**ightka*** (it is not right, is that not so?)
*kusasi**pecta*** (to suspect)

Because the majority of the sounds in the words are also found in Shona, some processors of the corpus felt the words could be considered as already accepted into the language. As a result they were left unmarked in some of the Shona corpus texts. Although the majority of the syllables have been remorphologised and rephonologised, some elements remain unchanged. So some corpus processors would mark only the unchanged part of the word as foreign as shown in the examples under (8).

- (8) *kusi*<foreign>*lai*</foreign>*dha* (to slide)
hazvisi ra<foreign>*ight*</foreign>*ka* (it is not right, is that not so?)
kusasi<foreign>*pect*</foreign>*a* (to suspect)

What was marked as foreign are just sections of words that have maintained their identity in spelling from the source language.

While the other parts of the words conform to the writing system of Shona, those marked as foreign are not full words, neither are they morphemes which have a clear meaning for the user of the corpus. The processors of the corpus have tried to indicate that there are notable foreign elements in the word structures but what has been marked is not useful for meaningful linguistic research.

In the absence of a clear policy, the way texts may be handled by different processors will vary, leading to inconsistencies. As has been demonstrated above, different aspects of the language may be perceived differently and uniformity in the processing of the corpus may be difficult to achieve. The guide formats just mention that foreign words should be marked, but as the above examples under (8) demonstrate, the criteria for identifying foreign words are

not laid down. They do not address these new challenges.

Consequently there are inconsistencies in the processing of various words particularly loanwords. The other alternative which was suggested in the discussion of the problem was to mark the whole structure as foreign, as shown in the examples under (9).

- (9) <foreign>kusilaidha</foreign> (to slide)
hazvisi <foreign>raightka</foreign> (it is not right, is that not?)
<foreign>kusasipecta</foreign> (to suspect)

This would override the fact that certain parts of these lexical items have already been adopted in and adapted to Shona.

However, when looked at more closely, the problem really stems from the fact that those handling the corpus are bilingual. The way a word may be presented in scripted form from the orally produced text is subjective. This is the reason why these words have various presentations in the corpus files. There has to be full communication with all involved in the processing of the corpus. In areas where there are divergent approaches, discussion is necessary and the consensus reached should be recorded. Such records should serve as a guide to the processors of the corpus. This is important for consistency. These records should be used as manuals guiding users of the corpus in the way some words were handled during the processing.

ALRI's standard formats for the processing of the Shona corpus have assumed that foreign words would be easily identified. No consideration has been given that some words would be changed in the process of adoption. As no prescriptions are given for the handling of partially adapted words, they are always problematic in the processing of the corpus. Without standard formats, individual preferences take precedence. This gives rise to the issue of the individual idiolect also coming into play. According to Crystal (1991: 170), an idiolect refers to the linguistic system of an individual speaker — one's personal dialect.

Varying idiolects

Individuals vary in their idiolects. This applies to both the text creator who in the case of oral texts is the interviewee and the processor of the text. This becomes evident in the punctuation of different texts. Certain aspects were handled differently because of a different conceptualisation of the punctuation of texts. One of the problematic areas is the 'probe statements' of the interviewee. Examples (10)(a), (b), (c) and (d) demonstrate the differences in the handling of the probe statements resulting from varying idiolects.

- (10) (a) *Zvinenge zvichida kuti kana wasvika wodini ... Wodzikama. Zvinhu zvobva zvodini ... Zvofamba nenzira yazvinofanira kufamba nayo. Zvozodini ... Zvoopera.*

It is necessary that when you arrive you what ... Remain cool. Then everything will what ... Go the way it should move. Finally it will what ... it comes to an end.

(b) *Zvinenge zvichida kuti kana wasvika wodini? Wodzikama. Zvinhu zvobva zvodini? Zvofamba nenzira yazvinofanira kufamba nayo. Zvozodini? Zvoopera.*

It is necessary that when you arrive you what? Remain cool. Then everything will what? Go the way it should move. Finally it will what? It comes to an end.

(c) *Zvinenge zvichida kuti kana wasvika wodini, wodzikama. Zvinhu zvobva zvodini, zvofamba nenzira yazvinofanira kufamba nayo. Zvozodini, zvoopera.*

It is necessary that when you arrive you what, remain cool. Then everything will what, go the way it should move. Finally it will what, come to an end.

(d) *Zvinenge zvichida kuti kana wasvika wodini! Wodzikama. Zvinhu zvobva zvodini! Zvofamba nenzira yazvinofanira kufamba nayo. Zvozodini! Zvoopera.*

It is necessary that when you arrive you what! Remain cool. Then everything will what! Go the way it should move. Finally it will what! It comes to an end.

Discussions by the ALRI team on how to handle such statements revealed divergent preferences. Different processors had different ways of punctuating the statement which resulted in four versions of the same statement. In example (10)(a), the processor marked the end of the seemingly unfinished statements with ellipses. In (10)(b), the processor preferred using question marks, feeling that the speaker was asking rhetoric questions he would immediately answer himself. Commas were preferred in example (10)(c), the reason given was that after the probe there was a pause before the statement was finally completed. The preferred punctuation in example (10)(d) was the exclamation mark, the reason advanced that the interviewee was penultimately stressing a point before the actual completion of the statements. After a common approach to such texts had been debated, the use of a comma was finally agreed upon.

Whatever decision made had implications on the whole set of words in the given sentences. For example: The use of a comma would mean the following word would start with a small letter. The use of exclamation or question marks would render the examples into complete statements, having implications for the final corpus text. It is in the corpus processing that a text may be given a particular value. The processor of the text may decide whether the speaker uttered an exclamation, made a full statement or asked a question. This shows the importance of decisions on how to handle certain aspects of idiolect and register. After this particular case had been discussed, it was agreed that using

a comma would be preferable. However, this background information will not be passed on to the user of the corpus text who might have an own opinion on how the statements should have been punctuated, hence the need for written records on how particular aspects were handled.

The ALRI team agreed that the best way was option (10)(c). This consensus was only reached after there had been divergent treatments of these statements resulting in four options. The original guiding formats indicated where punctuation marks should be placed before the various tags are used in the text. It did not, however, deal with challenges arising from the actual punctuation of a document by the processors. The team's agreement has become the standard way of punctuating such statements. It is nevertheless important to record this decision for the users of the corpus so that they could have a fuller understanding of how the documents were treated. All the background information on the standardisation of the formats should be recorded. Standardisation should be continuous since new challenges continue to become manifest. One of these problems was how to handle factual distortions.

Factual distortion

Few cases were found where texts were factually erroneous. These could result from either a lack of actual knowledge about a subject or an unintended mistake by the creator of the text, or a misrepresentation caused by an oversight in the line of corpus production. However, the resultant corpus product emerges with erroneous information. Some of the examples which had to be discussed by the team working on the corpus are given under (11).

- (11) *Jona akamedza hove mugungwa.* (Jonah swallowed fish in the sea.)
Pamadhigirii etriangle ari 25 (of the triangle's total of 25 degrees)

There may be a quick conclusion that this is obviously wrong information that has to be corrected. It may be easy to return to the original text which in these cases will be the cassettes on which the interviews were recorded to verify whether the text really represents what was said by the interviewee. The principle in corpus processing is to be faithful to the original text of the creator. It should be kept in mind that the misinformation might not have been a deliberate but an unintended mistake. When people do research they need to find factual material from the sources they use. The team agreed that these statements had to be left unaltered. It was felt there could one day be somebody with an interest to study these slips in language. The decision, though taken with good intentions, does not fully benefit the user. A special tag could have been developed to alert the user. Why this was not very problematic to the ALRI team is that they use the corpus in dictionary making. There would not be problems in selecting headwords from such statements. For broader research such statements would need to be marked to indicate that they contain some factual errors.

Conclusion

Corpus building is a laborious task. This stems from the fact that the corpus has to be fully representative of the spoken language: of what was said and what was meant to be said. This presents challenges and new problems appear, so that the way of treating these has to be standardised. It is necessary to discuss these problems so as to develop a common approach. As the corpus continues to be processed, new tag marks have to be formulated. All these aspects need to be standardised for the researcher who fully utilises the corpus resource materials. The outstanding problems facing individuals in the corpus production line have to be discussed, agreed upon and standardised. As a result, there is the need for a manual serving as guide to both the processor and the user of the corpus. The ALRI team should consider all the aspects where there were divergent approaches. All the agreed solutions of problems should be combined and presented as front matter or indexed as corpus guide. This also gives the opportunity for users of the corpus to form their own opinion about certain aspects of the corpus.

The user of the corpus may have a better understanding of the language situation and the historical background of this situation. This will help the user to appreciate how various aspects of the texts were handled to make the corpus useful as a research tool. Even the contradictions by the processors of the corpus may be appreciated. In this way, there is room for input from the creators, the processors and the users that can help improve the corpus. Like any research resource the corpus should be analysed and criticised to create the possibility for its improvement.

Bibliography

- Aarts, J and W. Meijs (Eds.). 1984. *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi.
- Chimhundu, H. (Ed.). 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. (Ed.). 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Crystal, D. 1991. *A Dictionary of Linguistics and Phonetics*. Third Edition. Oxford: Blackwell.
- Doke, C.M. 1931. *Report on the Unification of the Shona Dialects Carried Out under the Auspices of the Government of Southern Rhodesia and the Carnegie Corporation*. Hertford: Stephen Austin.
- McEnery, T. and A. Wilson. 2001. *Corpus Linguistics: An Introduction*. Second Edition. Edinburgh: Edinburgh University Press.
- Sibanda, M.J. 1989. Early Foundations of African Nationalism. Banana, C.S. (Ed.). 1989. *Turmoil and Tenacity: Zimbabwe 1890–1990*: 25–49. Harare: The College Press.

The Microstructure of *Isichazamazwi SesiNdebele*

Mandlenkosi Maphosa, *Department of African Languages and Culture,
Midlands State University, Gweru, Zimbabwe (maphosam@msu.ac.zw /
phosaz_m@yahoo.co.uk)*

and

Dion Nkomo, *Department of Afrikaans and Dutch, Stellenbosch University,
Stellenbosch; and Multilingualism Education Project, Centre for Higher
Education Development, University of Cape Town,
Republic of South Africa (deeouf@yahoo.co.uk)*

Abstract: This article analyses the microstructure of *Isichazamazwi SesiNdebele*. The analysis takes place on two levels: the level of availability of information and the level of accessibility of the available information to dictionary users. Data derived from two outreach exercises carried out to ascertain users' perspectives on the dictionary and their competence in using the dictionary is also scrutinised while the notion of user-friendliness and the general principles of dictionary-making form the framework within which the analysis is done.

Keywords: DICTIONARY, DICTIONARY INFORMATION, DICTIONARY STRUCTURE, MICROSTRUCTURE, MACROSTRUCTURE, DICTIONARY CULTURE, REFERENCE NEEDS, REFERENCE SKILLS, ACCESSIBILITY, USER-FRIENDLINESS

Opsomming: Die mikrostruktuur van *Isichazamazwi SesiNdebele*. Hierdie artikel ontleed die mikrostruktuur van *Isichazamazwi SesiNdebele*. Die ontleding vind op twee vlakke plaas: die vlak van die beskikbaarheid van inligting en die vlak van die toeganklikheid van die beskikbare inligting vir woordeboekgebruikers. Gegewens verkry van twee uitreikoefeninge uitgevoer om gebruikers se opvattinge van die woordeboek en hul vaardigheid in die gebruik van die woordeboek vas te stel, is ook ondersoek, terwyl die begrip van gebruikersvriendelikheid en die algemene beginsels van woordeboekmaak die raamwerk vorm waarbinne die ontleding gedoen word.

Sleutelwoorde: WOORDEBOEK, WOORDEBOEKINLIGTING, WOORDEBOEKSTRUKTUUR, MIKROSTRUKTUUR, MAKROSTRUKTUUR, WOORDEBOEKKULTUUR, NASLAANBEHOEFTE, NASLAANVAARDIGHEDE, TOEGANKLIKHEID, GEBRUIKERSVRIENDELIKHEID

1. Introduction

This article analyses the microstructure of *Isichazamazwi SesiNdebele* (henceforth the ISN) to evaluate the accessibility levels of information and the user-friendliness of the dictionary in this respect. Svensén (1993: 210) describes the microstructure as "the structure of individual dictionary entries:† their various parts

and the mutual relationship of these including the typographical conventions used". It is one of the guide or access structures defined by Louw (1999: 108, 109) as "the set of structures that provides a framework within which the availability and accessibility of information can be evaluated". The others are the macrostructure and the mediostructure. Since the microstructure operates within a dictionary article, this evaluation is concerned with the provision, presentation and arrangement of data categories that constitute an entry.

The provision of information should satisfy users' needs. Gouws (1996: 100) notes that the aim of any lexicographic project should be the production of dictionaries which satisfy the needs of the target users of a specific language community. Hartmann (1996: 102) expresses a similar sentiment when he says that the ultimate reason for the existence of dictionaries is the users' need for information. In line with this view, this article looks at the provision of information through the ISN microstructure against the users' reference needs.

However, Svensén (1993: 10) adds another aspect to the above when he writes: "Not merely information needs of different categories of users are important. An equally important factor is their ability to find and make use of the information given." This then necessitates the investigation of the reference skills of the users. The article considers whether, in addition to the provision of various information categories, the microstructure of the ISN facilitates users' access to information.

Also of importance is the dictionary culture. The term *dictionary culture* according to Hartmann and James (1998), refers to: "The critical awareness of the value and limitations of dictionaries and other reference works in a particular community." The dictionary culture of a community gives information on both the users' reference needs and their reference skills. As such, it is also important to shed light on the dictionary culture of the Ndebele community which should be borne in mind when looking at the reference needs and reference skills of this community.

The analysis conducted in this article is based on two fieldwork exercises in the Ndebele-speaking communities: a feedback outreach exercise by the editorial team (Maphosa included) in 2002 and another by Nkomo in 2003. The ISN editors solicited users' perspectives on the dictionary, including its structure (Maphosa 2003), while Nkomo (2003) evaluated the accessibility of the dictionary, mainly focusing on the microstructure. The two exercises provided empirical data on the reference needs, reference skills and dictionary-using culture of the Ndebele community. In this respect the analysis of the ISN microstructure and the final assessment of the dictionary are performed in the light of dictionary research.

The article presents the ISN microstructure first. The reference needs of the Ndebele community, followed by their reference skills are then presented so that the ISN microstructure is judged in terms of the provision and accessibility of information. The dictionary-using culture of the Ndebele community is then presented in order to contextualise the assessment of the user-friendliness of the ISN in respect of its microstructure.

2. The ISN Microstructure

The structure of any dictionary, be it the microstructure, the macrostructure, the mediostructure or even the megastructure, is largely determined by, among other factors, the type of dictionary, its purpose and the target users. If this line of argument is followed, it is therefore necessary to understand the dictionary itself in accordance with these factors before any attempt at judging its structure is made. In the case of the ISN, its type, its purpose, and its main target users should be defined.

The ISN is a monolingual general-purpose dictionary and its target users are mainly secondary schools, tertiary institutions and the general Ndebele speakers. Coming after *A Practical Ndebele Dictionary*, a bilingual English–Ndebele dictionary by Pelling (1966), it is the first monolingual dictionary in Ndebele.

The ISN has a general microstructure. Béjoint (2000) identifies two distinct types of microstructure, a general and a specialised. A general microstructure gives a "complete" programme of information in each article, including at least an explanation of the meaning, generally with the help of a definition (Béjoint 2000: 38). In addition to definitions, a general microstructure also includes other information categories like parts of speech, pronunciation, examples, etymology, etc.

The above description suits the ISN microstructure as dictionary articles contain a number of information categories, including information on spelling, variants, type of speech, meaning, word usage (examples), and synonyms. It does not have a specialised microstructure since it is not restricted to and characterised by a single information category like etymology or synonyms. Table 1 presents the information categories that constitute the ISN microstructure and the lexicographic conventions used to present them. Appendix 1 is a page extracted from the dictionary to complement Table 1.

Table 1: ISN information categories and conventions used

Information Category	Presentation and Conventions used
Spelling	Headwords in bold print
Variants	After headwords, upper case and in [] and BONA
Type of speech	After headwords, through abbreviations, e.g. bz 5, sz mwa
Meaning	Definitions as full sentences, sense numbers used
Examples	After definitions and in italics
Synonyms	After explanations through abbreviation FAN

From Table 1 and Appendix 1 it may be noted that headwords are keys to dictionary articles. Bold print and the indentation of other information categories in the successive lines make them distinct and easy to identify. In addition, as

Landau (2001: 112) stresses, headwords are "indicators of the preferred spelling". Their variants, if any, are presented immediately after them, printed in upper case letters and enclosed in square brackets. Variants present alternative forms of the word with regard to spelling and are defined using the main articles.

Speech-type information is then given by means of abbreviations, with the dominant ones being *bz* for *ibizo* (noun) and *sz* for *isenzo* (verb) as the majority of dictionary lemmata fall within these two word categories. Nouns are then marked for class, e.g. *bz 5* for nouns falling in class 5 following the Bantu noun classification system by Meinhof in terms of prefixes, while verbs are marked for transitivity and intransitivity using the abbreviations *mwa* and *gmwa* for *isenzo esilomenziwa* and *esingelamenziwa* respectively.

So far the presentation of formal and syntagmatic information categories has been described. The remaining types are basically semantic although others tend to overlap. Definitions are the most important of all semantic categories. In the ISN, definitions are given as full sentences with the headword being used in the definition. This follows one of the defining principles of the Collins Birmingham University International Language Database (COBUILD), whose main objective is to make definitions as comprehensive as possible. When a word carries at least two senses, a definition is provided for each sense with the primary sense presented first. In such instances the definitions are distinguished by sense numbers, with 1 for the primary sense, 2 for the second sense, etc. In some instances, definitions in the ISN are elucidated by means of examples. Printed in italics, examples illustrate the usage of the word, that is, its combination with other words in speech. Then lastly, synonyms are given after the abbreviation FAN for *amagama alengcazelo efanayo*. For space economy, lesser synonyms are defined by major synonyms.

3. ISN User Perspectives

User perspective considers a lexicographic work from the point of view of the user. In this case, focus is on the users' reference needs, the users' reference skills, and the dictionary culture which determines users' reference needs and reference skills.

3.1 The ISN Users' Needs

According to Hartmann (1983: 5), "all dictionaries are motivated by and judged against the lexical needs of the language user whom they serve". An analysis of the microstructure should therefore look at the provided information categories relative to users' needs.

The ISN editorial team managed to solicit the ISN users' needs from Ndebele dictionaries. According to Maphosa (2003: 6), "users were given a number of information categories and ... asked to state whether they had need for them

in dictionaries or not". Some information categories investigated are not provided in the ISN, but were included so as not to limit users to what is found in the dictionary, thereby giving them the possibility to express their needs, if at all, for the excluded categories. Furthermore, some of the information categories omitted from the ISN need to be considered for the *Advanced Ndebele Dictionary*. The need for information was also investigated with regard to English dictionaries. Since English dictionaries seem to be an integral part of the dictionary-using culture in the Ndebele community, this was done with the expectation that the findings could reveal interesting trends in users' needs. The influence of English dictionaries on users' needs and reference skills should not be underestimated. Table 2 sums up the ISN users' lexicographic needs obtained by the editorial team through the outreach exercise.

Table 2: Comparative figures on users' needs in Ndebele and English dictionaries.

Information categories	Total number of informants	Ndebele dictionaries		English dictionaries	
		+	-	+	-
Meaning	1250	425	825	1250	0
Spelling	1250	1100	150	1250	0
Examples	1250	1150	100	1196	54
Speech type	1250	850	400	1212	38
Synonyms	1250	775	475	1110	140
Variants	1250	900	350	1150	100
Etymology	1250	980	270	1035	215
Pronunciation	1250	455	795	1072	178
Tone	1250	320	930	220	1030

Key: + need for information category
 - no need for information category

A comparison of Table 1 on information categories provided in the ISN and Table 2 on the ISN reference needs shows that the ISN is deficient in meeting the reference needs of the users. Svensén (1993) is of the view that decisions on lexical treatment of lemmata in terms of provision and presentation of information should be in the order of the preference of the intended users. However, as the ISN is the first and so far the only monolingual Ndebele dictionary, this could not have served as a guide for the ISN editorial team when the dictionary was compiled. The dictionary was largely shaped, according to the description of Béjoint (2000: 170), by "what lexicographers knew, imagined, of those needs, and by what lexicographers knew about the language in general and words in particular". The comment on whether or not the editors' decisions were justified shall therefore be reserved until the genuineness of the reference needs of the ISN users is confirmed in view of the Ndebele dictionary culture.

An analysis of the findings of the ISN editorial team on users' needs presents a negative view of the Ndebele dictionary culture. The target users seem to be unaware of the value of dictionaries in Ndebele as becomes evident from the lower rates of the need for information categories compared to those from English dictionaries, with the worst being the need for meaning which is rated at only 36%.

Dictionaries are largely consulted for the meaning of words, and if the target users of a dictionary do not recognise the value of this important information category, then it means that they are unaware of the value of the dictionary itself. Interesting though is that the very same users, who claim not to need meaning from the Ndebele dictionary indicate a higher need for other information categories such as spelling, speech type, etymology, variants, synonyms and examples. Except in specialised dictionaries with specialised microstructures, these information categories cannot be separated from the meanings of words. In most cases, some of them complement the explanation of the words given by definitions. For instance, examples of word usage usually elucidate the meaning of a word and therefore a user who searches for the meaning of the word, will need to learn it from the examples of word usage.

A higher need for examples of word usage somehow reflects a genuine, but unacknowledged need for word meanings by ISN users. In respect of this view, it may be noted that these claims are merely a manifestation of attitude problems characteristic of a community with a young dictionary culture, which holds that people know their language and need no explanation of it. When Nkomo (2003) undertook a control test that involved assigning students in four Bulawayo high schools selected words to define, this attitude surfaced among them. The number of students who acknowledged the need for meaning from Ndebele dictionaries rose as many who had denied this need failed the test. Hadebe (2004: 98-99) explains the detrimental impact of attitude problems on reference needs, dictionary culture and reference skills. The users' perspectives on the information provided by the ISN, and their needs should therefore be treated with caution.

3.2 The ISN Users' Reference Skills

Emphasising the importance of users' reference skills, Svensén (1993: 10) notes that the form of a dictionary should be greatly determined, among other factors, by the users' capabilities to decode the provided information. In addition to the investigation of users' needs, the reference skills of the ISN users were also investigated during the two field-work exercises. The participant-observation method was employed in both outreach exercises to discover the reference skills of students at secondary schools. The method was effective in affording the researchers an opportunity to observe the skills of users engaged in actually consulting the dictionary. The participants and informants were given the task to decode the information categories in the dictionary and interpret some of the

conventions. The findings tabulated below are those of the ISN editors since they represent a bigger sample than those obtained from the schools.

Table 3: Users' capabilities to interpret conventions and decode information from the ISN.

Convention	+	-
Bold print	1200	50
Italics	987	263
Abbreviation bz	933	317
Abbreviation sz	918	332
FAN	1100	150
BONA	1210	40
Square brackets	775	475
Sense numbers	900	350

Key: + correct interpretation of the convention
 - wrong interpretation of the convention or no response

When the table is studied, it is safe to say that the ISN users possess an above average of the required reference skills in using the dictionary, although it is the first monolingual Ndebele dictionary, which they consulted. This may be attributed to the fact that users have at some time used English, and perhaps a few Zulu dictionaries. At this point, users' reference skills are not a serious problem in Ndebele dictionary use. Ilson (1984: 83) observes that "a dictionary is a social artefact ... whose users become familiar with its conventions unconsciously, just as lexicographers do". This is true of the ISN and its users who exhibited no difficulty in using the ISN in spite of it being the first monolingual Ndebele dictionary.

3.4 Dictionary Culture in Ndebele

By developing a dictionary culture as noted by Hartmann and James (1998), a community makes a considerable adjustment to lexicography. This means that the community acknowledges a gap in its linguistic knowledge and appreciates lexicography as a tool that may be exploited to fill this gap. This does not necessarily mean that the community waits for lexicographers to produce dictionaries. Rather, the community follows the dictionary-making trends closely, making sure that lexicographers produce good dictionaries. In a community with an advanced dictionary culture, users would criticise dictionaries and even reject some if they do not meet their lexicographic needs.

Svensén (1993) notes that the linguistic gap alluded to above, which necessitates the use of dictionaries is realised when writing, reading, speaking and/

or listening. These are daily activities which involve the use of language. What logically follows is that dictionaries have a vital role in peoples' everyday lives. In spite of this, a dictionary-using culture in the Ndebele community is still in its infancy. The community has not made a considerable adjustment to lexicography. Perhaps this is due to the fact that currently only three Ndebele dictionaries are available for use, namely Pelling's (1966) *A Practical Ndebele Dictionary*, which is bilingual, Hadebe's (2001) *Isichazamazwi SesiNdebele*, which is the only monolingual Ndebele dictionary, and Nkomo and Moyo's (2006) *Isichazamazwi SezoMculo*, which is a specialised Ndebele dictionary of musical terms.

One aspect relating to dictionary culture is dictionary ownership. Nkomo (2003) held eight discussions with students, with each discussion involving at least 25 Ndebele students at four high schools in Bulawayo. During the discussions, it emerged that only two out of at least two hundred students possessed a dictionary (none of these being the ISN). The rest of the students said they borrow dictionaries from school libraries, community libraries, friends or relatives. However, out of the four schools, only three had at least one ISN copy in their libraries although the majority of the students had not used it. The implications of this is that either the schools were unaware of the availability of the dictionary at bookshops two years after its publication or were not aware of the value of a Ndebele dictionary, hence their reluctance to buy it. This lack of awareness is therefore passed on to students and impacts negatively on the development of a dictionary culture. Interestingly, the schools were chosen on the basis that they produce the best Ndebele results at both ordinary and advanced levels annually. Therefore great awareness of language development activities in such schools would be expected.

Although the Ndebele community is not accustomed to Ndebele dictionaries, it has nevertheless used English dictionaries. Secondary school students could even name some English dictionaries they have consulted. The higher need for most of the information categories from English dictionaries as shown in Table 2 is part of the dictionary culture of the Ndebele community. However, the frequency of using dictionaries is very low, with most of the users consulting dictionaries only on rare occasions.

The above findings serve to confirm the infancy of a dictionary-using culture in the Ndebele community. The ISN users' reference needs and reference skills are a result of this culture. Whilst the method of measuring the reference skills might have been accurate, it should be noted that, with this dictionary culture, the reference needs could not be relied on either in compiling the dictionary or in passing a judgement on the provision of information in the dictionary. The reference needs should be treated with caution until the dictionary culture has been improved. Therefore a dictionary-using culture is a critical problem area for Ndebele lexicography, of which every form of dictionary criticism should be wary. In the light of this, the accessibility of information by means of the ISN microstructure, hence its user-friendliness, is evaluated in the next section.

4. User-friendliness of the ISN Microstructure

As stated earlier in the article, the ISN microstructure is evaluated on two levels. The first is that of meeting the reference needs of users by providing much needed information on words. The second is that of employing symbols to facilitate users' access to information without estranging them. The evaluation considers both the user perspective and the compiler perspective. Dictionaries are not only compiled according to the user perspective, but in accordance with lexicographic principles guiding this compilation. Therefore, while the ISN editors' efforts of meeting their target users' needs are investigated, their consistency in applying the principles of lexicography is not neglected.

Table 2 shows that only the last three information categories investigated are not provided in the ISN. As far as the excluded information is concerned, it should be remembered that no one dictionary is exhaustive of all aspects of a language. In the same manner, the ISN as a general dictionary could not satisfy all its target users, especially those whose needs were unknown, or could not be relied upon even if they were known because they are lexicographically unschooled and ill-informed.

The dictionary also reflects the trends in language study at the time of its compilation. At present, the study of Ndebele grammar has not reached advanced levels of analysis. For instance, mainly introductory courses of Ndebele grammar are offered at the University of Zimbabwe, Harare. Perhaps after a serious needs analysis, an advanced dictionary should provide some of the excluded information types. However, at the moment, the needs analysis for information shows that, with the improvement of users' awareness of the value of dictionaries, the dictionary can serve as a vital reference tool in Ndebele.

After deciding on the information the dictionary is to provide, the lexicographer further determines the conventions that will help him/her in presenting the information and at the same time help the user in accessing it. In section 3.2, it was noted that the ISN users displayed above average skills in decoding information from the dictionary. What this means is that, while the conventions enabled the editors to save space, they are explicit enough to guide users to the various information categories the dictionary provides.

In addition to expecting users to learn how to use the dictionary on their own, lexicographers usually give guidelines in the front matter on how the microstructure presents information and directs users. The microstructure cannot be fully studied and be comprehensively understood in isolation (Nkomo 2003: 5). In the same way, users may need to refer to the front matter to obtain instructions with regard to the microstructure since it does not function in isolation; it is only part of the guide or access structures. To provide users with guidelines, the ISN editors explain the microstructure in the front matter which is detailed enough for easy comprehension. Hadebe et al. (2001: xl) note the need for guidance by saying:

Kuqakathekile kosebenzisa isichazamazwi ukuthi alandele leziziqondiso ukuze athole lokho ayabe ekudinga. Lapha sichaza ngamafitshane impawu ezihamba legamanhloko.

(It is important to the dictionary user to follow these guidelines so that it becomes easy to access the required information type. Here we briefly explain the markers accompanying each headword.)

However, it may be argued that information must be presented in a simple manner rather than expecting the user to read the front matter or interpret the keys. The provision of the front matter has always been a good lexicographic principle although dictionary users tend to ignore it. For instance, investigating the dictionary-using culture of the ISN users, Nkomo (2003: 53) discovered that the majority of students, who are the main target users of the ISN do not read the front matter. While this serves to emphasise the need for simplicity in the presentation of dictionary information, especially for users in a community where a dictionary-using culture is still in its infancy, at the same time, it calls for user education with regard to lexicography in general and other aspects of lexicography like, among others, dictionary structure in particular. Hadebe (2004: 89-104) reiterates the importance of teaching dictionary skills as a way of developing a dictionary-using culture in Ndebele and other communities who are in a similar situation.

The presentation of the information and the use of conventions should be done consistently for related headwords. This is accentuated by Zgusta (1971: 248) who points out that "all entries should be constructed in as uniform a way as possible". Consistency is important, for the user has to master the microstructure from the start to continue using the dictionary with efficiency. Such a structure, the microstructure in particular, in this way becomes effective for educating its users on how to use the dictionary. A closer look at the microstructure of the ISN (see Appendix 1) shows that the ISN editors have used the conventions in presenting dictionary information consistently. This is likely to be a score on the lexicographers' attempt to achieve more user-friendliness.

After the ISN users' needs, reference skills and the dictionary-using culture which informs them have been identified, a fair assessment of the dictionary in terms of accessibility and user-friendliness can be made. Firstly, it should be remembered that the ISN is a pioneering work in Ndebele monolingual lexicography whose shape was not dictated by the real users' needs, but "assumptions by dictionary editors" (Hadebe 2004: 90). Even the needs on which this article is based are questionable since at the time the research was undertaken (2002 and 2003), the majority of the users had not seen, let alone used the dictionary. In this respect, attention must be shifted from the provision of information and the satisfaction of users' needs to the accessibility of the provided information.

Here it should be remembered that accessibility of information does not necessarily mean a simple microstructure with easily interpretable conventions. According to Béjoint (2000: 157), the skills necessary for the successful consultation of a dictionary are only those of the average educated language

user. When this view is accepted, it follows that if average secondary school Ndebele students could use the dictionary upon seeing it for the first time, it is accessible and user-friendly. Béjoint further compares the process of learning to use a dictionary to learning to read, a point which somehow emphasises the need for user-education rather than producing a dictionary outside the principles of dictionary-making. Children learn to read by receiving help or support from their educators. When the ISN microstructure is considered in relation to the ISN users' reference skills, an improved dictionary-using culture and practice will hopefully help and support more users to be more comfortable with the dictionary.

As a pioneering reference work in Ndebele monolingual lexicography, the ISN can even educate its users lexicographically. The consistency with which the microstructure presents information contributes towards users becoming used to the dictionary. In addition, the provision of a detailed and lexicographically educative front matter for a pioneering monolingual dictionary does not only help Ndebele language users in the consultation of the ISN, but also in introducing them to Ndebele lexicography. The conventions used in the ISN will be interpreted and used with greater ease and efficiency for the access of information in future Ndebele dictionaries.

5. Conclusion

The focus of this article has been on the ISN microstructure, assessing the accessibility of dictionary information and the user-friendliness of the dictionary. The assessment is based on the ISN users' reference needs in relation to the information types provided in the ISN and the ISN users' reference skills in relation to the lexicographic conventions used in presenting the provided information types. It emerges that a fair assessment could not be done without taking cognisance of the dictionary-using culture of the Ndebele community since it dictates both the users' reference needs and their reference skills. The dictionary-using culture of the community seems to be the most critical point that needs attention so that users can access the provided information in the ISN. What is vital is that they can fully utilise the information and hence see the value of a monolingual dictionary in their language. As for the provided information, the microstructure facilitates its access since, through it, users can reach the searched-for information types without any training. The consistent presentation of information is helpful for users who are not fully competent in using a monolingual dictionary in their own language, but with practice they are likely to learn quickly. This consistency is complemented by a detailed front matter which gives, among other aspects, a comprehensive explanation of the microstructure. In this regard, it can be concluded that the dictionary information is accessible and the dictionary itself is user-friendly, but the Ndebele community is not lexicographically well-cultured.

Note

† In accordance with Gouws and Prinsloo (2005: 15), among other metalexigraphers, *article/s* instead of *entry/-ies* will be used throughout this evaluation.

References

- Béjoint, H.** 2000. *Modern Lexicography: An Introduction*. Oxford/New York: Oxford University Press.
- Gouws, R.H.** 1996. A Sequence for Meeting Lexicographic Needs. *Lexicography as a Financial Asset in Multilingual South Africa*: 97-110. Pretoria: Department of Arts, Culture, Science and Technology.
- Gouws, R.H. and D.J. Prinsloo.** 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PreSS.
- Hadebe, S.** 2004. Improving Dictionary Skills in Ndebele. *Lexikos* 14: 89-104.
- Hadebe, S. et al. (Eds.).** 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hartmann, R.R.K.** 1983. On Theory and Practice: Theory and Practice in Dictionary-making. Hartmann, R.R.K. (Ed.). 1983. *Lexicography: Principles and Practice*: 3-11. London: Academic Press.
- Hartmann, R.R.K.** 1996. Lexicography as an Applied Linguistic Discipline. Hartmann, R.R.K. (Ed.). 1996. *Solving Language Problems: From General to Applied Linguistics*. Exeter: Exeter University Press.
- Hartmann, R.R.K and G. James.** 1998. *Dictionary of Lexicography*. London: Routledge.
- Ilson, R.F.** 1984. The Communicative Significance of Some Lexicographic Conventions. Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*: 80-86. Tübingen: Max Niemeyer.
- Landau, S.I.** 2001. *Dictionaries. The Art and Craft of Lexicography*. Second Edition. New York/Cambridge: Cambridge University Press.
- Louw, P.A.** 1999. Access Structures in a Standard Translation Dictionary. *Lexikos* 9: 108-118.
- Maphosa, M.** 2003. The Users' Perspectives on *Isichazamazwi SesiNdebele*. Paper presented at the Eighth International Conference of the African Association for Lexicography, organised by the Department of Germanic and Romance Languages, University of Namibia, Windhoek, Namibia, 7-9 July 2003.
- Nkomo, D.** 2003. *The Microstructure and the Accessibility of Dictionary Information: An Analysis of Isichazamazwi SesiNdebele*. Unpublished B.A. Hons. Thesis. Harare: University of Zimbabwe.
- Nkomo, D. and N. Moyo (Eds.).** 2006. *Isichazamazwi SezoMculo*. Gweru: Mambo Press.
- Pelling, J.N.** 1966. *A Practical Ndebele Dictionary*. Harare: Longman Zimbabwe.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- Zgusta, L.** 1971. *Manual of Lexicography*. The Hague: Mouton.

Appendix 1: Page 124 of the ISN showing the structure of entries

iKhalanga	ikhasitadi
iKhalanga bz 5. iKhalanga ngumuntu wosendo lomhlobo wamaKhalanga okhuluma isiKhalanga njengolimi lwakhe lomdabuko.	ikhaphophi bz 5. <i>Ikhaphophi</i> lizenge elilibhokisi elembeswa ngemuva kwemota eyisigengele nxa kuvalelwa izulu, umoya kumbe ilanga.
ikhathali bz 5. BONA <i>ikhathali</i> .	ikhansili bz 5. <i>Ikhansili</i> ngabantu abakhethiweyo esigabeni kumbe endaweni abayabe bekhangele inhlupo kanye lentuthuko yaleyo ndawo.
ikhallenga bz 5. <i>Ikhallenga</i> ngumbhalo otshengisa amalanga asenyangeni, lamaviki akhona njengokulandelana kwawo kulowo mnyaka.	ikhansi bz 5. BONA <i>ibhodo</i> .
ikhalkhuletha bz 5. <i>Ikhalkhuletha</i> ngumtshina omncane osetshenziwa ekuncediseni abantu emisebenzini yezinombolo.	ikhaphakhapha bz 5. <i>Ikhaphakhapha</i> yinto enkulu njalo ebanzi.
ikhamanzi bz 5. BONA <i>inkezo</i> .	ikhaphaphe bz 5. BONA <i>ikhaphakhapha</i> .
ikhambi bz 5. <i>Ikhambi</i> ngumuthi wokwelapha osetshenziwa ngabantu empilweni zabo zansuku zonke ongadingi sazi ukuthi uwuthole.	ikhaphelo bz 5. BONA <i>idlelo</i> .
ikhambi bz 5. <i>Ikhambi</i> yisifo somzimba esiqubula umzimba ikakhulu ubuso.	ikhaphethi bz 5. <i>Ikhaphethi</i> yimethi enziwe ngentambo eziqinileyo eziconjwe zaba ngumbala okhangayo.
ikhambi bz 5. <i>Ikhambi</i> licebo lokulungisa okonakeleyo.	ikhaphithalizimu bz 5. <i>Ikhaphithalizimu</i> yindlela yempilo lapho umnotho oyabe usezandleni zabalutshwane abazama ukwenza inzuzo enengi kakhulu.
ikhamera bz 5. <i>Ikhamera</i> ngumtshina wokuthatha amapikitsha.	ikharadi bz 5. <i>Ikaradi</i> ngumuntu ozalwa ngabantu bembala eyehlukeneyo njengekhiwa lomuntu omnyama.
ikhampani bz 5. <i>Ikhampani</i> liqembu labantu elisebenza ndawonye likhangele ngezamabhezimisi.	ikharathi [ikhathali] bz 5. <i>Ikarathi</i> ngumdlobo wobuciko bokukhabana lokuzivikela ngezandla.
ikhampasi bz 5. <i>Ikhampasi</i> yithuluzi esetshenziwa ukudweba ikakhulu ezifundweni zenombolo.	ikhasethi bz 5. <i>Ikhasethi</i> ngumcephe omncane olentambo lamavili amabili odlalwa eredyweni.
ikhampasi bz 5. <i>Ikhampasi</i> ngumtshina okhombisa umuntu amagumbi omhlaba.	ikhasi bz 5. 1 <i>Ikhasi</i> yisembeso esithelo kumbe esesilimo. 2 <i>Ikhasi</i> licele elilodwa elephepha egwalweni.
ikhanda bz 5. <i>Ikhanda</i> yisitho esiyinhloko yomuntu.	ikhasi bz 5. <i>Ikhasi</i> lihlamvu elomileyo.
ikhandlela bz 5. <i>Ikhandlela</i> yisibane esingamafutha anqumileyo alentambo phakathi.	ikhasi lenyoka bz 5. BONA <i>inwebu</i> .
ikhangaru bz 5. <i>Ikhangaru</i> yinyamazana etholakala ikakhulu kwele Australia edume ngesikhwama sayo esingaphansi kwesisu ethwala ngaso abantwabayo.	ikhasi lenyoka bz 5. <i>Ikhasi lenyoka</i> yisikhumba senyoka esiphuma lapho inyoka ihluba.
ikhangarukhothi bz 5. <i>Ikhangarukhothi</i> ngumthethwandaba ongekho emthethweni ongenziwa yizibotshwa ejele, inhlanganiso yezisebenzi kumbe loba yiphi inhlanganiso ukuthonisisa elinye lamalunga alo eliyabe liphambanisile.	ikhasili bz 5. <i>Ikhasili</i> ngumhlobo wotshwala bamambodlela.
ikhango bz 5. <i>Ikhango</i> yinkomitsho enkulu.	ikhasimende bz 5. <i>Ikhasimende</i> ngumuntu ohlezi ethenga impahla yakhe emuntwini othile kumbe esithole esithile aze azakale kulesositolo.
ikhanka bz 5. <i>Ikhanka</i> yinyamazana yeganga elingana lenja edla izifuyo ezinjengembuzi.	ikhasino bz 5. <i>Ikhasino</i> yindawo okuyonyelwa khona uzulu ukuthi abheje ngemali emitshineni lokuthi azithokozise.
	ikhasitadi bz 5. <i>Ikhasitadi</i> yimputshana elithanga okuphekwa ngayo ilambazi lochago.

Not Mere Lexicographic Cosmetics: The Compilation and Structural Features of *Isichamazwi SezoMculo*

Nobuhle Moyo, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe (noxy_auree@yahoo.co.uk)*

and

Dion Nkomo, *Multilingualism Education Project, Centre for Higher Education Development, University of Cape Town; and Department of Afrikaans and Dutch, Stellenbosch University, Stellenbosch, Republic of South Africa (deeouf@yahoo.co.uk)*

Abstract: This article offers a brief overview of the compilation of the Ndebele music terms dictionary, *Isichamazwi SezoMculo* (henceforth the ISM), paying particular attention to its structural features. It emphasises that the reference needs of the users as well as their reference skills should be given a determining role in all lexicographic decisions leading to the publication of a dictionary. Dictionary structure should, therefore, be conceived and evaluated in terms of its data constituents and the accessibility of these data. Accordingly, this article demonstrates that the structure of the ISM is not a case of mere cosmetics but a lexicographic mode of communication between the dictionary compilers and users who are participants in a lexicographic communication process. In this way, the article highlights some of the challenges encountered during the compilation of the ISM and the strategies the compilers employed to facilitate the communication process between the lexicographers and the users regarding dictionary contents and the arrangement thereof. From such a perspective, this article may provide useful insights for LSP lexicography in African languages, prospects of which are based on the increased need for knowledge acquisition and dissemination as well as the multilingual nature of African societies.

Keywords: LSP LEXICOGRAPHY, DICTIONARY STRUCTURE, MEGASTRUCTURE, MACROSTRUCTURE, MICROSTRUCTURE, MEDIOSTRUCTURE, REFERENCE NEEDS, REFERENCE SKILLS, OUTER TEXTS, LEMMATA, DICTIONARY ARTICLE

Opsomming: Nie slegs leksikografiese kosmetiek nie: Die samestelling en strukturele eienskappe van *Isichamazwi SezoMculo*. Hierdie artikel bied 'n kort oorsig van die samestelling van die Ndebelewoordeboek van musiekterme *Isichamazwi SezoMculo* (voortaan die ISM) deur besondere aandag te gee aan sy strukturele eienskappe. Dit beklemtoon dat die naslaanbehoefte van die gebruikers sowel as hulle naslaanvaardighede 'n bepalende rol gegee behoort te word in alle leksikografiese besluite wat tot die publikasie van 'n woordeboek lei. Woordeboekstruktuur behoort derhalwe beplan en beoordeel te word ooreenkomstig sy databestanddele en die toeganklikheid van hierdie data. Hierdie artikel toon gevolglik dat die struktuur van die ISM nie net 'n geval van blote kosmetiek is nie, maar 'n leksikografiese kommunikasiewyse tussen die woordeboeksamstellers en -gebruikers wat deelnemers is in 'n leksikografiese kommunikasieproses. Op hierdie manier belig die artikel 'n aantal van die uitdagings wat teenge-

kom is tydens die samestelling van die ISM en die strategieë wat die samestellers gebruik het om die kommunikasieproses tussen die leksikograwe en die gebruikers betreffende woordeboekinhoud en die ordening daarvan te vergemaklik. Vanuit so 'n perspektief kan hierdie artikel nuttige insigte verskaf in TSD-leksikografie in Afrikatale, vooruitsigte wat gebaseer is op die toenemende behoefte aan kennisverwerwing en -verspreiding sowel as die veeltalige aard van Afrikagemeenskappe.

Sleutelwoorde: TSD-LEKSIKOGRAFIE, WOORDEBOEKSTRUKTUUR, MEGA-STRUKTUUR, MAKROSTRUKTUUR, MIKROSTRUKTUUR, MEDIOSTRUKTUUR, NASLAANBEHOEF- TES, NASLAANVAARDIGHEDE, BUITETEKSTE, LEMMAS, WOORDBOEKARTIKEL

1. Introduction

This article offers a general overview of the compilation of a Ndebele dictionary of music terms, *Isichazamazwi SezoMculo* (henceforth the ISM), but its main focus is on the structure of this dictionary. Dictionary structure is seen as a system of communication between the compilers and users. Such an approach is critical in modern-day lexicography. To start with, the purpose of any lexicographic project should be to satisfy the needs of the target users (Gouws 1996: 100), since "all dictionaries are motivated by and judged against the lexical needs of the language user whom they serve" (Hartmann 1983: 5). The lexicographer may claim to have satisfied the needs of his/her users by providing the relevant data categories required to address the problems which initiate dictionary consultation. However, the following should be taken into account:

Not merely information needs of different categories of users are important. An equally important factor is their ability to find and make use of the given information (Svensén 1993: 10).

The data on offer in the dictionary should be presented in such a way that the knowledgeable target user can access it in order to retrieve the kind of information that motivated a given dictionary consultation procedure (Gouws and Prinsloo 2005: 165).

A well-formulated dictionary structure is the only way of ensuring that the users receive the maximum assistance from a dictionary. Thus, Yong and Peng (2007: 83) assert that:

Effective lexicographic communication requires the communicative system and its subsystems to be coordinated and structured in such ways as to facilitate interaction between the compiler and the user and to enhance ready access to different parts of the dictionary where the required information is located.

At the planning stages of dictionary production, the lexicographer needs to focus on the reference needs and reference skills of the target users by constructing a user profile (Bergenholtz and Tarp 1995, 2003; Tarp 2000, 2002, 2004). Although the authors of this article cannot give a detailed account of

such planning regarding the ISM since they joined the project after its planning stages, they hope to show convincingly that the procedures they adopted after taking over the project, especially regarding the structural aspects of the dictionary, were motivated by the intention to satisfy the needs of the target users. While the decisions of the editors are largely reflected in the ISM, it should be noted that the style guide drawn up by the initiators of the project was always the point of departure. Some members of the initial editorial team and the reference working group also continued to assist in various ways until the publication of the dictionary.

Some of the challenges related to the typological nature of the ISM are discussed in Section 2. This will hopefully present some insights into specialised lexicography (henceforth LSP lexicography) especially for African language communities. The increased growth of knowledge in various subject fields owing to global technological developments as well as the need to nurture multilingualism in countries such as South Africa increases prospects for LSP lexicography in African languages. Together with the general lexicographic principles, the experiences of other LSP lexicographic projects such as the ISM may provide the necessary guidance. Section 3 discusses the frame structure, also called the megastructure, of the dictionary. The front-matter components of the ISM are subsequently described in Section 4. This is followed in Section 5 by the discussion of the macrostructure of the dictionary. The microstructure is then described in Section 6. The back matter is the subject of Section 7. This is followed by a description of the mediostructure in Section 8, the final one before the conclusion.

2. Some Challenges of a Typological Nature in the Compilation of the ISM

The ISM is the first specialised dictionary (henceforth LSP dictionary) in Ndebele. It appeared in 2006, five years after the publication of the first monolingual Ndebele dictionary *Isichazamazwi SesiNdebele* (henceforth the ISN). Like the ISN, it was also published under the auspices of the African Languages Lexical (ALLEX) Project which led to the establishment of the African Languages Research Institute (ALRI). Unlike the ISN which attempts a general lexicographic coverage and treatment of the Ndebele vocabulary, the ISM covers the music domain in Ndebele. Its target-user groups are mainly music students, lecturers, teachers and other practitioners who need to interpret and understand music genres, instruments and other music concepts in Ndebele. Such a group of users generally shares a low level of dictionary culture with the entire Ndebele community. To most of these users, the ISM arguably became the first Ndebele dictionary to which they were introduced as they have remained unacquainted with the ISN published five years earlier. This is an indication of a poor dictionary culture, which is not peculiar to Ndebele, but generally prevalent in many African language communities.

The very fact that the ISM is an LSP dictionary required special knowledge of the music domain. The editors, including their predecessors in the ISM project, did not possess such expert knowledge adequate for lexicographic purposes. The production of LSP dictionaries requires the combined effort of specialists in the specific subject field as well as lexicographic expertise (Bergenholtz and Tarp 1995, 2003; Tarp 2000, 2002). Although the ISM compilers possessed adequate lexicographic training, they had to rely on experts in the music industry and music education to produce a dictionary of substantial quality. They benefited from the expertise of a reference working group which comprised Ndebele-speaking music lecturers, music practitioners, linguists and individuals who, one way or another, are involved in music teaching and linguistic research.

The dictionary has a trilingual dimension. While Ndebele is the lemmatising and explicating language, Shona and English translation equivalents are provided. As a result, the ISM is generally, though erroneously, regarded as a trilingual dictionary. Scholars such as Burkhanov (2004), Gouws (2004a) and Hartmann and James (1998: 14) have argued in a convincing way that the provision of translation equivalents in a dictionary does not typologically make that dictionary bi- or multilingual. The main argument is that not every dictionary which provides translation equivalents may serve the typical functions of bi- or multilingual dictionaries. Nevertheless, this trilingual dimension of the ISM posed a number of challenges during its production.

The evident challenges encountered in the compilation of the ISM in respect of the trilingual dimension related to the issue of translation equivalents. This affected the formulation of the dictionary basis at the initial compilation stages. Firstly, the dictionary had to include terms representing concepts from traditional Ndebele music which continues to influence contemporary Ndebele musicians. Secondly, terms representing musical concepts which some Ndebele musicians have adopted from Shona traditional music also deserved inclusion. Finally, the dictionary had to include music terms representing concepts, be it of English, Italian, French and German origin, which are now central in the music domain. Either in the Ndebele term creation for the dictionary (Nkomo 2005) or the search for English and Shona translation equivalents, challenges were encountered because, as far as LSP lexicography is concerned, music is a highly culture-dependent domain (Bergenholtz and Tarp 1995). Thus, cultural knowledge was required in many instances as well as specialised music knowledge and linguistic competence to facilitate translation. It has already been indicated that music experts were helpful regarding specialised knowledge, but the editors also needed to develop competence to grasp and process that knowledge for lexicographic purposes. Linguistic competence in English and Shona was thus another important factor.

As far as English translation equivalents were concerned, the editors relied on their competence in English which is the main language of education in Zimbabwe. Furthermore, general monolingual, bilingual and music dictionar-

ies in English were amply available for consultation throughout the compilation stage. Music experts who were members of the reference working group also evaluated the English translation equivalents to ensure that content was not seriously lost during translation.

Translation between Ndebele and Shona would be equally difficult as it appeared that there are certain musical concepts in Shona not found in Ndebele and vice versa. The editors' linguistic competence in Shona was also too limited for lexicographic purposes. Unlike in English, there were no appropriate dictionaries that could give assistance. Thus, the ISM editors could not determine the correct translation equivalents themselves. In fact, determining translation equivalence between Ndebele and Shona was more difficult than between Ndebele and English despite the fact that cultural distance is greater between the latter two. This meant that the editors had to rely heavily on their Shona-speaking colleagues at ALRI. Conveniently, the editors of *Duramazwi reMimhanzi* (Mheta 2005), the Shona music terms dictionary, from their side needed similar assistance throughout the compilation of their dictionary. Nevertheless, the manifestation of the cultural factor and linguistic anisomorphism in the ISM remain evident as will be shown in Sections 5, 6 and 7.

Another matter which the editors had to contend with during and even after the compilation of the ISM pertains to the attitude of the users towards other languages. The attitude of Ndebele speakers towards other languages in Zimbabwe are discussed by Hadebe (2005) with reference to the prospects of a bilingual Ndebele–Shona/Shona–Ndebele dictionary and also by Ncube (2005) regarding loan words in the ISN. Both writers rightly indicate that these attitudes result from certain historical, economic and socio-political factors. Although they cannot be overstated, they surfaced during and after the production of the ISM. Some users, including some music practitioners who were members of the reference working group questioned the inclusion of translation equivalents, especially in Shona. They argued that a purely monolingual music dictionary should be compiled, especially following a major breakthrough in mother-tongue lexicography inaugurated by the successful publication of the ISN. Yet others questioned the manner in which transliteration was used to create Ndebele terms for some Shona and English music concepts. While lexicographic planning needs to be cognisant of such factors, it should ensure that appropriate and user-friendly dictionaries are ultimately produced. The typological nature of the ISM is best suited for the music industry and music education in Zimbabwe. Multilingualism is both a resource and a tool which needs to be developed in various spheres of Zimbabwean life. Lexicography can make significant contributions in this regard.

3. The Frame Structure (Megastructure)

Metalexigraphic literature indicates that the central list bias in lexicographic practice, be it in the compilation or evaluation of dictionaries, is now disap-

pearing. The 'word book structure' (Hausmann and Wiegand 1989) which regards the so-called 'outer texts' as functional text components of dictionaries has been adopted. Not only the word list is considered as the dictionary, but also the 'front matter' and 'back matter' texts. Accordingly, a dictionary is now regarded as a 'carrier of texts' (Wiegand 1996, Kammerer and Wiegand 1998). The resultant type of dictionary structure is called the 'frame structure' (Kammerer and Wiegand 1998) or the 'megastructure' (Hartmann and James 1998, Yong and Peng 2007). Various studies and publications by scholars such as Gouws (2002, 2004, 2007), Gouws and Prinsloo (2005), Tarp (2005), Bergenholtz and Nielsen (2006) and Yong and Peng (2007) demonstrate several ways in which the outer texts ought to complement the central list in the realisation of lexicographic functions. For the purpose of this section, it can simply be noted that the ISM conforms to this type of dictionary structure. It consists of the front matter (pp. i-xxvii), the central word list (pp. 1-124) and the back matter (pp. 125-195). In the remainder of the article, it will chiefly be demonstrated how these different text components complement each other in the dictionary.

4. The Front Matter

The front matter of the ISM contains the following text components:

- (a) the title page,
- (b) the imprint page,
- (c) the contents page, *Okumunyethweyo*,
- (d) the editorial and participants page, *Abaphathekileyo*,
- (e) the preface, *Isandulelo*,
- (f) the foreword, *Isethulo*,
- (g) the acknowledgements, *Amazwi Okubonga*,
- (h) the bibliography, *Umthombo Wolwazi*,
- (i) the introduction, *Isingeniso*,
- (j) the guide for the use of the dictionary, *Okumele Kulandelwe Ngosebenza Isichazamazwi*, and
- (k) the editorial request, *Isicelo Sabahleli*.

Each of the ISM front matter components listed above falls within the distinct categories of integrated or non-/unintegrated outer texts (Gouws 2004). The distinction is based on whether the genuine purpose of the dictionary or any of its lexicographic functions which primarily occur in the central list are realised or complemented in a particular outer text. It is in the light of this distinction that the above front matter texts of the ISM will be discussed.

Gouws (2004: 72) describes the difference between these outer texts as follows:

Unintegrated outer texts complement the central list and are not needed to retrieve the information presented in the articles of the central list, neither do they

add to the treatment of the subject matter of the dictionary. Integrated outer texts function in co-ordination with the central list and are aids in ensuring an optimal and full retrieval of information and an accomplishment of the genuine purpose.

The imprint page provides the bibliographic information about the dictionary. This includes its title (already given on the cover and title page), names of the editors, the publisher and the ISBN number. This information is useful for library cataloguing and database indexing, but not for using the dictionary and solving problems which motivate dictionary consultation. It is thus a typical example of an unintegrated outer text. It has nothing or very little to do with the contents of the dictionary, their access and use. The same would generally apply to the editorial and participants page, the preface, the foreword, the acknowledgements, the introduction and the editorial request pages, although this would depend on how much information related to the contents and use of the dictionary are included in a particular text. For example, because it was written by a board member of ALRI who was not closely involved in the actual production of the dictionary, the ISM foreword is less informative than the preface which was written by the coordinator of the ALLEX Project, or more so the introduction written by the editors.

On the other hand, the title page, the contents page, the bibliography page and the guide to dictionary use may be very helpful as far as the retrieval of information in the dictionary is concerned. Therefore, they fit into the category of integrated outer texts, which may be distinguished further as either function-adhering or non-function-adhering (Gouws 2007). Function-adhering outer texts are those from which lexicographic information may be retrieved to achieve a given lexicographic function which may not necessarily occur in the word list (Gouws 2007: 82). Those which do not contain lexicographic data but merely facilitate its access through guidance are integrated non-function-adhering outer texts. This description suits the title page, the contents page, the bibliography page and the guide to dictionary use as they provide guidance, but not information which may initiate a look-up procedure. In addition to describing how these texts work in the ISM, it is attempted to draw attention to how such texts may serve a more comprehensive purpose.

The title of the dictionary is its identity, just like the cover entries. According to Gouws and Prinsloo (2005: 165), cover entries "are not only there for promotional purposes or as a form of lexicographic cosmetics", but rather "functional entries" which "form an integral part of lexicographic presentation". Therefore, the title, as it appears on the front cover, the spine and title page of the dictionary, informs the user what to expect in the dictionary, for example, whether it is the most suitable for a particular consultation procedure or not. There being only three Ndebele dictionaries in circulation, the user who needs information about a music term would tell from the title that the ISM would be the best available dictionary. Gouws (2007) suggests that users would benefit more if the title page were to provide as its subtitle a hint of the lexicographic functions which the dictionary covers. Although this is not given in the ISM, it

should be considered for future dictionaries, LSP dictionaries included. Still, the title page would remain an integrated non-function-adhering outer text.

The same would apply to the bibliography. The bibliography of the ISM lists the sources which provided terms, and definitional and etymological data contained in the dictionary. From these sources, users may obtain more comprehensive specialised and cultural music knowledge. However, since users obtain, except such guidance, no lexicographic information from the bibliography, it therefore remains an integrated non-function-adhering text. Such a text could be made more helpful if it were clearly linked to certain articles in the main text through cross-referencing. This would, as suggested by Nielsen (1999), enable the user to know the relevant source for a particular type of information.

The ISM guide to dictionary use (pp. xix-xxv) explains how the data contained in the dictionary and enumerated on the contents page may be accessed. The distribution of lexicographic data as well as their arrangement and use of lexicographic conventions is explained and illustrated by means of articles extracted from the main text. The following table is also part of the ISM guide to dictionary use.

Table 1: Symbols used in the ISM and their explanations (p. xxiv).

UPHAWU (indicator)	NGESINDEBELE (Ndebele)	NGESILUNGU (English)
[]	<i>Enye indlela yokubiza igama</i>	Variant enclosed
< >	<i>Imbali yegama</i>	Etymology enclosed
FAN	<i>Amagama alengcazelo ezifanayo</i>	Synonyms
BONA	<i>Bona ingcazelo</i>	Find meaning
KHA	<i>Khangela</i>	Check or Compare
(SH)	<i>Ligama lesiShona</i>	Shona equivalent
(EN)	<i>Ligama lesiLungu</i>	English equivalent

The first column of Table 1 lists the symbols and abbreviations used in the ISM. The second column explains each symbol in Ndebele while the third provides English explanations. The article will deal with the use of these symbols and abbreviations in the sections which follow, especially Sections 6 and 8. While front matter texts have often been viewed as occupying valuable dictionary space, because users rarely refer to them, they play a crucial role even in the realisation of the genuine purpose and functions of the dictionary. This has been indicated by describing the ISM front matter texts according to the distinction between integrated and unintegrated (Gouws 2004) as well as function-adhering and non-function-adhering (Gouws 2007). Gouws (2007) argues that front matter texts are in fact consulted, although the frequency of their consultation may not be expected to equal that of the main text since users would read the guide to dictionary use once, but after having mastered its conventions, often continue to consult only the main text. The challenge for practising

lexicographers, therefore, is to evince maximum caution in the planning and utilisation of the front matter texts for data distribution and user guidance.

5. The Macrostructure

This section describes the elements of the ISM macrostructure. It also highlights and explains the procedures adopted to formulate the macrostructure, the challenges encountered as well as the counter-strategies employed in the process.

The macrostructure of the ISM consists of 1 228 articles. The lexical items represent the conceptual and thematic structure of the music domain which is based on traditional Ndebele and contemporary Zimbabwean music. As a result, the macrostructure consists of indigenous and contemporary music terms. Indigenous music terms are mainly categorised as follows:

- (a) Names of traditional dances such as **isitshikitsha**, **ingquzu**, **indlamu**, **amajukwa** and **amantshomane**.
- (b) Names of traditional instruments such as **umqangala** (mouth bow), **uphondo** (horn) and **izikeyi** (wooden clappers).
- (c) Names of traditional ceremonies in which music plays an important role such as **iNxwala** (first-fruits ceremony), **umthethelo** (ancestral appeasement ceremony), **umbuyiso** (home-bringing ceremony) and **indunduzelo** (initiation ceremony).
- (d) Names of various types of traditional costume and attire such as **ama-bhetshu** (skin dress) **imisisi** (skin attire for women), **indlukula** (headgear), **isidlodlo** (headgear) and **amahlwayi** (leg-rattles).
- (e) Names of music/song categories according to ceremonies and situations of rendition such as **ingoma yeNxwala** (first-fruits ceremony song), **ingoma zokuzingela** (hunting songs) and **ingoma zokugiga** (pounding songs).

Such terms represent the core of traditional music knowledge in Ndebele. Furthermore, terms representing contemporary music knowledge were lemmatised, since they are integral in the production, study, analysis and appreciation of music and hence future developments in the music domain. The Ndebele people are now exposed not only to traditional Ndebele music but also to Zimbabwean and international music. To provide for this reality, the production of the ISM had to, as stated in Section 2, consider borrowed terms from mainly Shona and English. A significant number of lexical items in the ISM are names of dances and instruments which Ndebele people learnt or borrowed from Shona. The adoption and adaptation of Shona terms which resulted from this interaction is illustrated in the following table:

Table 2: Adoption and adaptation of Shona music terms into Ndebele

Shona Term	Lemmatised ISM Term
<i>nyunganyunga/nyungwenyungwe</i>	inyunganyunga/inyungwenyungwe
<i>nhare</i>	inare
<i>muchongoyo</i>	umtshongoyo
<i>jerusarema</i>	ijerusalem
<i>kongonya</i>	ikongonya
<i>shangara</i>	ishangara
<i>mhande</i>	imande
<i>mbira</i>	imbila/imbira
<i>mbakumba</i>	imbakumba
<i>magagada</i>	amagagada
<i>maghavu</i>	amagavu
<i>madhebhe</i>	amadebhe

The first procedure for the adoption and adaptation of Shona terms into Ndebele is the prefixing of these terms with an initial vowel. This is in agreement with the morphological structure of nouns in the Nguni languages, of which Ndebele is one. By including an initial vowel as a pre-prefix, the ISM editors already took a significant step in adapting Shona terms to Ndebele morphology. As illustrated in Table 2, this is the most remarkable difference between the Ndebele terms and their Shona equivalents in the ISM.

However, the inclusion of an initial vowel alone before the prefix proper does not change all Shona terms into Ndebele. Some terms have sounds not fully accepted in Ndebele, while the sounds of others are represented differently in Ndebele. The trill [r] is one sound which is resisted in Ndebele although most speakers find it unavoidable in other words, some of which are borrowed from English. Instead of [r], some speakers would prefer the lateral [l]. As a result, /r/ and /l/ in the fourth and fifth syllables respectively of the Shona term *jerusarema* and the Ndebele term **ijerusalem** (for an original type of Shona dance) distinguish the two terms. Although Ndebele borrowed this term from Shona, it should be noted that Shona adopted the Biblical 'Jerusalem' to name this traditional dance whose survival was severely threatened by the influence of early missionaries who found it obscene. Thus, it is actually Shona which rejected /l/ in the fourth syllable of 'Jerusalem' in preference of /r/. In other cases, it is not only difficult but also futile to replace /r/ with /l/ because words become awkward and difficult to pronounce. In this regard, relevant examples from Table 2 are **inare** (a type of mbira), **ishangara** (a type of dance) and **imbira** (mbira). However, in the case of **imbira**, the term was treated as a variant of **imbila** which, according to music experts, is equally popular. What is interesting though is that **imbila** already existed in Ndebele, referring to a rabbit. If this does not cause problems in the music domain, the rephonologised term may become the standard term, but if it does, **imbira** will have to be fully

accepted. This flexibility is provided for by a proscriptive lexicographic approach as opposed to purely prescriptive and descriptive approaches (Bergenholtz 2003, Tarp and Gouws 2008).

Another phonological and orthographic feature which the ISM editors had to deal with in the adaptation of Shona terms to Ndebele is the distinction between certain sounds and their breathy counterparts. Ndebele does not make this distinction, but Shona does by using the voiced glottal fricative [fi]. Some of the sounds involved include the nasals [m] and [n], the voiced velar stop [g] and the voiced alveolar stop [d]. In Table 2, the Shona terms *mhande* (type of traditional dance) and *nhare* (type of mbira) contain the breathy voiced nasals [ṃ] and [ṅ] respectively. Adapting them to Ndebele resulted in **imande** and **inare** respectively. Although the two breathy nasals are realised in spoken Ndebele as in the first syllables of *unina* (his/her mother) and *mana* (stop) respectively, this phenomenon is not given special orthographic presentation as in Shona. In cases where *mh* is realised in Ndebele writing, the pronunciation differs from that of Shona. For instance, in the word *umhawu* (jealousy), *m* becomes a separate syllable as a contracted form for *mu*. Thus *Mhangura* (the name of a town) can be read in two completely different ways by a Ndebele and a Shona speaker. The adaptation of some Shona terms to Ndebele, therefore, included the dropping of *h* from breathy nasals as it causes phonological inconsistencies in some instances. The same procedure was applied to the Shona terms *maghavu* (leg-rattles) and *madhebhe* (a type of mbira) which contain the breathy glottal stop [ʔ] and the alveolar stop [d] respectively to result in **amagavu** and **amadebhe**.

For the adaptation of Shona terms to Ndebele the differences in the presentation of the voiceless palatal affricate [tʃ] also had to be considered. In Shona, the sound is orthographically written as *ch* while it is rendered as *tsh* in Ndebele. As shown in Table 2, the Shona term *muchongoyo* (a type of traditional dance) becomes **umtshongoyo** in Ndebele. While the spelling *ch* exists in Ndebele, it is a dental click [ɬ] as in *-chitha* (spill) or *umchamo* (urine). Thus, simply prefixing the Shona term *muchongoyo* with an initial vowel would produce a different pronunciation and fail to capture the meaning of the original term.

A fairly large number of lexical items in the ISM are terms borrowed from English. Similar principles of adopting and adapting English as those discussed previously in the case of Shona terms were used. Hadebe (2002: 138) points out that "a loanword might be accepted into the language as part of vocabulary growth ... or coined as a term and be accepted but has to fit in the Ndebele writing system". This remained an important factor in the adoption of music terms from English. Specifically, the challenges encountered pertained to what could have been misconceived as vowel sequencing when initial vowels were prefixed on vowel-commencing noun stems from English and other unacceptable sound clusters. The challenges of dealing with vowel-commencing noun stems is summarised in Table 3 below.

Table 3: Lemmatising vowel-commencing adoptives from English in the ISM

English Term	VSS	Hyphenated	Semi-vocalised	ISM Lemmata
alto	<i>ialitho</i>	<i>i-alitho</i>	<i>iyalitho</i>	<i>alitho</i>
encore	<i>ienkho</i>	<i>i-enkho</i>	<i>iyenkho</i>	<i>enkho</i>
earphone	<i>iiyefoni</i>	<i>i-iyefoni</i>	<i>iyiyefoni</i>	<i>iyefoni</i>
oboe	<i>iobho</i>	<i>i-obho</i>	<i>iwobho</i>	<i>obho</i>
audition	<i>iodishini</i>	<i>i-odishini</i>	<i>iwodishini</i>	<i>odishini</i>
organ	<i>iogani</i>	<i>i-ogani</i>	<i>iwogani</i>	<i>ogani</i>

As noted earlier, prefixing noun terms from Shona with initial vowels was the first step to adapting them to Ndebele orthography. This also applies to English terms and terms from other languages with a different noun structure. This seemingly simple procedure resulted in several lemmatisation dilemmas in the ISM with regard to adopted vowel-commencing English nouns like those in the first column of Table 3. The initial vowel, which is characteristic of nouns in Ndebele and other Nguni languages would result in orthographic words such as those in the second column of Table 3. The Ndebele orthography which needs updating does not permit vowel sequencing. Nouns with vowel-commencing stems are treated inconsistently by different writers, some of whom identify vowel sequencing in them. However, the reality is that vowel sequencing is impermissible within syllables and that what appears in the second column of Table 3 is not vowel sequencing but vowel-syllable sequencing (VSS). Although the ISM editors were fully aware of this, they had to avoid controversy, especially regarding a dictionary with a very limited user-group, by not adopting a strategy which was not even used in the ISN. Had it been used in the ISN, there would have been entries such as *ioda* (order), *iorenji* (orange) and *ioyili* (oil), which have, however, been lemmatised as *iwoda*, *iworenji* and *iwoyili* respectively. This involves the process of semi-vocalisation, whose results are presented in the fourth column of Table 3. By adopting this strategy, the ISN editors avoided yet another strategy used by many Ndebele writers, which consists of inserting a hyphen between the initial vowel and the adopted noun stem as shown in the third column of Table 3. Following the principles of the ISN, the ISM editors also avoided the use of the hyphen which is mainly employed in breaking compounds. Neither did they lemmatise the adoptives with semi-vowels bridging the perceived vowel sequences as the ISN editors have done. The reason for this is that semi-vocalisation is not predictable in all contexts. Blending the letters *i* and *a* may result in inserting either a *y* or an *h* so that the user may be uncertain whether to search the lemma using *iy-* or *ih-*. Relevant examples in the ISM regard the English terms *alto*, *encore* and *earphone*, which would be lemmatised as *iyalitho*, *iyenkho* and *iyiyefoni* respectively. Yet informal discussions with Ndebele linguists have indicated that *ihalitho*, *ihenkho* and *ihiyefoni* would have been better. In the ISM, only the rephonologised English stems were lemmatised without initial vowels as shown in the last column, thereby avoiding the problems associated with each of the first three strategies.

The structure of the article headed by such lemmata is illustrated by the following example:

alitho [ikhontiralitho] aruto (SH) alto (EG) ...
Iyalitho liphimbo lowesifazana lokuhlabela
 eliphansi kulawo wonke. ... (ISM, p. 2).

The access position of the adopted term in the macrostructure of the ISM is determined by the first letter of the vowel-commencing stem of the English term *alto*, as it is also the case in the English alphabetic equivalent register. Together with the statement in the front matter drawing users' attention to this exception to the rule that all noun terms are lemmatised according to their initial vowels, the alphabetic equivalent register may be useful in guiding users to the lemma type in question, given that the English terms are better known. Semi-vocalisation is then employed in the full-sentence definitions which include the defined lemmata. In the definitions of this type of lemmata under discussion, lemmata are preceded by initial vowels and semi-vocals. The definition of *alitho* in the exemplified article thus begins with '*Iyalitho* ...' and the lemma is italicised as in the rest of the articles (see Section 6).

In addition to the challenge resulting from vowel-commencing English stems, there are yet other sound clusters in English which are problematic in Ndebele and had to be avoided in the borrowing process. For example, the adoptives from *orchestra* and *quadrille* were lemmatised **okhesitira** and **ikhwadirili**. While *iokhestra* and *ikhwadri* respectively are realised in speech, their acceptance in writing is still contentious, so that the ISM editors had to avoid these representations. The problematic sound clusters in this regard are [str] and [dr] respectively.

In an attempt to represent contemporary musical knowledge in the ISM, loan translation was also used with the result that compound terms became part of macrostructural elements. Examples include **amazwingoma** from *amazwi engoma*, literally meaning 'words of a song' for *lyrics* and **inqalakutsha** from *-qala kutsha*, literally meaning 'a fresh start' for *da capo*.

Multi-lexical lexical items are also part of the ISM macrostructure. Terms such as **ihubo lenkozo** (chorale), **ikhandensi ephanjanisiweyo** (interrupted cadence), **isikali somculo** (musical scale) and **isikali esincane** (minor scale) are some examples. While the decision to include them was made at the initial planning stage, multi-word units posed a challenge that only surfaced at the final editorial stage carried out together with the publishing editor. Some lemmata appeared where they were least likely to be searched for in the alphabetic macrostructure, thereby posing a danger of making them less accessible to the users. For example, the users, as well as the editors, are likely to take it as given that the article stretch including lemmata ranging from **ingoma yeNxwala** to **ingoma zokuzingela** would be preceded by articles headed by **ingomabafi**, **ingomabusuku**, **ingomanyoni** and **ingomasililo**. However, the computer interprets the space between compound words as a letter coming before the

letter *a*. Accordingly, the multi-word units made an earlier alphabetic appearance than anticipated. However, although users may not fail to find such lemmata as they immediately follow the lemma **ingoma**, a statement in the front matter, especially in the guide to dictionary use, would have sufficed in drawing the users' attention to this fact.

The tonic solfa notes, namely *s, f, m, r, t, l, d* and *d*, were entered as they are pronounced and sung to render them respectively as **so, fa, mi, re, thi, la, do** and **do** without initial vowels, just like terms with vowel-commencing stems as exemplified by **alitho** earlier on. Otherwise, the arrangement of lemmata in the ISM is strictly letter by letter, with spaces between words in multi-lexical items being treated as alphabetical letters.

6. The Microstructure

This section focuses on the structure of dictionary articles. Before looking at those articles with more data items, the structure of articles with the least number of data items is presented first. This presentation is accompanied by a discussion of structural indicators which facilitate information retrieval from the various data categories.

The following structure represents those articles with the least number of data items:

Lemma → Shona Equivalent → English Equivalent → Definition

As in most dictionaries, bold print is used for lemmata so that they are distinct and therefore easy to find. As noted earlier, some terms have been created to represent concepts which are less known to Ndebele users with a low level of encyclopaedic competence. Such lemmata are important for presenting spelling information.

Since the dictionary intends to facilitate an understanding of musical concepts whose lemmata are represented primarily in Ndebele, definitions follow in terms of importance. The definitions explaining the concepts consist of full sentences. The term defined is italicised in the definition. In contradistinction to the ISN, not all words used in the definitions are entered and defined in the ISM since some are 'common words', not specifically musical terms, which users should know or could look up in the ISN. To give users more details about the music concepts, ISM definitions are longer and more encyclopaedic than those of the ISN. This becomes clear in cases where certain lexical items are lemmatised in both dictionaries. To achieve this, defining principles were formulated to ensure that terms for similar concept categories such as dances, music genres, instruments, ceremonies and various types of costumes are defined exhaustively and at the same time uniformly.

Shona and English equivalents give the ISM a trilingual dimension, but as noted earlier, not a trilingual dictionary typology. The indicators SH and EN,

being abbreviations for Shona and English respectively, are used. Equivalent terms have been included because the cultures represented in Shona and English have affected Ndebele culture, the music domain being no exception. In some cases, the equivalent is more familiar to the users than the Ndebele lemmatised term. The reason for this is that the equivalents and the terms in the dictionary are not, as mentioned in Section 2, direct equivalents owing to the cultural factor and linguistic anisomorphism. This, as indicated in Section 7, is reflected in the equivalent registers of the dictionary.

Those articles with more data items are now dealt with. An ISM article with all data items that the dictionary provides would show the following structure:

Lemma → Variant → Shona Equivalent → English Equivalent → Etymology → Definition → Synonym → Comparative reference → Reference to illustrations

No article in the dictionary contains all the data items given in the above structure. The inclusion of a particular data item was determined by the type of term and the concept represented with a view to its value for the user. The value of the data items discussed regarding articles with the least number of microstructural entries is the same, even in those articles which give more data. Repeating a discussion of them is therefore unnecessary.

Where required, variants are provided after the lemma, enclosed in square brackets (refer to Table 1). They highlight the fact that the language is not standardised, so that transliterations may be rendered in more than one form. A good example is **iphiyano** [ipijano] for *piano*. Some Ndebele speakers pronounce this borrowed term with an aspiration as [p^h] [ip^hijano] while others pronounce it without aspiration but as an ejective [p'] [ip'ijano]. The former rendition was defined as the main lemma and the latter by a cross-reference to the main lemma (see Section 8). This suggests to users which term is preferred while space is saved at the same time. Besides these instances, variants appear in the ISM as transliterations for some acronyms from English. Examples are **isidi** [idisiki] for *compact disk* and **idakhapho** [idisi] for *da capo*. The former is defined as the main lemma because it is more popular than the latter whose rendition in Ndebele is problematic. Instead of **idisiki**, the preference may have been for *idiski* which is written as it is pronounced, disregarding the [sk] sound combination which remains contentious in writing. This is different from the latter case where the transliteration of the fuller form was preferred to that of the acronym *DC* which is used along with the full form.

Etymology was provided for musical concepts whose terms were adopted from English yet having either Greek, French, Latin or Italian origins. The information that users can retrieve includes the source language, the word form in that language and the literal meaning which elucidates the contemporary meaning of the term. *IsiGirikhi*, *isiFurentshi*, *isiLathini* and *isiTaliyana* re-

spectively represent the four main source languages listed above. The following example illustrates this.

iligetsha ... ligature (EN) <IsiTaliyana: *ligature* 'okubambanisayo'>
(that which links) ... (ISM, p. 39).

The provided etymological data enhances a better understanding of the term through the link that can be established from the literal meaning to the specialised meaning. For instance, *ligature* from the above example refers to a notation symbol which links the first note to the last one in a musical phrase. Literally it means "that which links", but in a musical sense it refers to a specific notation symbol.

Just like variants, synonyms in the ISM are partly the result of borrowing terms from English where transliterations are in a synonymous relationship with Ndebele coinages. Examples are **ilizwi/phimbo/iphathi** (part of polyphonic music) and **umaqondisekude/irimothi** (remote controller). There are also synonymous indigenous terms such as **ingoma/iculo** (song) and **-haya/-bonga** (praise mainly through poetry). In all such cases where both terms seem to be equally familiar, the compilers could not prescribe by excluding one of the terms but only by treating one as the main term and the other as (a) lesser synonym(s) which is defined by 'FAN' (refer to Table 1).

Reference indicators or markers are also part of the ISM microstructure. These include instances where users are referred from one article to another to effect paradigmatic relations as in **itshelo** (cello) where users are referred to all other types of violins or from an article in the main list to the section containing illustrations in cases where a concept is illustrated.

As far as the data items constituting the structure of individual articles are concerned, their provision depends largely on the term involved and the concept it represents. Information is provided in order to meet user needs in respect of a particular term but this provision is systematic for related terms. Once users learn the communicative significance of structural indicators, which are explained in the front matter, they can consistently apply it to all related articles.

7. The Back Matter

The back matter of the ISM contains four major sections in the following order: the section containing illustrations, *Imifanekiso*, the section containing songs, *Ingoma ezikhethiweyo*, the Shona alphabetic equivalent register, *Amagamanhloko ngesiShona*, and the English alphabetic equivalent register, *Amagamanhloko ngesiLungu*. These sections, which complement the main text can easily be accessed through the contents page. In the way in which they complement the central text, they may, just like the front matter texts, be understood according to the distinctions between integrated versus non-integrated outer texts on the one

and function-adhering versus non-function-adhering outer texts on the other hand.

Illustrations serve to elucidate some of the concepts whose terms are lemmatised and defined in the central list. They are not just meant to brighten or enliven the dictionary. A cross-reference link is provided between illustrated terms and their illustrations. The section containing songs illustrates the practical use of the music notation symbols whose terms are lemmatised and defined in the central list. The mediostructure also facilitates the user search path between the main text and this section since the notations are named and defined in the main text. Thus, both the section containing illustrations and the section containing songs are integrated function-adhering outer texts.

Being in the access alphabetical order, the alphabetic equivalent registers, also called reverse indices, guide users not familiar with the Ndebele terms for the alphabetised Shona and English terms to the central list in which the Ndebele terms are lemmatised. This is necessary because some of these Ndebele terms were created for the first time during the production of the ISM. As indicated in Section 7, the alphabetic equivalent registers make the dictionary poly-accessible. However, it should be noted that in the many cases where direct equivalence does not exist, the equivalents are mainly explanatory and sometimes very brief paraphrases of the Ndebele terms. In the Shona register, *mutambo wechinyakare* has more than 20 different types of traditional dances recognised in Ndebele, while English has more than twice this number for *traditional dance*. Still, the information contained in the equivalent register may suffice for dictionary consultations whereby the user just wants to know the equivalent. In this way, the alphabetic equivalent registers are also integrated function-adhering outer texts.

8. The Mediostructure

Cross-referencing is a useful lexicographic device which is an integral component of the access structure. It guides the dictionary-user from one point to another for more information or for comparison. Cross-referencing can be explicit or implicit: in the former, the user is clearly advised or commanded to go to another point in the dictionary, while in the latter, the user is made to feel the need to go to another point without clear advice or command (Gouws and Prinsloo 1998, 2005; Louw 1999; Yong and Peng 2007).

Explicit cross-referencing was employed in the ISM in the following cases:

- (a) Between synonyms whereby lesser synonyms are defined using other terms which would be more preferable. 'FAN' (refer to Table 1) is employed to link the more preferable lemma to its synonyms which the preferable lemma defines. For instance, 'FAN' *-khalisa* in the article corresponding to the lemma *-tshaya* (play a musical instrument or gadget) informs the user that while the two terms are synonymous, *-tshaya* is the

more preferable term. The term **-khalisa** is then defined using *-tshaya* which directs the user to the more informative article for the meaning.

- (b) Between variants whereby a less popular variant is defined using a more popular one. A less popular variant is enclosed in square brackets immediately after the main form as in **ipansula** [isipansula] (a dance). The term **isipansula** is then linked to the main form using the guide word 'BONA' which directs the user to **ipansula** which has been given a more comprehensive lexicographic treatment.
- (c) From terms to illustrations whereby users are advised to refer to the section containing illustrations for a picture of an instrument, a piece of costume or a musical notation symbol. 'BONA *umfanekiso*' (SEE *the illustration*) at the end of the article corresponding to the lemma **isikali esikhulu** (major scale) (ISM, p. 73) advises the user to refer to the diagrammatic illustration of the major scale for effective conceptualisation.
- (d) Between terms denoting concepts which are in a paradigmatic relationship, the user is advised through 'KHA' (refer to Table 1) to compare the term **ivayola** (viola), for instance, with **ivayolini** (violin), **itshelo** (cello) and **ibhesimbili** (double bass) which denote the other types of violins (ISM, p. 86). Unlike 'BONA', 'KHA' is not obligatory since all the necessary information on the term would be given in the reference article. Cross-referencing in such a case gives information that is not sought in the initial look-up process although it may benefit the user.

Implicit cross-referencing may be noted in cases where the user is not explicitly advised to refer, but may feel the need to do so, such as in the following:

- (a) From the illustrations, especially of musical notation symbols to selected songs and vice versa. After moving from a term denoting a musical symbol, the user may refer to illustrations and proceed to check how the symbol is used in the selected songs. The user may also start from a particular song in the section containing songs, then check the symbol under the illustrations and eventually go to the main text to find out how it is used in singing.
- (b) From the alphabetic equivalent registers to the main text whereby a user who may know the English or Shona term may start from the English or Shona alphabetic equivalents respectively to find out what a concept is called in Ndebele. The user may then proceed to the main text to find more information on the concept in Ndebele. This would be a case of the user taking advantage of the poly-accessible structure of the dictionary.
- (c) From the ISM to other references outside the dictionary whereby the user may go through the list of references (dictionaries and other books) entitled *Umthombo Wolwazi* (p. xv) to obtain extra details which the dic-

tionary could not provide. This is called dictionary external cross-referencing since access to the reference address involves exceeding the boundaries of the dictionary in use (Gouws and Prinsloo 2005: 180).

9. Conclusion

This article has given some insights into the production of the Ndebele music terms dictionary, *Isichazamazwi SezoMculo* (also referred to as the ISM). It has also outlined the structure of the dictionary, especially regarding dictionary contents, their distribution and presentation. It has been shown that the dictionary was compiled within the parameters of lexicographic principles, taking cognisance of the users' reference needs and skills. Thus, the resultant dictionary structure is not just a case of convention or a form of lexicographic cosmetics, but a vital component of dictionary planning which was intended to facilitate both data presentation and data retrieval. It is a mode of lexicographic communication. The article by no means claims that the ISM structure is without limitations as indications are given of how the dictionary could have been made more user-friendly. A critical analysis of the ISM by dictionary users and other metalexigraphers might offer insights which may be useful for future LSP lexicography in Ndebele and other African languages. The insights may also be of value to lexicography in general.

References

- Bergenholtz, H.** 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Bergenholtz, H. and S. Nielsen.** 2006. Subject-field Components as Integrated Parts of LSP Dictionaries. *Terminology* 12 (2): 281-303.
- Bergenholtz, H. and S. Tarp.** 1995. *Manual of Specialised Lexicography*. Amsterdam: John Benjamins.
- Bergenholtz, H. and S. Tarp.** 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes: Journal of Linguistics* 31: 171-196.
- Burkhanov, I.** 2004. Requirements for an "Ideal" Bilingual L1-L2 Translation-Oriented Dictionary. *Lexikos* 14: 17-32.
- Gouws, R.H.** 1996. A Sequence for Meeting Lexicographic Needs. *Lexicography as a Financial Asset in a Multilingual South Africa*: 97-110. Pretoria: Department of Arts, Culture, Science and Technology.
- Gouws, R.H.** 2002. Using a Frame Structure to Accommodate Cultural Data. Emejulu, J.D. (Ed.). 2002. *Éléments de Lexicographie Gabonaise. Tome II*: 54-69. New York: Jimacs-Hillman Publishers.
- Gouws, R.H.** 2004. Outer Texts in Bilingual Dictionaries. *Lexikos* 14: 67-88.
- Gouws, R.H.** 2004a. Monolingual and Bilingual Learners' Dictionaries. *Lexikos* 14: 264-274.
- Gouws, R.H.** 2007. A Transtextual Approach to Lexicographic Functions. *Lexikos* 17: 77-87.
- Gouws, R.H and D.J. Prinsloo.** 1998. Cross-referencing as a Lexicographic Device. *Lexikos* 8: 17-36.

- Gouws, R.H. and D.J. Prinsloo.** 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PReSS.
- Hadebe, S.** 2002. *The Standardisation of the Ndebele Language through Dictionary Making*. Unpublished D.Phil. Dissertation. Harare/Oslo: University of Zimbabwe/University of Oslo.
- Hadebe, S.** 2005. Reflections on the Proposed Ndebele–Shona/Shona–Ndebele Dictionary. *Lexikos* 15: 265-275.
- Hadebe, S. et al. (Eds.).** 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hartmann, R.R.K.** 1983. *Lexicography: Principles and Practice*. London: Academic Press.
- Hartmann, R.R.K. and G. James.** 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Hausmann, F.J. and H.E. Wiegand.** 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Berlin/New York: Walter de Gruyter.
- Kammerer, M. and H.E. Wiegand.** 1998. Über die textuelle Rahmenstruktur von Printwörterbüchern. Präzisierung und weiterführende Überlegungen. *Lexicographica: International Annual for Lexicography* 14: 224-238.
- Louw, P.A.** 1999. Access Structures in a Standard Translation Dictionary. *Lexikos* 9: 108-118.
- Mheta, G. (Ed.).** 2005. *Duramazwi reMimhanzi*. Gweru: Mambo Press.
- Ncube, C.** 2005. Language Development or Language Corruption: A Case of Loan-words in *Isichazamazwi SesiNdebele*. *Lexikos* 15: 294-305.
- Nielsen, S.** 1999. Mediostuctures in Bilingual LSP Dictionaries. *Lexicographica: International Annual for Lexicography* 15: 91-113.
- Nkomo, D.** 2005. Compiling Specialised Dictionaries in African Languages: *Isichazamazwi SezoMculo* as a Special Reference. *Zambezia: The Journal of Humanities of the University of Zimbabwe* 32 (1–2): 64-77.
- Nkomo, D. and N. Moyo (Eds.).** 2006. *Isichazamazwi SezoMculo*. Gweru: Mambo Press.
- Pelling, J.N.** 1966. *A Practical Ndebele Dictionary*. Harare: Longman Zimbabwe.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- Tarp, S.** 2000. Theoretical Challenges to Practical Specialised Lexicography. *Lexikos* 10: 189-208.
- Tarp, S.** 2002. Basic Elements of a Lexicographic Theory. Emejulu, J.D. (Ed.). 2002. *Éléments de Lexicographie Gabonaise. Tome II*: 7-20. New York: Jimacs-Hillman Publishers.
- Tarp, S.** 2004. Basic Problems of Learners' Lexicography. *Lexikos* 14: 222-252.
- Tarp, S.** 2005. The Pedagogical Dimension of a Well-Conceived Specialised Dictionary. *Iberica: Journal of European Languages for Specific Purposes* 10: 7-21.
- Tarp, S. and R.H. Gouws.** 2008. A Lexicographic Approach to Language Policy and Recommendations for Future Dictionaries. *Lexikos* 18: 232-255.
- Wiegand, H.E.** 1996. A Theory of Lexicographical Texts. An Overview. *South African Journal of Linguistics* 14(4):134-149.
- Yong, J. and H. Peng.** 2007. *Bilingual Lexicography from a Communicative Perspective*. Amsterdam/Philadelphia: John Benjamins.

Sense Relations in the Treatment of Meaning in *Isichazamazwi* *SesiNdebele*

Eventhough Ndlovu, *Department of African Languages and Literature,*
University of Zimbabwe, Harare, Zimbabwe (endlovu@arts.uz.ac.zw)

Abstract: This article evaluates how meaning has been treated in *Isichazamazwi SesiNdebele* (2001) through the application of the sense relations hyponymy, meronymy and synonymy. It attempts to establish the adequacy of meaning presentation in definitions given to headwords. It assesses how the application of sense relations contributes to the user-friendliness of the definitions, the accessibility of the dictionary contents and the semantic analysis of lexemes in dictionaries. It aims to give insights that will help refine and improve future dictionary definitions. The article reveals that hyponymy and meronymy facilitate the formulation of more user-friendly and accessible dictionary definitions. On the other hand, synonym definitions reduce the user-friendliness and accessibility of the definitions in dictionary types that are not dictionaries of synonyms. As dictionary making is a part or a form of standardizing a language, it ought to be as accurate as possible in the presentation of semantic facts which is a reflection of the cultural values and ideologies of the society.

Keywords: HYPONYMY, MERONYMY, SYNONYMY, USER-FRIENDLY, ACCESSIBILITY, DEFINING FORMAT, DEFINING PRINCIPLE, CIRCULARITY, ONOMASIOLOGICAL, SEMASIOLOGICAL, CROSS-REFERENCING

Opsomming: **Betekenisverwantskappe in die behandeling van betekenis in *Isichazamazwi SesiNdebele*.** Hierdie artikel beoordeel hoe betekenis in *Isichazamazwi SesiNdebele* (2001) behandel is deur die aanwending van die betekenisverhoudings hiponimie, meronimie en sinonimie. Dit probeer om die toereikendheid van betekenis-aanbieding vas te stel in definisies wat aan trefwoorde gegee is. Dit bepaal hoe die aanwending van betekenisverhoudings bydra tot die gebruikersvriendelikheid van die definisies, die toeganklikheid van die woordeboekinhoud en die semantiese inhoud van lekseme in woordeboeke. Dit beoog om insigte te bring wat sal help om toekomstige woordeboekdefinisies te verfyn en te verbeter. Die artikel toon dat hiponimie en meronimie die formulering van gebruikersvriendeliker en toegankliker woordeboekdefinisies vergemaklik. Aan die ander kant verminder sinoniemdefinisies die gebruikersvriendelikheid en toeganklikheid van die definisies in woordeboeksoorte wat nie sinoniemwoordeboeke is. Omdat woordeboekmaak 'n deel of 'n vorm van die standaardisering van 'n taal is, behoort dit so noukeurig as moontlik te wees in die aanbieding van semantiese feite wat 'n weerspieëling van die kulturele waardes en ideologieë van die maatskappy is.

Sleutelwoorde: HIPONIMIE, MERONIMIE, SINONIMIE, GEBRUIKERSVRIENDELIK, TOEGANKLIKHEID, DEFINIËRINGSFORMAAT, DEFINIËRINGSBEGINSEL, SIRKULARITEIT, ONOMASIOLOGIES, SEMASIOLOGIES, KRUISVERWYSING

1. Introduction

Definitions are probably the most central information in a dictionary, and most users consult a dictionary more for definitions than for any other information presented in it. Most critics of dictionaries therefore focus mainly on definitions. Because of the importance of definitions, it is necessary to establish defining formats and principles that will help formulate more user-friendly and accessible dictionary definitions adhering to the essential principles of defining. This article looks at how the sense relations hyponymy, meronymy and synonymy can contribute to definitions which avoid circularity, define the entry word, define every word used in the definition and avoid using words more difficult to understand than the entry word. In short, the article assesses the contribution(s) any sense relations make towards the semantic analysis of lexemes and phrases in lexicography.

Sense relations are concerned with the meaning relations words contract with each other in the meaning system of the vocabulary of a language (Lyons 1977: 241). It deals with meaning relations within the vocabulary system of the language. Jackson (1988: 43) defines sense relations as meaning relations between words themselves, a complex system of relationships that exist between the linguistic elements themselves. Sense relations are therefore the intralinguistic relations of lexical items making up the vocabulary of a language.

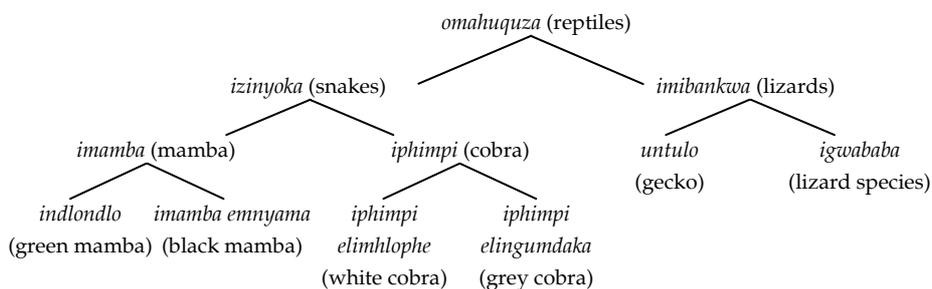
2.1 Hyponymy

According to Jackson (1988: 65), hyponymy refers to the hierarchical relationship between the meanings of lexemes, in which the meaning of one lexeme is included in or under the meaning of another lexeme. Lyons (1977: 242) defines hyponymy as a less familiar relationship that refers to a relation of inclusion, whereby one can say that an X is a kind of Y. Hyponymy is the relation between more specific (hyponymous) terms and less specific (superordinate) terms. According to Crystal (1995: 105), the more general term is called the superordinate or hypernym, and the less general ones are hyponyms.

Hypernyms and hyponyms relate to each other through a hierarchical relation of generality and specificity. Hyponymy is thus a notion that refers to the meaning relation of inclusion of meaning of specific terms into more general terms. This can be illustrated by the following examples: *indlondlo* (green mamba) is a hyponym of *imamba* (mamba), a hyponym of *izinyoka* (snakes), a hyponym of *omahuquza* (reptiles) and *iphimpi elimhlophe* (white cobra) and *iphimpi elingumdaka* (grey cobra) are hyponyms of *iphimpi* (cobra), a hyponym of *omahuquza* (reptiles). If these hyponyms share the same hypernym they are called co-hyponyms. According to this, *indlondlo* (green mamba) and *imamba emnyama* (black mamba) are co-hyponyms of the hypernym *imamba* (mamba), and *iphimpi elimhlophe* (white cobra) and *iphimpi elingumdaka* (grey cobra) of the

hypernym *iphimpi* (cobra), while *imamba* and *iphimpi* are co-hyponyms of the hypernymic term *izinyoka* (snakes), which in turn is a hypernym of *omahuquza* (reptiles).

Crystal (1995: 105) notes that much of the vocabulary of a language is linked by such systems of inclusion which yield semantic networks forming the hierarchical taxonomies. The hierarchical ordering of lexemes can be represented formally using tree diagrams such as the following:



The notion of inclusion contained in the above definition of hyponymy, can be illustrated with the following entry from ISN:

- (1) **iphimpi** bz 5. *Iphimpi* luhlobo lwenyoka encane, elobuhlungu njalo empofu ngokunsundu. Yande ukutholakala endaweni ezilamatshe.
(cobra noun cl 5. A *cobra* is a type of a small snake that has venom and is dark brown in colour. It is usually found in rocky areas.)

In definition (1), the sense of *iphimpi* (cobra) is included in the sense of *inyoka* (snake).

Hyponymy involves entailment, for example to say *leli liphimpi* (this is a cobra) entails that it is an *inyoka* (snake) and an *umahuquza* (reptile). To say *yinsumbelume* (he is one who is past the normal or expected age of marrying) entails that he is a human being and specifically a man, and means that *kathathanga* or *kalamfazi* (he is unmarried). The involvement of entailment in hyponymy forms, according to Palmer (1981: 87), the basis of Carnap's meaning postulates, where it is suggested that the meaning of lexical items can be stated in terms of such entailments. Hyponymy is therefore defined in terms of a one-way entailment. Thus, if one says *lo ngumthakathi* (this is a witch/wizard), it entails that she/he is a human being and means that *ukuthi umuntu uyaloya* (the person practises witchcraft). From this example it can be deduced that being a witch/wizard is hyponymous to being a human being and synonymous with practising witchcraft. Saying *le yinyumba* (this one is barren) entails that she is a female human or animal and means that *kazali* (she is unable to conceive). Here, being barren is hyponymous to being a female human or animal and synonymous with being unable to conceive. If the idea of entailment involved in hyponymy is followed, *inyumba* (barren female human

or animal) could be defined as in the following example proposed by the author.

- (2) **inyumba** bz 9. *Inyumba* ngumuntu wesifazana ongazaliyo loba inyamazana ensikazi engazaliyo.
(**barren female human or animal** noun cl 9. *A barren female human or animal is a female human or animal unable to conceive.*)

In view of the idea of entailment involved in hyponymy, Palmer (1981: 87) concludes that hyponymy can be defined in such a way that synonymy counts as a special case of hyponymy. Recognizing this relationship between hyponymy and synonymy, Carter (1998: 21) interprets hyponymy as a kind of asymmetrical synonymy. The two synonyms *yinyumba* (is barren) and *kazali* (is unable to conceive) are co-hyponyms of "is a female *umuntu* (human) or *inyamazana* (animal)", hence synonymy can be regarded as a special case of hyponymy.

Hyponymy is a transitive relation, meaning that it can be seen 'in transit' all along the line so that if X is a hyponym of Y and Y is a hyponym of Z, then X is a hyponym of Z (Saeed 2003: 69). For example, if *igwababa* (a lizard species) is a hyponym of *imibankwa* (lizards) and *imibankwa* is a hyponym of *omahuquza* (reptiles), then *igwababa* is a hyponym of *omahuquza*. The fact that hyponymy is transitive therefore means that a specific lexeme is a hyponym of any other lexeme that dominates it in the tree as shown in the diagram above.

Hyponymic relations vary from language to language, for example, in English a tomato and a potato are included under vegetables but in ISN *itamatisi* (tomato) and *igwili* (potato) are both defined as *isihlahlakazana* (shrub) not as *umhlobo wembhida* (type of vegetable). The word *imibhida* (vegetable) refers specifically to leaf and green vegetables.

There are, however, some lexemes that do not have any hyponymic relations, so that trying this formula for these lexemes results in somewhat vague and all-inclusive general terms such as those denoting state, place or thing. Abstract nouns as well as verbs and adjectives are among the most difficult in this respect. For example, is *umsindo* (noise) a kind of *umdumo* (sound) or is *umdumo* (sound) a kind of *umsindo* (noise)? In these cases, the level of abstraction is difficult to determine.

2.2 Meronymy

According to Saeed (2003: 70), meronymy refers to part-whole relationships between lexical items. Meronymy involves defining lexemes by saying that something forms part of something else or something contains or possesses something else. Saeed explains this relationship using the sentence frame: X is part of Y or Y has X. It can therefore be said that the *ingono* (nipple) is part of the *ibele* (breast) and the *ucilikicane* (little finger) is part of *isandla* (hand) or the *ibele* (breast) has an *ingono* (nipple) and the *isandla* (hand) has an *ucilikicane* (little finger). Apart from saying *yingxeny*, one can say *yisitho somzimba* (it is an

organ/a part of the body), *yisiqa se ...* (it is a piece/a part of ...) or *lucezu lwe ...* (it is a part of or a piece of ...), which all capture meronymy.

According to Crystal (1995: 107), meronymy reflects hierarchical classifications in the lexicon somewhat similar to taxonomies, but the hierarchies of meronymy are less clear-cut and regular than those of taxonomies. Meronyms vary in terms of how necessary the part is to the whole. Networks identified as meronymy are therefore lexical, that is, it is conceptually possible to segment an item in countless ways, but only some divisions are coded in the vocabulary of a language.

Hyponymy and meronymy therefore differ in terms of transitivity. As illustrated earlier in the article, hyponymy is always transitive but meronymy may or may not be. In other words, meronymy is not consistently transitive in the way hyponymy is. For example, despite the fact that *umunwe* (finger) is a meronym of *isandla* (hand) and *isandla* (hand) a meronym of *ingalo* (arm), one might have some hesitation about the statement *umunwe* (finger) is a part of *ingalo* (arm). A transitive example of meronymy can be illustrated by positing the following: *uzipho* (finger nail) is part of *umunwe* (finger), and *umunwe* (finger) is part of *isandla* (hand), so *uzipho* (nail) is therefore part of *isandla* (hand).

2.3 Synonymy

According to Jackson (1988: 66), two words can be described as synonyms if they can be used interchangeably in all sentence contexts. Saeed (2003: 65) defines synonyms as different phonological words having the same or very similar meaning. Likewise Yule (2006: 104) describes synonyms as two or more forms with very closely related meaning(s), which are often, but not always, intersubstitutable in sentences. In a general and broader sense, synonymy is a relation of sameness of meaning.

From the above observations, one notes that apart from the sameness of meaning involved in synonymy, synonymy is a matter of degree, which therefore yields two types of synonyms: absolute, strict, true or total synonyms and partial, loose or near synonyms. Matthews (1997: 327) defines absolute synonyms as those having meanings identical in all respects and contexts, for example *-jonga/-hlosa* (resolve), *-jwamula/-hluthuna* (snatch), and *-qila/-dlelezela* (cheat). Partial synonyms are those having meanings identical in some contexts or identical only as far as replacing one with the other does not change the truth conditions of a sentence, for example *-zimuka/-nona* (get fat), and *-zala/-khululeka* (give birth). These are partial synonyms, because replacing one with the other does not change the truth condition of the sentence but they cannot be used interchangeably in all contexts, because the rules of usage are not similar and the rules of predication and reference are not the same in habitual texts. The interchangeability of partial synonyms can cause embarrassment to both speakers and listeners, for example:

- (3) (a) *Inkazana izithwele.* (The girl is pregnant.)
 (b) **Inkomo izithwele.* (*The cow is pregnant.) (i.e. of a human being)
 (c) *Inkomo imithi.* (The cow is in calf.)
 (d) **Inkazana imithi.* (*The girl is pregnant.) (i.e. of an animal)
- (4) (a) *Inkazana ikhululekile.* (The girl gave birth.)
 (b) **Inkomo ikhululekile.* (*The cow gave birth.) (i.e. of a human being)
 (c) *Inkomo izele.* (The cow calved.)
 (d) **Inkazana izele.* (*The girl gave birth.) (i.e. of an animal)

Palmer (1981: 89) notes that there are few absolute or real synonyms in natural languages. Similarly Ullmann (1962: 142) asserts: "Very few words are completely synonymous in the sense of being interchangeable in any context without the slightest alteration in objective meaning, feeling-tone or evocative value."

Yule (2006: 104) explicates that it is usually possible to find some nuances that separate synonyms or contexts in which one of the lexemes can appear, but the other(s) cannot. Therefore, when this fact is taken into consideration, it becomes clear that absolute synonymy is rare, because synonyms often have different distributions along a number of parameters. Synonymous forms differ in terms of dialect, formality, language, style, emotion, evaluative meanings, collocation restrictions, connotation and grammaticality among others. Ndebele is particularly rich in synonyms for the historical reason that its vocabulary has come from various languages, among them, Nguni varieties, English, Afrikaans, Kalanga, Shona and many others.

Synonymy is sometimes seen as a special type of hyponymy. It can be defined as symmetrical hyponymy (Yule 2006: 105). Yule describes this relationship using this sentence frame: If X is a hyponym of Y and Y is a hyponym of X, then X and Y are synonyms. For example, *scarlet/crimson* is a hyponym of *red*, and *buy* a hyponym of *get*, implying or entailing that these are synonyms.

3. Hyponymy, meronymy and synonymy and the treatment of meaning in ISN

When definitions in ISN are considered, it emerges that hyponymy has been extensively and effectively used in defining headwords. How ISN compilers have utilized hyponymy in their definitions, can be illustrated by the following selected ISN definitions of semantic sets:

- (5) **umlaza** bz 3. *Umlaza* ngamayezi angahlangananga.
 (**cirrus** noun cl 3. *Cirrus* are clouds that are scattered in the sky.)
- (6) **indlondlo** bz 9. *Indlondlo* yimamba eluhlaza ehlala ezihlahleni elolaka kakhulu.
 (**green mamba** noun cl 9. A *green mamba* is a green type of snake of the mamba family that is very vicious and lives in trees.)

Many headwords in ISN are defined using the notion that hyponyms are a kind or type of a specific hypernym. The definitions of the headwords provide the hypernym together with various distinguishing hyponymic features. Although only a few examples of such definitions can be given, it appears that hyponymy forms the core relationship within the dictionary.

An evaluation of the treatment of meaning in ISN reveals that the definitions of headwords have been structured in a manner reflecting the extensive application of meronymy too. The definitions of the following ISN entries show the use of meronymy:

- (7) **ixolo** bz 5. 1 *Ixolo* yingxenye yesihlahla evikela umzimba waso. 2 ...
(**bark** noun cl 5. 1 *Bark* is the outer covering of a tree that protects its trunk. 2 ...)
- (8) **ilunga** bz 5. 1 *Ilunga* yingxenye yolutho, ikakhulu imfe. 2 ...
(**joint** noun cl 5. 1 A *joint* is a part of a something, for example the stem of the sweet reed. 2 ...)
- (9) **isiduku** bz 7. Ingxenye yenduku okutshaywa ngayo esakhanda layo *yisiduku*.
(**knob** noun cl 7. The thicker round end of a knobkerrie that is used for hitting is the *knob*.)

Following the application of the notions of hyponymy and meronymy in the evaluation of the treatment of meaning in ISN, one notes that these sense relations are a useful and an important defining format. Using hyponymy and meronymy facilitates defining the headword. According to Landau (2001: 162), the definition must define and not just speak about the word or its usage. It must answer the question "What is it?" directly and immediately. The definition must not fail in its basic purpose of giving users enough direct and immediate information to enable them to surmise, at least approximately, its meaning in context.

For example, to say *ulude* (greens) *ngumbhida* (are a type of vegetable) or *umthala* (the thick muscles) *yingxenye yolusu* (are part of the stomach of cattle), form the first and most important part of this essential principle of defining. Various distinguishing features of *ulude* as a type of vegetable and *umthala* as part of the stomach of cattle further explain and define them. This approach to defining decomposes the meaning of a word and describes it with a meaningful sequence of words, which makes the definition more accessible and specific. It leads to accurate and precise definitions, which are the aim of every lexicographer. Moreover hyponymy and meronymy help the lexicographer adhere to good defining practices, especially that of the prioritization of its characteristics, where the most essential elements of meaning come first while the more incidental ones come later. When hyponymy and meronymy is used in defining, the essential elements of the meaning of the headword are stated immediately and directly. For example, to define *ikhasi* (page) as *ingxenye yogwalo yokulobela* (the internal part of a book) or *umkhiwa* (wild fig) as *yisithelo seganga* (a wild fruit tree), immediately and directly brings forward the essential ele-

ments of the meaning of *ikhasi* and *umkhiwa* respectively. Hyponymy and meronymy guarantee simple and straightforward definitions characteristic of a user-friendly dictionary. An application of hyponymy and meronymy in defining helps one adhere to the essential principles of defining, guarding against circularity and repetition.

Hyponymy and meronymy also prove to be a solution to defining taboo and/or offensive words that appear to have been a problem for ISN compilers. For example, the lexicographer can adopt an approach of entering the culturally acceptable synonyms of the taboo and/or offensive word as the main entries that carry the entire definition and use these synonyms as cross-references indicating their tabooedness and/or offensiveness. The definitions of the main entries could be structured as shown in the following examples proposed by the author to illustrate the usefulness of hyponymy and meronymy in defining taboo and/or offensive words.

- (10)(a) **iphambili** bz 5. *Iphambili* yisitho sensitha sezidalwa zesilisa loba ezinduna esigcina njalo sidale inhlanyelo loba ubudoda. FAN *isende*.
(**private part** noun cl 5. A *private part* is a reproductive organ of males that stores and produces semen. COMP *isende*.)
- (b) **isende** bz 5. *Isende* ligama elingahloniphiyo elitsho iphambili alithiwa bhadla emphakathini. BONA *iphambili*.
(**testicle** noun cl 5. *Testicle* is a taboo word that refers to a private part of males, and it cannot be used in public. SEE *iphambili*.)

Hyponymy and meronymy yield definitions that are more explicit about and representative of the content aspect of the entry since they allow the lexicographer to be as specific as wished. They are fundamental to semantic analysis of lexemes in lexicography because they express the basic logical relationships that are represented widely throughout the lexicon.

ISN definitions reveal that synonymy has been extensively used in defining a number of headwords. A look at a selected sample of verbs and taboo words in ISN shows the extent to which ISN compilers have used synonymy as a defining format:

- (11) **-zonda** sz mwa. 1 ... 2 Ukuzonda yikucaphuka lokufuthelana emoyeni.
(**hate** transitive verb. 1 ... 2 *Hate* is to have a strong feeling of dislike and emotional frustration.)
- (12) **isende** bz 5. *Isende* liphambili lowesilisa njalo kulapho okulahlanyelo yobudoda. Igama leli liyahloniphisa alikhulunywa emphakathini nje.
(**testicle** noun cl 5. A *testicle* is a male person's reproductive organ where semen is stored. This term is a taboo word; it cannot be used in public.)
- (13) **-thwala** sz gmwa. Nxa izulu *lithwala* liyabe lihloma selilungela ukuna.
(**condense** intransitive verb. When clouds *condense* they would be building up ready for a shower of rain.)

For illustrative purposes these three definitions suffice to support the observation that synonymy was extensively used as a defining format for quite a number of headwords in ISN.

To assess the contribution(s) that synonymy makes to user-friendliness, accessibility of definitions and semantic analysis of headwords in lexicography, it must be considered whether synonym definitions contribute to the essential principles of defining, ensuring the compilation of clear, systematic, straightforward and specific definitions, which are culturally and contextually appropriate and relevant, making dictionaries user-friendly and accessible. For the assessment of the use of synonymy in dictionaries, it is necessary to redefine what a synonym is in lexicographic contexts. The definition to be adopted for lexicography is that of Wiegand (1983: 146) which views lexical synonymy as the occurrence between two lexical items when the rules of usage are to such an extent similar that the rules of reference and predication are the same in habitual texts.

With regard to standard or smaller general monolingual dictionaries, the prevailing view is that synonym definitions are necessary. Landau (2001: 398) states that synonym definitions are not necessarily a mark of slack or inept editing, rather they may be the best solution to the problem of too little space. Sharing the same opinion, Svensén (1993: 119) also notes that the use of synonyms and near synonyms as definitions save space. They are entirely valid when the need for semantic precision is not too great, but one has to be on guard against polysemous synonyms. Words used in the context of the definition must be unambiguous. In other words, as Hanks (1987: 116) rightly puts it, the definition must be clear and unambiguous. If a word used in the definition is polysemous, the particular sense in which it is intended must be made clear by the rest of the definition (Landau 2001: 170). However, if it is considered that synonym definitions are only one-word definitions, there is no possibility for a definition that will explain the particular sense intended for the polysemous word, as is the case with the following ISN entries *ishinga* (rascal) and *amachaphazi* (larvae). For these entries the polysemous synonyms *ongezwayo* and *izibungwana* respectively are presented as complete definitions. The word *ongezwayo* can be interpreted in two ways, meaning either someone who is naughty or stubborn (*isiqholo*) and/or someone who is deaf (*isacuthe*), which is the primary sense of *ukungezwa*. The word *izibungwana* can be taken to signify either the early stages of an insect's development or small insects (in the diminutive sense), which is the primary meaning of *izibungwana*. A quick glance through ISN reveals that avoidance of ambiguous words was often overlooked. Synonyms cannot be used as definitions unless they are disambiguated through exemplification or contextualization. Although the sense of these defining words captures the meaning of the headwords they define, the use of polysemous synonyms ambiguates the definition, compromising the dictionary's accuracy and accessibility of its definitions. If it is considered that space is limited in most dictionaries, especially in dictionary types like ISN

which are general-purpose dictionaries, examples are kept as minimal as possible, hence synonym definitions prove to be an inexact way of specifying the meaning of words.

Svensén (1993: 119) stresses that if a synonym is polysemous, it must not be given as a complete definition, but must be disambiguated. Adopting Svensén's stance of disambiguation through the addition of further synonyms is also problematic not only because it wastes space as the words used in the definition must also be entered and defined, but also because it compromises precision since the words used in the definitions are partial synonyms. In addition to these problems, this approach takes up users' time for they have to go to at least three or more distinct but related entries.

This type of synonym definitions furthermore yields culturally unacceptable definitions that are inaccurate and imprecise since these words cannot be used interchangeably in all contexts. Evaluated according to the adopted lexicographic definition of lexical synonymy that upholds the rules of usage, reference and predication, ISN definitions identified as culturally inappropriate and unacceptable are those that use a class of synonyms whose contextual differences are great and their interchangeability impossible, causing embarrassment to speakers and writers. In this way, the value of the dictionary is compromised as a trusted and respected repository of facts about a language and its culture, and as a decisive authority on "good" as opposed to "bad" usage. In addition to this, the use of ambiguous synonym definitions causes the dictionary to fail as a guardian of the moral and ideological values of the society (Béjoint 2000: 124). A dictionary must reflect these societal values in some aspects of its macrostructure and microstructure, particularly the selection and presentation of headwords, the wording of definitions and the choice of examples. Apart from yielding culturally inappropriate and unacceptable definitions, use of synonym definitions in ISN lead to unnatural circumlocutory definitions. This can be illustrated by the following ISN examples:

- (14)(a) **-khwela** sz mwa. Leli ligama lokuhlonipha elitsho ukuzeka. FAN *zeka*.
(**mate** transitive verb. This is a taboo word that refers to having sexual intercourse. COMP sexual intercourse.)
- (b) **-zeka** sz. Uma inkunzi ikhwela inkomokazi iyabe iyizeka.
(**have sexual intercourse** transitive verb. When a bull is covering a cow, it will be *having sexual intercourse*.)

The above definitions do not define the entry by answering the question "What is it?" directly and immediately. These two words cannot be used interchangeably since *-khwela* (cover) only refers to animals and *-zeka* (have sexual intercourse) is a taboo word used only in respect of humans; therefore they cannot define each other.

Commenting on circularity, Landau (2001: 157) notes that, since the primary purpose of a dictionary is to inform users of the meaning of words, any-

thing that denies them this information makes a dictionary defective. Circularity does not just complicate the consulting process, it makes it impossible. No amount of diligence on the part of users can overcome the barrier of circularity (Svensén 1993: 33). The use of synonym definitions often results in inaccurate definitions. Near-synonyms compromise precision and at times the cultural relevance of the definition, leading to miscommunication as in defining *ukukhuluphala* (being stout) with *ukunona* (being fat) or vice versa. Although these words refer to the same quality, the former refers to humans only while the latter refers to animals only. A similar case is *-zaca* (become slim) which in ISN has been cross-referenced to *-caka* (become thin). These two headwords cannot be used interchangeably since the former refers to humans only while the latter refers to animals only. In this way, the value of the dictionary as a tool for improving communication is compromised.

A definition must not fail its basic purpose of giving users enough direct and immediate information to enable them to surmise, at least approximately, its meaning in context. Definitions compromising the principle of defining the entry lengthen the consultation process by sending users to more than one entry, a procedure which most, if not all, users are unwilling to follow. The following ISN definitions could be considered as examples:

- (15)(a) **-khathaza** sz mwa. Nxa *ukhathaza* uyabe uhlupha abanye abantu.
(**pester** transitive verb. If you are *pestering* other people, you will be bothering them.)
- (b) **-hlupha** sz mwa. Nxa *uhlupha* uyabe uphatha kabi abanye abantu usenza badubeke.
(**bother** transitive verb. If you are *bothering* people you will be treating them badly making them suffer.)
- (16)(a) **inhlanhla** bz 9. 1 ... 2 *Inhlanhla* lithamsanqa lokusinda engozini.
(**luck** noun cl. 19. A stroke of *luck* is to be fortunate to survive an accident.)
- (b) **ithamsanqa** bz 5. *Ithamsanqa* yinhlanhla engalindelwanga eyehlela umuntu.
(**fortune** noun cl 5. *Fortune* is unexpected luck which happens to someone.)
- (17)(a) **-depha** sz gmwa. Ukudepha yibude kumbe ukutshona kolutho phansi.
(**to have depth** intransitive verb. How much *depth* something *has*, is how far it goes down into the earth.)
- (b) **-tshona** sz gmwa. 1 ... 2 Ukutshona komgodi yisilinganiso sawo sokudepha usiya phansi.
(**to be deep** intransitive verb. 1 ... 2 How *deep* a hole *is*, is how far it goes down from the surface.)

The above definitions compromise the principle of defining the entry by sending users to other entries which they also find defined circularly. Examples (15)–(17) are defined by their synonyms, *-hlupha* (bother), *inhlanhla* (luck) and

-depha (deep) respectively. Béjoint (2000: 193) reiterates that simple equivalences or synonyms make the definitions inaccessible as compared to those that decompose the meaning of a word using an analytic definition which is more accessible to users.

Wiegand (1983: 150) argues that the use of synonyms and antonyms as definitions can be acceptable in dictionaries of synonyms or antonyms or when for one reason or another the analytic type does not work in dictionaries, which are not dictionaries of synonyms or antonyms. In his view synonym and antonym definitions are not definitions at all. According to him, synonyms belong in the semantic commentary yet not in the semasiological, but in the onomasiological part where they are best listed under an explicit cross-reference for synonyms, for example, 'FAN' or 'COMP' which can be understood as semantic commentaries functioning as cross-references.

Synonyms are a necessary part of the information that a dictionary must contain, hence they must be included. Fieldwork findings by Nkomo (2003: 44) reveal that synonyms are an important information category to the ISN users as 60% of the informants interviewed expect to find information on synonymy in Ndebele dictionaries. This shows synonyms to be an important component of the ISN microstructure. In including them, they must be printed in such a way that the letter type differs clearly from that of the meaning explanation. In as much as meaning is relational, a structural semantic definition, while important, still does not contain the full information, it cannot serve as a complete definition. In this regard, synonyms are not definitions, but serve as supplements to the lexicographic meaning explanation. Synonym definitions give users intralinguistic knowledge and deny them extralinguistic knowledge. As it is difficult to use a word without some extralinguistic knowledge, it must be combined with intralinguistic knowledge if users are to genuinely claim they know the entry. Synonyms are best used as cross-references, which help, apart from saving users' time, avoid compromising the essential defining principles and avoid unnecessary repetition of definitions.

From the outset it is important to establish and follow a set way of handling synonyms. The selection of synonyms as entries in ISN often violates its style manual, which stipulates that synonyms will be avoided as much as possible in definitions and will be included after the definition proper. With regard to headword selection, only indigenous synonyms will carry the entire definitions with the borrowed ones being used as cross-references. Compromising the stipulations of the style manual lead to inconsistencies both in the selection of synonyms as entries and their use as cross-references, making it difficult to master the microstructure and to use the dictionary efficiently.

As Zgusta (1971: 248) rightly points out, the presentation and treatment of the information need to be consistent if the dictionary is to be user-friendly and accessible. A cursory look through ISN reveals inconsistencies in the selection of synonyms as headwords and their use as cross-references. In some cases users are sent to entries that are not entered in the dictionary or are sent back to

the entries with which they started. Examples (15)–(17) show such circular definitions.

Synonym definitions also compromise the defining principle of using common, everyday words. Landau (2001: 157) states that the lexical definition should not contain words "more difficult to understand" than the word being defined. Using "more difficult words" to define the entry wastes space since every word used in the definition must be entered. For these it would be necessary to formulate a clearer definition that will simplify the "more difficult entry". When using a general monolingual dictionary, users would expect that, if they do not know the meaning of a word used in a definition, they can look up that word and find it defined. The use of words which are not common everyday words compromises the naturalness of definitions, making them more difficult to understand (Moon 1987: 89). As far as ISN is concerned, the principle of defining using common, everyday words is compromised. When these "more difficult words" are entered and defined, in some cases they are defined circularly and in some cases users are referred back to the very entry with which they started. For example, **inhlanhla** (luck) is defined using the less common word *ithamsanqa* (fortune), which as headword is explained circularly with the word it defines. In fact, the first definition of **inhlanhla** accurately captures the meaning of these two absolute synonyms; it will therefore be proper just to use *ithamsanqa* as a cross-reference. Another example is **-tshona** (to be deep) which is defined using the less common word *-depha* (to have depth) which as headword is explained circularly with the word it defines. The practice of defining using less common words makes it difficult for users to understand the definitions.

4. Relationship between sense relation, semantic field theory and componential analysis

A close look at hyponymy reveals that the idea of semantic field theory is also found in it. On a very general level, one can say that the words in a semantic field, though not synonymous, are all used to speak about the same general phenomenon. There is a meaning inclusion relation between the items in the field and the field category itself. The study of word meaning reveals that the lexicon of a language is not simply an unorganized list of words. Semantic relations such as hyponymy, meronymy, synonymy and antonymy are all involved in semantic fields, serving to link certain words with other words. Hyponymy is often important in displaying the semantic relation between items within a lexical field. In short, co-hyponyms form a semantic field of a particular field. For example, the semantic field of *inyoka* (snakes) has as some of its members *inhlathu* (python), *imamba* (mamba), *iphimpi* (cobra), *ibululu* (puff adder) to name a few. A semantic field is also composed of meronyms. For example, the fingers make up a semantic field and so also the parts of the body. The notion of semantic field can therefore be extended to any set of terms with

a close relation in meaning, all of which can be subsumed under the same general label. The field then becomes the context within which to establish meaning relations. A semantic field makes it possible to ascertain meaning relations between words sharing the semantic space.

Not only does hyponymy relate with semantic field theory, it also relates with componential analysis in terms of logical inclusion. The idea of entailment involved in hyponymy is also evident in componential analysis, which is an alternative way of establishing, or at least confirming sense relations. Components or semantic features are the factors, or contrastive elements, which it is necessary to posit in order to account for all significant meaning relations (Leech 1969: 20). Componential analysis is founded on the notion of semantic contrast — antonymy. The componential notation provides a simple characterization of semantic relationships of logical inclusion-hyponymy, logical exclusion-antonymy and synonymy seen in componential synonymy. Componential analysis goes some way in explicating semantic relations such as hyponymy, synonymy and antonymy. Semantic components are hypernyms of particular co-hyponyms, for example +human, +male and +female.

5. Conclusion

The close-knit relationship between sense relations, semantic field theory and componential analysis is evidence enough that these approaches are necessary parts of a semantic theory worth adopting in lexicography as a holistic approach to the analysis of the meaning of lexemes. It can help account for more than a selection of semantic facts of lexemes in lexicography, and above all help adhere to the essential principles of defining. A holistic approach to meaning can go a long way in approaching meaning from a lexicographer's point of view. Reference on its own is not enough to specify the meaning of an entry in lexicography, for meaning encompasses both the intra- and extralinguistic elements of the entry. Hyponymy and meronymy form the basis of the analytic and classical definition, therefore lexicographers ought to consider them in their formulation of dictionary definitions in order to adhere to the essential principles of defining. On the other hand, as Landau (2001: 398) argues, synonym definitions waste, instead of saving space. They compromise the essential principles of defining, particularly that of defining the entry addressing the question "What is it?" directly and immediately, which in turn leads to circular, imprecise, inaccurate and culturally irrelevant definitions. In this regard, synonym definitions compromise the dictionary's user-friendliness and accessibility. They fail in their basic purpose to give users enough immediate information to enable them to surmise, at least approximately, the entry's meaning in context. In addition to the above, use of synonym definitions in monolingual general-purpose dictionaries lead to miscommunication. They prove to be an inexact way of specifying meaning, especially in dictionary types which are not synonym dictionaries. They take up time as users are sent to at least more than

two entries, therefore disrupting the activity in which they are engaged. Hyponymy and meronymy prove to be productive and useful defining formats that help the lexicographer adhere to the essential principles of defining, in this way facilitating the user-friendliness and accessibility of the dictionary. Hyponymy and meronymy can also be useful formats for defining taboo and/or offensive words which proved problematic for ISN compilers and also prove to be a challenge in languages having the *hlonipha* variety (employed to show respect).

References

- Béjoint, H.** 2000. *Modern Lexicography: An Introduction*. Oxford: Oxford University Press.
- Carnap, R.** 1942. *Introduction to Semantics*. Cambridge, MA: Harvard University Press.
- Carter, R.** 1998. *Vocabulary: Applied Linguistic Perspectives*. London: Routledge.
- Crystal, D.** 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- Hadebe, S. et al. (Eds.)**. 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hanks, P.** 1987. Definitions and Explanations. Sinclair, J.M. 1987: 116-136.
- Jackson, H.** 1988. *Words and their Meaning*. London/New York: Longman.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Leech, G.N.** 1969. *Towards a Semantic Description of English*. Harlow: Longmans.
- Lyons, J.** 1977. *Semantics. Volume 1*. Cambridge: Cambridge University Press.
- Matthews, P.H.** 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford/New York: Oxford University Press.
- Moon, R.** 1987. The Analysis of Meaning. Sinclair, J.M. 1987: 86-103.
- Nkomo, D.** 2003. *The Microstructure and the Accessibility of Dictionary Information: An Analysis of Isichazamazwi SesiNdebele*. Unpublished B.A. Hons. Thesis. Harare: University of Zimbabwe.
- Palmer, F.R.** 1981. *Semantics*. Second Edition. Cambridge/London: Cambridge University Press.
- Saeed, J.I.** 2003. *Semantics*. Second Edition. Malden, MA: Blackwell.
- Sinclair, J.M. (Ed.)**. 1987. *Looking Up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London: Collins ELT.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford/New York: Oxford University Press.
- Ullmann, S.** 1972. *Semantics: An Introduction to the Science of Meaning*. Oxford: Basil Blackwell.
- Wiegand, H.E.** 1983. Synonyms Appearing in Major Alphabetical Dictionaries of Contemporary German. Wiegand, H.E. 1999. *Semantics and Lexicography. Selected Studies (1976–1996)*: 139-152. Edited by A. Immken and W. Wolski. Tübingen: Max Niemeyer.
- Yule, G.** 2006. *The Study of Language*. Third Edition. Cambridge: Cambridge University Press.
- Zgusta, L.** 1971. *Manual of Lexicography*. The Hague/Paris: Mouton.

The Treatment of Borrowed Nouns in *Isichazamazwi SesiNdebele* and *Isichazamazwi SezoMculo*

Eventhough Ndlovu, *Department of African Languages and Literature,
University of Zimbabwe, Harare, Zimbabwe (endlovu@arts.uz.ac.zw)*

Abstract: This article focuses on the lemmatisation of vowel-commencing borrowed nouns and the allocation of borrowed nouns to noun class prefixes in *Isichazamazwi SesiNdebele*, the first monolingual general-purpose Ndebele dictionary, and *Isichazamazwi SezoMculo*, the first specialised Ndebele dictionary of musical terms. It adopts a comparative approach, also highlighting the controversies surrounding the status of the initial vowel of the prefix or the pre-prefix in Ndebele and other Nguni languages. It further looks at the challenges and limitations of lemmatising the noun using either the initial vowel of the prefix or the initial letter of the noun stem. It is found that there are some inconsistencies in the lemmatisation of vowel-commencing borrowed nouns and the allocation of borrowed nouns to noun class prefixes in the two dictionaries. These inconsistencies impact negatively on the standardisation and treatment of borrowed nouns.

Keywords: LEMMATISATION, MICROSTRUCTURE, MACROSTRUCTURE, USER-PREFERENCE, USER-PERSPECTIVE, USER-FRIENDLY, ACCESSIBILITY, REDUNDANCY, NOUN STEM, NOUN CLASS PREFIX, AUGUMENT, PREFIX PROPER, BORROWED NOUN

Opsomming: Die behandeling van leensubstantiewe in *Isichazamazwi SesiNdebele* en *Isichazamazwi SezoMculo*. Hierdie artikel fokus op die lemmatisering van leensubstantiewe met aanvangsklinkers en die toewysing van leensubstantiewe aan substantiefklasprefikse in *Isichazamazwi SesiNdebele*, die eerste eentalige meerdoelige Ndebelewoordeboek, en *Isichazamazwi SezoMculo*, die eerste gespesialiseerde Ndebelewoordeboek van musiekterme. Dit gebruik 'n vergelykende benadering sodat ook aandag gegee word aan die strydpunte rondom die status van die aanvangsklinker van die prefiks en die pre-prefiks in Ndebele en ander Ngunitale. Dit kyk verder na die uitdagings en beperkings van die lemmatisering van die substantief deur óf die aanvangsklinker van die prefiks óf die aanvangsletter van die substantiefstam te gebruik. Daar word gevind dat daar 'n aantal inkonsekwensies in die lemmatisering van leensubstantiewe met aanvangsklinkers en die toewysing van leensubstantiewe aan substantiefklasprefikse in die twee woordeboeke is. Hierdie inkonsekwensies het 'n negatiewe uitwerking op die standaardisering en behandeling van leensubstantiewe.

Sleutelwoorde: LEMMATISERING, MIKROSTRUKTUUR, MAKROSTRUKTUUR, GEBRUIKERSVOORKEUR, GEBRUIKERSPERSPEKTIEF, GEBRUIKERSVRIENDELIK, TOEGANKLIKHEID, OORTOLLIGHEID, SUBSTANTIEFSTAM, SUBSTANTIEFKLASPREFIKS, VOORVOEG, EGTE PREFIKS, LEENSUBSTANTIEF

Introduction

A brief discussion of the morphological structure of the Ndebele noun, considering the grammatical status of its constituent morphemes, will serve as an introduction to this article. The noun class prefix distinguishes the class to which the noun belongs. It also denotes whether the noun is singular, plural or neuter. Furthermore it is the basis for the derivation of all substantival prefixes and concordial agreements dependent on it.

In Ndebele, as in all other Bantu languages, nouns are divided into different classes. The arrangement of Ndebele noun classes follows the accepted numbering applicable to all Bantu languages. There are 18 classes for Ndebele, with no nouns in classes 12 and 13. Classes 16, 17 and 18 reflect locative nominalisation. Some Ndebele noun class prefixes have allomorphs, which are phonologically conditioned according to the stem types they take. For example, class 7 /*isi-*/ has its allomorph /*is-*/, /*isi-*/ taking consonant-commencing stems such as *isilwane* (lion), *isizwe* (nation), *isisu* (stomach), and /*is-*/ taking vowel-commencing stems such as *isono* (sin), *isothamlilo* (human being) and *isoni* (sinner).

This system of noun classes must be mastered, as it is important for the correct speaking of the language. The noun class prefix is central, being the base of other substantival prefixes and concordial agreements of other word categories. Like other Nguni languages, Ndebele exhibits a peculiar feature in its noun class prefixes, the presence of the initial vowel or pre-prefix, or the augment, as Khumalo (2003: 80) calls it. Some noun class prefixes are just vowels as in classes 1a, 2a and 5. The initial vowel distinguishes Nguni languages from other Bantu languages. Apart from the initial vowel, a simple Ndebele noun is made up of the prefix proper and the stem as shown in *u* (initial vowel) -*mu-* (prefix proper) -*thi* (stem) = *u-mu-thi* (medicine). The initial vowel is the same as the vowel in the prefix proper, for example in /*ĩĩ-*/, /*ũũ-*/ and /*izi-*/. When a noun is used in its evocative sense the initial vowel is discarded, for example *ndoda!* (man!), *mfana!* (boy!) and *mntwana!* (child!). The initial vowel plays a complementary role to the prefix proper, in other words it augments it.

The grammatical status of the initial vowel is highly contentious, as there is no consensus on whether it is a syntactic, semantic or morphological feature. This initial vowel is, however, one of the unique distinguishing features of Ndebele. The controversies surrounding its grammatical status have led to inconsistencies among lexicographers and linguists in the lemmatisation of the noun in Ndebele and other Nguni languages, which share this same feature (Maphosa 1997: 1).

If the central role of the prefix in the derivation of substantival prefixes and concordial agreements is considered, it is undeniable that the subject and object concords, substantival prefixes and qualificative concords ought to be used as central criteria in deciding the class affiliation of borrowed nouns. These concords help clarify and distinguish the class affiliation of each given noun.

The noun stem carries the central meaning of the word and does not change when the noun changes from singular to plural; it remains constant. It does not have any specific meaning until a prefix is attached to it (Hadebe 2006: 87). It is a bound morpheme, as it needs the prefix in order to give it a specific meaning. Ndebele noun stems come in different phonological forms; some are consonant-commencing, for example *in-ja* (dog), *isi-nkwa* (bread), *in-doda* (man), while others are vowel-commencing, for example *is-ono* (sin), *is-ona* (type of weed) and *is-othamlilo* (human being). Phonologically the noun stem in Ndebele ends in a vowel. Ndebele noun stems can be derived from other word categories, for example from verbs, relative stems, and adjectival stems among others. Some noun stems are primitive stems whose derivation cannot be traced, for example *-ntu* for *umuntu* (person), *-fazi* for *umfazi* (woman) and *-hlahla* for *isihlahla* (tree).

In pioneering dictionaries, the lemmatisation of the noun in Ndebele and other related Nguni languages has been on the basis of the initial letter of the noun stem. The inconsistencies in the lemmatisation of the nouns in Ndebele and other Nguni languages are a clear testimony of the lack of consensus on the grammatical status of the initial vowel of the noun class prefix.

Apart from the omission of the initial vowel in vocatives, Von Staden (1973: 170) notes that the initial vowel is obligatorily omitted after possessive concords when they belong to class 1a, in a few instances after the locative possessive concord *kwa-*, in relative constructions, when they appear as second components of compound nouns, when used conjunctively, after negative associative copulatives, in interrogatives (such as in *Ubona muntu bani?* 'What sort of a person are you seeing?'), and when employed as relative stems.

Vowel elision and more commonly vowel coalescence result in the disappearance of the initial vowel. These processes are a means towards retaining the acceptable Ndebele syllable structure of CV or CCV which avoids cases where vowels follow each other in writing. Any one of these processes takes place if CVV, VVC or CCVV syllable structure occurs.

The realisation that the initial vowel is an individualising morpheme, because it distinguishes the meaning of nouns, proves beyond doubt that it is a semantic feature for it denotes class, person, gender and number. The fact that the initial vowel does not appear in the selected phonological conditions above is a clear testimony that it is part of the noun in Ndebele. The semantic value of the initial vowel makes it a necessary and indispensable morpheme that must be considered in lemmatising nouns, especially considering the primary and central role of meaning in lexicography. The semantic value of the initial vowel makes it a central feature that cannot be ignored in a field dealing primarily with meaning; hence, because of this semantic value, the Ndebele noun should be lemmatised using the initial vowel of the prefix.

Lemmatisation of the noun in Nguni languages

A cursory look at selected dictionaries in the Nguni family reveal that the noun

has been lemmatised using the initial letter of the stem. Kropf (1899) lemmatises nouns according to the initial letter of the stem, on the grounds that the use of the initial vowel would crowd the majority of the headwords under the letters *i* and *u* since most of the nouns in these varieties have /i/ and/or /u/ as their initial vowel.

Doke and Vilakazi (1948) also follow the set principle in Nguni of lemmatising the noun using the initial letter of the stem, citing similar concerns and fears as Kropf. Of interest in their justification for this usage is their acknowledgement of the difficulties in identifying the initial letter of the stem. These two pioneering works clearly reveal the challenges of demarcating the noun stem as there are some remarkable inconsistencies in their identification of the initial letter of the stem of similar nouns, for example, Kropf's stem for *impuphu* (mealie-meal) is *-puphu*, and for Doke and Vilakazi the stem is *-mpuphu*. Doke and Vilakazi include the prefix proper /m/. This reveals the challenges that accompany lemmatising according to the initial letter of the stem, especially with regard to nasal-commencing stems.

Doke (1958) continues lemmatising the noun using the stem. Like the pioneers of Nguni lexicography, Nyembezi (1992) also lemmatises the noun according to the initial letter of the stem. Sometimes Nyembezi's initial letter of the stem begins with the prefix proper, revealing, as in the works of the other pioneers of Nguni lexicography, the controversies of what a stem is. Dent and Nyembezi (1969) follow Doke's way of lemmatising according to the initial letter of the stem, though with inconsistencies, since sometimes the stem incorporates the prefix proper, for example, *-ntombi* for *intombi* (maiden) and *-mpi* for *impi* (war). Nkabinde (1982 and 1985) also adopts Doke's way of lemmatising according to the initial letter of the stem. To mark the difference between the stem and the prefix, Nkabinde writes the stem in capital letters and the prefix in small letters.

It is clear that Nguni dictionaries reveal some significant inconsistencies in what a noun stem is. In some cases, the noun stem includes the prefix proper. In other cases, the noun stem is uncertain, especially when it is nasal-commencing. The inconsistencies reveal the challenges of demarcating the noun stem in Nguni and using it for lemmatising purposes.

Lemmatisation of the noun in Ndebele

The bilingual Ndebele dictionary of Pelling (1966) lemmatises the noun using the initial letter of the stem. The first monolingual general-purpose dictionary in Ndebele, *Isichazamazwi SesiNdebele* (henceforth ISN) lemmatises the noun using the initial vowel of the prefix. On the other hand, the first specialised Ndebele dictionary of musical terms, *Isichazamazwi SezoMculo* (henceforth ISM), reveals some inconsistencies in its lemmatisation of the noun. Vowel-commencing borrowed nouns are lemmatised using the initial letter of the stem while indigenous nouns are lemmatised using the initial vowel of the prefix. In the style manual the ISM editors note:

Amagamanhloko angamabizo okwebolekwa aqala ngabonkamisa ethulwe ngalabonkamisa njengokuthi alitho, enkho ... Lapho atholakala khona phakathi kwengcazelo alotshwe ngamayitaliki elandela onsingankamisa /w/ loba /y/ kusiya ngabonkamisa beziqalo.

(Borrowed noun headwords which are vowel-commencing are lemmatised using those vowels as in *alitho; enkho* ... Where they are found in the definition they are written in italics following the semi-vowels /w/ or /y/ depending on the initial vowel of the prefix.)

The method of lemmatising vowel-commencing borrowed nouns is part of the major concern of this article, alongside the use of semi-vowels to break vowel sequencing in the ISN and ISM. The semi-vowel used for each borrowed noun also indicates the class to which that noun has been allocated, for example both *iwolintshi* (orange) and *iyinki* (ink) are allotted to class 5 in ISN. ISM's departure from ISN in the lemmatisation of vowel-commencing borrowed nouns impacts negatively on the standardisation of these terms, leaving users with an uncertain choice on how these should be written. ISN and ISM's use of the semi-vowels /w/ and /y/ to break vowel sequencing might not reflect all users' preference and speakers' usage of these terms, therefore violating acceptability (cf. Gilreath 1993: 91). Hadebe (2006: 183-184) notes that words such as *iwoyili* (oil) and *iwoveni* (oven) are pronounced by some people with the semi-vowel, while the majority does not do so. A form like **iyayini* (iron) seems never to be used. The hyphen is sometimes used to break vowel sequencing not only in writing but also in pronouncing these terms, as in *i-ayini*, *i-inki*, *i-olintshi*, *i-oyili* and *i-oveni*. The use of the semi-vowels /w/ and /y/ affects the essential principles of term creation. For example, the ISM entries *akhusitikhi/iyakhusitikhi* (acoustic), *adaphutha/iyadaphutha* (adaptor) and *ampulifaya/iyampulifaya* (amplifier) sound ungrammatical, violating linguistic correctness (cf. Gilreath 1993: 89). However, there are cases such as *iyuniformu* (uniform), *iyunivesithi* (university) and *iyembe* (shirt) where the use of the semi-vowel is acceptable as long as these borrowed nouns conform to the principles of term creation and the speakers' pronunciation of them.

Significant in these examples from ISM are also the omission of the noun prefix. The treatment of vowel-commencing borrowed nouns as headwords without the noun prefix in ISM, for example *i-akaphela* as *akhaphela* (acapella), *i-ophera* as *ophera* (opera) and *i-eriyeli* as *eriyeli* (aerial), does not reflect their use by Ndebele speakers. They cease to be nouns if they are represented without the prefix. Here they are in their evocative state as interjections.

In their attempt to be normative, the ISN and ISM compilers should have striven to strike a balance between being descriptive and prescriptive in their treatment of vowel-commencing loan nouns. In accepting a loan word, it has to be adopted and adapted to suit the lexical, morphological, phonological and orthographic patterns of the recipient language. According to Béjoint (2000: 101), total prescriptiveness does not work if it is estranged from the realities of usage, as it is the case with ISN and ISM's use of borrowed headwords containing semi-vowels. The same is true of ISM's lemmatisation of vowel-com-

mencing borrowed nouns using the initial vowel of the loan noun and denying them prefixes. This way of lemmatisation overlooks user-perspective because it is contrary to speakers' common usage. Users will therefore be challenged to locate these lemmas in the dictionary. This compromises the dictionary's accessibility which in turn translates into a less user-friendly dictionary. The inconsistencies in the lemmatisation of the noun in ISN and ISM impact negatively on the standardisation of these terms because of the variations in their treatment as headwords.

Apart from violating the above-mentioned terminological principles, the treatment of borrowed nouns presents challenges to users in deducing their word categories, especially considering the fact that ISM does not indicate word categories, granted that it is meant for specialised users. For the users to master the microstructure and continue using the dictionary without challenges, consistency in the treatment of common word categories is essential.

Allocation of borrowed nouns to noun class prefixes in ISN

In the Nguni languages and particularly in Ndebele, each noun belongs to a specific noun class prefix according to the following criteria (Canonici 1995: 14):

- (a) shape of the noun prefix,
- (b) correspondent singular/plural class,
- (c) syntactic behaviour or the concords it requires, and
- (d) semantic content.

Canonici argues that all these criteria must be used in combination to decide the class affiliation of a noun. Applying these criteria to borrowed nouns would therefore help to determine their class affiliation. A close look at most borrowed nouns in Nguni reveal that the majority of them fall under classes 5 and 9. However, there are some which are unambivalent in nature and fall either in class 5 or class 9, for example, among others, *irandi* (Rand) and *isenti* (cent). Testing the class affiliation of these unambivalent loan nouns, by using Canonici's criteria, confirm their unambivalent nature.

Table 1

Entry	Subj Con	Adj Con	Rel Con	Poss Con	Abs Pro	Dem Pro	Quant Pro	Sel
<i>irandi</i> (Rand)	li-/i-	eli-/en-	eli-/e-	la-/ya-	lo-/yo-	leli/le	lo-/yo-	li-/i-/yi-
<i>isenti</i> (cent)	li-/i-	eli-/en-	eli-/e-	la-/ya-	lo-/yo-	leli/le	lo-/yo-	li-/i-/yi-
<i>ithenisi</i> (tennis shoe)	li-/i	eli-/en-	eli-/e	la-/ya-	lo-/yo-	leli/le	lo-/yo-	li-/i-/yi-

Abbreviations: **Subj Con:** Subject Concord, **Adj Con:** Adjectival Concord, **Rel Con:** Relative Concord, **Poss Con:** Possessive Concord, **Abs Pro:** Absolute Pronoun, **Dem Pro:** Demonstrative Pronoun, **Quant Pro:** Quantitative Pronoun, **Sel:** Selector

A cursory look at ISN reveals that most of the borrowed nouns are slotted into class 5, yet, as will be shown later in this section, they belong to class 9 allomorph or a proposed class 9a, for example: *iwigi* (wig), *ithara* (tar), *itsheki* (cheque), *itafula* (table), *isepa* (soap), *irula* (ruler), *inalithi* (needle), *ipenda* (paint), *ifoni* (telephone), *ibhasi* (bus), *ibhenki* (bank), *ivayisi* (vice), *idivosi* (divorce), *isemantiki* (semantics), *iwaranti* (warrant), *isophurano* (soprano), *ikhiye* (key), *iyunifomu* (uniform), *isawa* (sour milk), *irekhodi* (record), *inayiloni* (nylon), *iKhisimusi* (Christmas), *ifulegi* (flag), *ikhofi* (coffee), *ifeni* (fan), *ibhata* (butter), *irisithi* (receipt), *igirisi* (grease), *ipheni* (pen), *ithoyi* (toy), *iselifoni* (cell phone), *irubhu* (rubber rope), *iyunivesithi* (university), *iyembe* (shirt), *ikhoma* (comma), *igirineda* (grenade), *ifenitsha* (furniture), *irediyo* (radio), *idisikoti* (discord) and *ibhera* (bell).

According to Canonici (1995: 22), each class commands its own concordial system, dependent nominals and predicates, and these ought to be used alongside the other mentioned criteria to decide the class affiliation of borrowed nouns. There are, however, some Ndebele noun classes that share their concordial agreement. In these cases, testing for class affiliation must consider other tests such as shape of the prefix and correspondent singular/plural class further to ascertain the class affiliation of the nouns in question. Examples are classes 1, 1a and 3, 2 and 2a, 4 and 9, class 9 allomorph or a proposed class 9a, as well as classes 8 and 10. Classes 1 and 1a share concords because 1a is an allomorph of class 1. The same is true for 2 and 2a, 9, 9a and the proposed class 9 allomorph. For classes 1, 1a and 3, 4 and 9, class 9 allomorph or a proposed class 9a, shape of the prefix and correspondent singular/plural class ought to be considered in these cases further to ascertain the class affiliation of the nouns in question. Class 5 is a singular class which is morphologically represented as /ili-/, this class has its allomorph /i-/ which is phonologically conditioned, mainly taking multisyllabic stems. The subject concord for this class is /li/. It realises its plurals in class 6.

Class 9 is a singular class with most of its plurals in classes 10 and 6. Examples of class 9 nouns that realise their plurals in class 6 are, among others, *inkomokazi* (cow), *inkomitsho* (cup), *intombazana* (young girl) and *indoda* (man). It is morphologically realised as /iN-/, with the capital *N* standing for either the nasal /n/ or /m/. Following the entering of borrowed nouns in Ndebele, there seems to be a growing new class of borrowed nouns which simply prefix /i/ without the nasal to the loan stem (Canonici 1995: 14). The noun class system of Ndebele already has this prefix in class 5 as an allomorph and this has tempted linguists and lexicographers to think that nouns that take /i-/ belong to class 5, yet Canonici's test for class affiliation reveals that these borrowed nouns belong to class 9 allomorph or a proposed class 9a.

For purposes of diminishing the number of classes, the proposed class 9a can be used as an allomorph for class 9 which will then have its allomorph /i-/ similar to other classes which have allomorphs, for example class 1 which has its allomorph /um-/. Canonici's proposal can even be adopted, as the Ndebele noun class system already reflects the class *na* as in class 1a. What is essential is

to realise that some borrowed nouns which take /i-/ as their prefix are not class 5 nouns. Instead, they share the same concordial system, dependent nominals and predicates with class 9.

The above realisation can be substantiated using the following examples:

- iringi* (ring)
I-iringi i-lerusi * *I-iringi li-lerusi* (The ring is rusted).
I-iringi ya-mi * *I-iringi la-mi* (My ring).
- ijemu* (jam)
I-jemu i-litshukela * *I-jemu li-litshukela* (Jam is sweet).
I-jemu ya-mi * *I-jemu la-mi* (My jam).
I-jemu e-bomvu * *I-jemu eli-bomvu* (The red jam).
- isenta* (centre)
I-senta i-banzi * *I-senta li-banzi* (The centre is wide).
I-senta e-banzi * *I-senta eli-banzi* (The wide centre).

These nouns have their plurals in class 6, just as the majority of those of class 5. However, it should be noted that, as illustrated before, class 9 also has some of its plurals in class 6. An application of Canonici's criteria reveals that the majority of the borrowed nouns in ISN are wrongly placed in class 5. Instead, most of them belong to the class 9 allomorph or the proposed class 9a, as reflected by their syntactic behaviour or the concords they require (see Table 2).

Table 2

Entry in ISN	Subj Con	Adj Con	Rel Con	Poss Con	Abs Pro	Dem Pro	Quant Pro	Sel
<i>iwigi</i> (wig)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ithara</i> (tar)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>itsheki</i> (cheque)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>itafula</i> (table)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>isepa</i> (soap)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>irula</i> (ruler)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>inalithi</i> (needle)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ipenda</i> (paint)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ifoni</i> (telephone)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ibhasi</i> (bus)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ibhenki</i> (bank)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ivayisi</i> (vice)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>idivosi</i> (divorce)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>isemantiki</i> (semantics)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>iwaranti</i> (warrant)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>isophurano</i> (soprano)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ikhiye</i> (key)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>iyunifomu</i> (uniform)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>isawa</i> (sour milk)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>irekhodi</i> (record)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>

Entry in ISN	Subj Con	Adj Con	Rel Con	Poss Con	Abs Pro	Dem Pro	Quant Pro	Sel
<i>inayiloni</i> (nylon)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>iKhisimusi</i> (Christmas)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ifulegi</i> (flag)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ikhofi</i> (coffee)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ifeni</i> (fan)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ibhata</i> (butter)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>irisithi</i> (receipt)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>igirisi</i> (grease)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ipheni</i> (pen)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ithoyi</i> (toy)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>i-/yi-</i>
<i>iselifoni</i> (cell phone)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>i-/yi-</i>
<i>irubhu</i> (rubber rope)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>iyunivesithi</i> (university)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>iyembe</i> (shirt)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ikhoma</i> (comma)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>igirineda</i> (grenade)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ifenitsha</i> (furniture)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>irediyo</i> (radio)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>idisikoti</i> (discord)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>
<i>ibhera</i> (bell)	<i>i-</i>	<i>en-</i>	<i>e-</i>	<i>ya-</i>	<i>yo-</i>	<i>le</i>	<i>yo-</i>	<i>iphi/yiphi</i>

Abbreviations: **Subj Con:** Subject Concord, **Adj Con:** Adjectival Concord, **Rel Con:** Relative Concord, **Poss Con:** Possessive Concord, **Abs Pro:** Absolute Pronoun, **Dem Pro:** Demonstrative Pronoun, **Quant Pro:** Quantitative Pronoun, **Sel:** Selector

If the test for class affiliation in Table 1 is used, it becomes clear that these ISN entries belong to class 9 allomorph or Canonici's proposed class 9a. The following concordial agreements are derived:

(1) Adjectival Concords

Noun Class Qualificative Formative + Noun Class Prefix = Adjectival Concord
 9/9a *a* *i-* *en-*

(2) Relative Concords

Noun Class Qualificative Formative + Noun Class Prefix = Relative Concord
 9/9a *a* *i-* *e-*

In the derivation of the adjectival and relative concords in (1) and (2), vowel coalescence takes place when the qualificative formative, *a-*, is prefixed to the full prefix. For the relative concord, all nasal consonants and vowels that follow them are elided in its derivation.

(3) Possessive Concords

Noun Class Subject Concord + Qualificative Formative = Possessive Concord
 9/9a *i-* *a* *ya-*

In deriving the possessive concord in (3), the subject concord is consonantised

and becomes a semi-vowel /y/. This happens to retain the acceptable Ndebele syllable structure.

(4) Selector Concorde

Noun Class	Subject Concord + Selector	Roots = Selector Concord
9/9a	i-	-phi iphi/yiphi

In (4), the selector concord is derived from the copulative prefix /yi/.

(5) Absolute Pronoun

Noun Class	Subject Concord + Pronoun	Formative = Absolute Pronoun
9/9a	i	o yo-

In deriving the absolute pronoun in (5), the subject concord is consonantised and becomes a semi-vowel /y/. As in (3), consonantisation takes place to retain the acceptable Ndebele syllable structure.

(6) Demonstrative Pronoun

Noun Class	Demonstrative Formative + Noun Class	Form = Demonstrative Pronoun
9/9a	la	i- le

In (6), vowel coalescence takes place between the final vowel of the demonstrative formative and the initial vowel of the prefix. In cases where the prefix is followed by nasal consonants, the nasal and the vowel that follow it are elided.

(7) Quantitative Pronoun

Noun Class	Subject Concord + Pronoun	Formative = Quantitative Pronoun
9/9a	i-	o yo-

In (7), consonantisation of the subject concord takes place to retain the acceptable Ndebele syllable structure.

The classes dealt with here regarding borrowed nouns are traditionally problematic classes; hence, this problem is not unique to Ndebele. Zawawi (1979: 127), writing about Swahili, suggests that such nouns may first be incorporated into class 9, and may later be recategorised as class 5, the latter having the advantage of distinguishing singular from plural. Eastman (1991: 61 ff.) argues the reverse: that borrowed nouns should start out in class 5 and should later be recategorised as belonging to class 9.

The confusion of classes for borrowed nouns has a negative influence on the language since the mastering of the noun class system is central to speaking Ndebele correctly. Grammatical information has to be as accurate as the definitions in the dictionary. For a language like Ndebele that lacks comprehensive documentation and grammar books, the inclusion of grammatical information cannot be underestimated. Because of this, the centrality of the noun class prefix makes it essential to give the correct class to each noun, since it is important

for the derivation of the subject and object concords, other dependent substantival prefixes and qualificative concords in languages like Ndebele.

The merits and demerits of lemmatising nouns using the stem

Lemmatisation of nouns in the ISN and ISM cannot be discussed without focusing on user-perspective, an issue that has emerged as one of the important lexicographical subjects in the selection and treatment of lexical items, in this case the adoption and adaptation of loan words. As Prinsloo (2001: 18) notes, a dictionary should not reflect the attitude of the lexicographer but should rather be aimed at the specific needs of well-defined target users. It must also be compiled according to the needs and usage skills of the target users. It is against this background that the merits and demerits of lemmatising the noun using the initial letter of the stem or the initial vowel of the prefix are discussed.

There seems to be more demerits than merits in using the initial letter of the stem for lemmatising the noun in Ndebele and other Nguni languages. According to Maphosa (1997: 27), one advantage of lemmatising according to the initial letter of the stem is that related nouns are entered in the same place, that is, nouns sharing semantically related stems are grouped together. However, users usually do not consult a dictionary in search of semantically related words. They consult a dictionary when they come across separate words whose meaning and use they want to clarify. Furthermore, if thesauruses can equally and effectively list semantically related terms together, lemmatisation for this reason can be left to them. The naturalness of the language is compromised when lemmatisation takes place according to the initial letter of the stem since the initial vowel of the prefix is the one defining and distinguishing feature of Nguni languages.

The systematic ordering of nouns using the initial letter of the stem to achieve a semantic ordering of entries gives the impression that language consists of strings of semantically related words, which is not the case. In fact, some of these identical stems are semantically unrelated. For example:

- zulu *i-zulu* (rain) vs -zulu *u-m-zulu/a-ba-zulu* (Zulu person/Zulu people).
- fo *u-m-fo aba-fo* (enemy/enemies) vs -fo *i-si-fo/i-zi-fo* (infection/infections).
- tshe *i-li-tshe/a-ma-tshe* (stone/stones) vs -tshe *i-n-tshe/i-zi-ntshe* (ostrich/ostriches).
- ngane *i-ngane* (child) vs -ngane > *u-m-ngane* (friend).
- thebe *i-si-thebe* (eating vessel) vs -thebe > *i-thebe* (hip bone).

Maphosa (1997: 23) also notes that the use of the initial letter of the stem in lemmatising noun headwords in Nguni dictionaries enhances the quality of the layout of these dictionaries. In as much as the use of the initial letter of the stem allows for the even distribution of headwords in the dictionary across all the

letters of the alphabet, it is however a natural characteristic of Nguni languages that the majority of their entries belong to certain letters. The major limitation of the system of lemmatising nouns using the initial letter of the stem is that it compromises the naturalness of the language by denying the noun its major distinguishing feature. The stem is also not easily identifiable as reflected within and among dictionaries of Nguni languages that often contain inconsistencies in and variations of the stems for similar nouns. This in turn renders the method less user-friendly for users might look for the entry under the wrong letter, wasting their time. It also makes the dictionary less accessible, because, in most cases when users consult dictionaries, they would normally be engaged in some linguistic activity, for example translation, which they would not want to interrupt for too long. In addition, this method is not in line with common usage of the noun in Nguni. In everyday speech or interaction, speakers of Nguni languages do not use the noun stem but the complete noun with its prefix and stem. It is only in evocative and interrogative usage that they drop the initial vowel of the prefix.

Another major limitation of using the initial letter of the stem in lemmatising the noun is that it requires prior linguistic knowledge and training for users to access the entry for which they would be searching. This same linguistic knowledge and training would be required from lexicographers to lemmatise nouns according to their stems.

According to Doke (1993: 196), stems are not words, because they are bound morphemes, not having meaning on their own. As the noun class prefixes, they cannot stand on their own, and hence they cannot be seen and used as entries. The noun stem needs a prefix to enable it to distinguish class, number, gender and person, so as to convey its full meaning. It is only when stems are attached to the prefix that they become words. In common usage, users will never come across a text that refers only to the noun stem. A noun consists of two components; therefore the initial vowel of the prefix ought to be considered in lemmatising nouns.

The above limitations of lemmatising using the initial letter of the stem all point to the dictionary's compromise of user-friendliness because this method impacts negatively on the relative ease with which information can be accessed in the dictionary. It leads to an ambiguous retrieval of information on both the macro- and microstructural levels of the dictionary.

The merits and demerits of lemmatising using the initial vowel of the prefix

Lemmatising noun headwords using the initial vowel of the prefix is more user-friendly, making the information in the dictionary more accessible to users. It provides a mode of access that is convenient for and common to them. Consulting a dictionary does not require any prior linguistic knowledge and training for users, therefore making it easy and quick to access the information

in the dictionary. This method is therefore in line with common usage of the noun in Ndebele, users' expectations and preferences.

Use of the initial vowel of the prefix in lemmatising noun headwords represents the noun in its original form in the language and in the way in which users utter and write the Ndebele noun. It leads to an unambiguous retrieval of information presented in both the macro- and microstructure of the dictionary. The approach is in line with the morphological structure of the Ndebele noun. Above all, it reflects the semantic value of the initial vowel, indicating that it is not merely a syntactic feature.

Maphosa (1997: 35) notes that the demerit of lemmatising noun headwords using the initial vowel of the prefix is that it results in an uneven distribution of headwords across the letters of the alphabet. This is not much of a demerit as Maphosa also acknowledges, it is just a reflection of the nature of a specific language.

The argument that the lemmatisation of headwords using the initial vowel of the prefix takes up much space can be addressed by entering only the singular and neutral forms as headwords. Plural forms with existing singular forms would then be avoided, since the noun class system will be clearly and adequately explained, allowing users to deduce the plural forms by themselves, guided by the table of noun classes in the front matter of the dictionary. Inclusion of such information will complement efforts of making an exposition of Ndebele grammar available.

Because space is critical in a dictionary, no dictionary is spared the necessity of saving space (Landau 2001: 375). Entering nouns in their singular form is one method ISN adopts as a space-saving device. Hadebe (2002: ixxxiv) notes:

Isichazamazwi lesi siqukethe amabizo aphawula ubunye. Lokhu kwenzelwe ukulondoloza indawo egwalweni kanti futhi uma igama selivele lachazwa akudingeki ukubona lichazwe njalo selikhomba ubunengi ngoba ingcazelo iyabe ilokhu ifana.

(This dictionary contains nouns that are in their singular form. This has been done to save space in the dictionary and when a noun has been defined there is no need to see it defined again in its plural form since the meaning will still be the same.)

However, in ISM some nouns are entered in both their plural and singular forms, for example *izihlangu* (shields) and *amabhetshu* (loin skins). It should not be denied that for some, common usage reflects their use in the plural or collective form, hence requiring them to be entered as plurals, for example *amahlwayi* (leg rattles), *izikeyi* (wooden clappers) and *amangqongqo* (beads). A plural such as *amanothi* (notes) is for example defined to capture a slightly different sense from the singular *inothi* (note). It is, however, necessary to avoid entering both the plural and singular forms where a singular form can be used.

For the sake of the user-friendliness of the dictionary and the accessibility of its contents, it is worthwhile to lemmatise using the initial vowel of the prefix. This approach preserves and reflects the language in its natural state, representing the noun headword in a manner that is familiar to the user in speech

and writing. It does not require prior linguistic knowledge and training from the user to accurately identify the stem of the noun and locate it in the dictionary. In other words, as Hadebe (2002: xxii) notes, the method of lemmatising the stem does not reflect common usage of the noun in a way familiar to users:

Lokhu kwethula amabizo kutsho ukuthi nxa obalayo engenelisi ukwahlukanisa igama ngezakhi zalo angeke alithola lula ibala alidingayo.

(This method of lemmatising nouns means that if the user is not able to segment the noun by its morphemes he/she cannot find the word he/she is searching for.)

The controversies on where the stem begins complicate decisions for the compilers, leading to inconsistencies within a dictionary and among dictionaries for similar noun headwords and those sharing the same stem. From the perspective of the users, the consultation process is even more complicated because they would not know for certain under which letter to find the entry.

Furthermore, the centrality of the noun prefix in the derivation of concordial agreements and other dependent substantival prefixes makes it necessary to lemmatise using the initial vowel of the prefix, which, with the prefix proper, is the base prefix. In addition to this, it is the noun class prefix in Ndebele, which denotes number in terms of singular, plural or neuter. In the Nguni languages, the noun prefix is therefore more than a syntactic feature. Except for natural gender, it expresses the categories of number and person, making it a semantic feature than cannot be ignored in a field where meaning is the focal point. The fact that the noun class prefix and the noun stem are bound morphemes means they cannot stand separately as entries because in this form they do not constitute words. It is therefore linguistically proper to lemmatise using the initial vowel of the prefix for the entries to be called words and more precisely Ndebele nouns.

Conclusion

It has been argued that the treatment and lemmatisation of vowel-commencing borrowed nouns should occur in a way that satisfies the decoding needs of the target users. It must at the same time help the encoding needs of the users in providing sufficient help for them to produce correct vowel-commencing loan nouns in speech and writing. It has also been pointed out that lemmatising using the initial vowel of the prefix enhances information retrieval strategies, being in line with user-perspective and -preference because it presents the noun in a way familiar and common to the user. This way of lemmatising has the advantage of easy reference, rendering the dictionary more user-friendly and accessible. In addition to this, it is in line with the principles of term creation discussed in the article. On the other hand lemmatising using the initial letter of the stem is subject to many limitations and demerits which render the dictionary less user-friendly and accessible. With regard to the adoption and

adaptation of vowel-commencing borrowed nouns, it has been noted that the use of semi-vowels to break vowel sequencing is not ideal in all instances. Although there are cases where it works, by and large this does not seem to be the best method for breaking vowel sequencing. Except that this method violates the principles of term creation, it compromises user-preference and –perspective. With regard to the lemmatisation and treatment of vowel-commencing loan nouns in ISN and ISM the compilers' use of semi-vowels to break vowel sequencing overlooks the linguistic and communicative needs of the target users in both the selection and treatment of these nouns at the micro- and macrostructural levels. Finally, with regard to the allocation of borrowed nouns to noun class prefixes, Canonici's set of criteria ought to be used in determining the class affiliation of each borrowed noun and an emerging class 9 allomorph or class 9a which is morphologically realised as /i-/ has been recommended for the majority of the borrowed nouns allocated to class 5 in ISN.

References

Primary Sources

- Dent, G.R. and C.L.S. Nyembezi. 1969. *Scholar's Zulu Dictionary: English–Zulu, Zulu–English*. Pietermaritzburg: Shuter and Shooter.
- Doke, C.M. and B.W. Vilakazi. 1948. *Zulu–English Dictionary*. Johannesburg: Witwatersrand University Press.
- Doke, C.M. et al. 1958. *English–Zulu Dictionary*. Johannesburg: Witwatersrand University Press.
- Hadebe, S. et al. (Eds.). 2002. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Kropf, A. 1899. *A Kaffir–English Dictionary*. King William's Town: Lovedale Mission Press.
- Nkabinde, A.C. 1982. *Isichazamazwi 1*. Pietermaritzburg: Shuter and Shooter.
- Nkabinde, A.C. 1985. *Isichazamazwi 2*. Cape Town: Oxford University Press.
- Nkomo, D. and N. Moyo (Eds.). 2006. *Isichazamazwi SezoMculo*. Gweru: Mambo Press.
- Nyembezi, C.L.S. 1992. *AZ: Isichazamazwi sanamuhla nangumuso*. Pietermaritzburg: Reach Out Publishers.
- Pelling, J.N. 1966. *A Practical Ndebele Dictionary*. Harare: Longman Publishers.

Secondary Sources

- Béjoint, H. 2000. *Modern Lexicography: An Introduction*. Oxford: Oxford University Press.
- Canonici, N.N. 1995. *Elements of Zulu Morph-Syntax*. *Zulu Language and Literature*. Durban: University of Natal.
- Doke, C.M. 1993. *Trekking in South Central Africa, 1913–1919*. Edited by R.K. Herbert. Johannesburg: Witwatersrand University Press.
- Eastman, C.M. 1991. Loanwords and Swahili Nominal Inflection. Blommaert, J. (Ed.). 1991. *Swahili Studies: Essays in Honour of Marcel van Spaandonck*: 57-77. Ghent: Academia Press.

- Gilreath, C.T.** 1993. *Onometrics: The Formal Evaluation of Terms*. Strechlow, R.A. and S.E. Wright (Eds.). *Standardising Terminology for Better Communication*. Oxford: Basil Blackwell.
- Hadebe, S.** 2006. *The Standardisation of the Ndebele Language through Dictionary-Making*. Unpublished D.Phil. Dissertation. Harare/Oslo: University of Zimbabwe/University of Oslo.
- Khumalo, L.** 2003. *A General Introduction to Ndebele Grammar*. Harare: University of Zimbabwe Publications.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Maphosa, M.** 1997. *The Morphological Structure of the Ndebele Noun and Implications on the Ordering of Entries in Ndebele Dictionaries*. Unpublished B.A. Hons. Thesis. Harare: University of Zimbabwe.
- Prinsloo, D.J.** 2001. The Compilation of Electronic Dictionaries for the African Languages. *Lexikos* 11: 139-159.
- Von Staden, P.M.S.** 1973. The Initial Vowel of the Noun in Zulu. *African Studies* 32 (3): 163-181.
- Zawawi, S.M.** 1979. *Loanwords and their Effect on the Classification of Swahili Nominals*. Leyden: E.J. Brill.

Looking beyond Meaning in the *Advanced Ndebele Dictionary**

Langa Khumalo, *Department of Linguistics, University of KwaZulu-Natal,
Durban, Republic of South Africa (la_nga@yahoo.co.uk)*

Abstract: It is an established view in lexicography that the most important function of early dictionaries was to provide information on the meaning of words of a particular language. Over the years, tendencies have emerged with modern dictionaries providing detailed linguistic information resulting in more informative dictionaries. This article discusses the presentation of grammatical information, pronunciation, tone marking and usage labels and the structure and content of the back matter in the prospective *Advanced Ndebele Dictionary* (henceforth AND), which will be a successor to *Isichazamazwi SesiNdebele* (2001), the first-ever monolingual dictionary in Ndebele.¹ It is therefore the inclusion of this additional information that is examined in this article. The AND is still restricted to the planning stages. The work that has been done on the dictionary has been confined to academic articles about the dictionary's structure and content. The current article is a third instalment on the AND following on Khumalo 2003 and 2007.

Keywords: ADVANCED NDEBELE DICTIONARY, GRAMMATICAL INFORMATION, MEANING, TONE MARKING, PRONUNCIATION, USAGE LABELS, BACK MATTER

Opsomming: Om verder te kyk as betekenis in die *Gevorderde Ndebele-woordeboek*. Dit is 'n gevestigde siening in die leksikografie dat die belangrikste funksie van vroeëre woordeboeke was om inligting te verskaf oor die betekenis van die woorde van 'n bepaalde taal. Oor die jare het neigings na vore gekom met moderne woordeboeke wat uitvoerige taalkundige inligting verskaf wat gelei het tot meer informatiewe woordeboeke. Hierdie artikel bespreek die aanbieding van grammatiese inligting, uitspraak, toonaanduiding en gebruiksetikette en die struktuur en inhoud van die agterwerk in die *Gevorderde Ndebelewoordeboek* (voortaan GNW) wat 'n opvolger sal wees van *Isichazamazwi SesiNdebele* (2001), die heel eerste eentalige woordeboek in Ndebele. Dit is dus die insluiting van hierdie bykomende inligting wat in hierdie artikel ondersoek word. Die GNW is nog tot die beplanningstadiums beperk. Die werk wat aan die woordeboek gedoen is, het nog net gevorder tot akademiese artikels oor die woordeboek se struktuur en inhoud. Die huidige artikel is 'n derde aflewering oor die GNW in opvolging van Khumalo 2003 en 2007.

Sleutelwoorde: GEVORDERDE NDEBELEWOORDEBOEK, GRAMMATIESE INLIGTING, BETEKENIS, TOONAANDUIDING, UITSPRAAK, GEBRUIKSETIKETTE, AGTERWERK

* This article was presented as a paper at the Fourteenth International Conference of the African Association for Lexicography, organised by the Xhosa Departement, University of the Western Cape, Bellville, Republic of South Africa, 6–8 July 2009.

Introduction

Monolingual lexicography in Zimbabwe was initiated by the African Languages Lexical Project, known by its acronym as the ALLEX Project. The main aim of the ALLEX Project was the compilation of a variety of general and specialised monolingual dictionaries in Zimbabwe's two national languages, Shona and Ndebele. Work in the ALLEX Project, established in 1992 and institutionalised as the African Languages Research Institute in 2000, has culminated in the publication of two monolingual Shona dictionaries, *Duramazwi reChiShona* (henceforth DRC) (Chimhundu 1996) and *Duramazwi Guru reChiShona* (henceforth DGC) (Chimhundu 2001) and a general medium-sized Ndebele monolingual dictionary, *Isichazamazwi SesiNdebele* (henceforth ISN) (Hadebe 2001) (see also Chabata 2007: 281). Evidently, these dictionaries were completed in phases, with phase one being the pioneer DRC followed by phase two, which saw the production of both the DGC and the ISN. Since the ISN was preceded by the DRC, it was modelled on the latter, however, with some substantial improvements on its predecessor. Likewise the compilation of the *Advanced Ndebele Dictionary* (henceforth AND) will follow the model of its predecessor, the DGC. It is the AND, a successor to the ISN, which this article seeks to discuss.

The main objective of the ALLEX Project, which was expanded and transformed into the African Languages Research Institute (ALRI) is, according to its institutional mandate, to research, record, reanimate and develop the African languages of the country, encouraging their ever wider use in all the multifarious affairs of life by the general public of today and, increasingly, by generations to come (Fortune 2005: 1). The compilation of the AND is one of the undertakings being planned by ALRI in fulfilling this institutional mandate. As explained by Khumalo (2007: 1):

The AND will not just be larger than the ISN, but will be more advanced with regard to the depth and scope of its lexical items and definitions. ... Unlike its forerunner, the AND will provide grammatical information for each lexical item, including phonetic transcription, tone marking and etymology.

It is the inclusion of detailed grammatical information, pronunciation, tone marking and usage labels and the structure and content of the back matter in the AND that is critically examined in this article.

The Target User

It is befitting to begin with Gouws's (2006: 85) observation that at the core of all the decisions pertaining to the compilation of any dictionary stands the person who will open the dictionary and use it, the target user. Gouws further observes that modern-day lexicography demands a clear projection of the genuine purpose of each dictionary. The fulfilment of this genuine purpose is meas-

ured by the ease with which the target user can successfully retrieve the information provided in the dictionary. It must, therefore, be clear to the lexicographer who the target user is and how best to present information to allow maximum access to knowledge and facilitate easy use by the target user.

The target groups for the ISN and AND are different. The ISN was mainly designed to fulfil the needs of lower secondary school pupils. The primary needs of the users at this level are identified as reading and comprehension, composition, basic structure of the language and general language use (Chabata and Mavhu 2005). Accordingly the ISN was crafted to contain a basic Ndebele vocabulary and a basic grammatical description with a mini-grammar in the front matter. Furthermore, it was designed to be inexpensive and easy to handle so as to be accessible to secondary school pupils. The AND on the other hand is intended to be a comprehensive reference work, which will serve as a resource for more advanced users. Its target users are those at higher levels at schools and university undergraduate level, i.e. the higher secondary and tertiary education levels. The AND is designed to be a general-purpose, medium-sized, synchronic, monolingual dictionary, meant to be a comprehensive description of the Ndebele language. Hence it is conceived to be as inclusive as possible of the vocabulary and usage of the language as gleaned from the ALLEX Ndebele corpus.² While traditionally the most central information in a dictionary is the description of words and their meaning, and that meaning is even more central when it comes to monolingual dictionaries (Landau 2001: 8), it has also become the norm in modern lexicography to include information on aspects such as the spelling, pronunciation and etymology of words. By highlighting the inclusion of this important information, this article therefore aims at looking beyond meaning in the AND.

Pronunciation

A monolingual dictionary in a native language assumes that the target user's knowledge of the language is to some extent limited and/or imperfect (Svensén 1993: 20). However, it has been argued by Jackson (2002: 102) that pronunciation is not information that native speakers regularly search in a dictionary. If they do, Jackson (2002) contends, it is likely to check the pronunciation of a word that they have encountered only in writing. How a word is pronounced is very important as this is one of the idiosyncratic facts of a language.

It has been the practice of some lexicographers to proscribe the transcription of the pronunciation of what they refer to as "ordinary, everyday words". This raises the question of determining what constitutes an "ordinary, everyday word". It results in an act of pronouncing judgement on a particular aspect of the language under description. Hence the counter-argument to this has been that the exclusion of information on pronunciation represents a move away from a subservience to the recording function of general-purpose dictionaries towards considerations of what the user might or might not need (Jackson

2002: 102). Chimhundu (1997: 154) observes that "the dictionary is [...] the chief instrument in interpreting rules about spelling and pronunciation". Therefore the inclusion of pronunciation is important to a fuller description of the language.

Following Landau's (2001: 154) observation that "every lexicographer, like any good author, has his readers very much in mind", the AND will include pronunciation because the main target users would find it very helpful. The target users are tertiary and university level students who are introduced to language courses, which among others cover aspects of phonetics and phonology. Further, the dictionary may be used by second language speakers who may find the transcriptions useful to acquire the correct pronunciation of words. The inclusion of pronunciation is thus well motivated.

The transcription system to be used for representing pronunciation is the International Phonetic Alphabet (IPA). This roman-based system can be used for transcribing the speech of any language. The example below shows a typical entry envisaged for the AND.

(1) **itshabi** /itʃʰaβi/ KKP bz 5. THUKA

The headword in example (1) is in bold lower-case roman letters which is then followed by the IPA transcription between slashes.

Grammatical Information

It is one of the conventions of lexicography to identify the word class(es) or part(s) of speech to which each lexeme in a dictionary belongs. The traditional terms, usually abbreviated, are *ibizo* (bz) (noun), *isenzo* (sz) (verb), *isandiso* (sd) (adverb), and so on. Verbs are usually marked for their different senses such as transitive (*sz mwa*) and intransitive (*sz gmwa*). The ISN gives all this information, including number in the case of nouns. However, the AND will present more grammatical information such as tone marking for each headword, which was not provided in the ISN. It was considered that since the ISN is a monolingual dictionary, a mother-tongue speaker would hardly require the tone marking of two or more homonyms because context can easily disambiguate such words. Further, it was considered that the main target users of the ISN, i.e. secondary school pupils, are not at a level where such grammatical detail would have been introduced to them, hence complicating the accessibility of the information.

However, whereas Ndebele like most Bantu languages is tonal, the exclusion of tone marking therefore denies the complete grammatical representation of the language. It was decided therefore to include tone marking, together with the phonetic transcription of each headword in the AND, as shown in example (1). Tone, like pronunciation, is an important distinctive detail of each word in Ndebele. Example (1), here given as (2), can now be further examined.

(2) **itshabi** /itʃʰaβi/ KKP bz 5. *THUKA*

The headword and its pronunciation is followed by tone marking with K, *khweza/khuphula*, for high tone, and P, *phansi* for low tone. However, as a space-saving device in the dictionary, high and low tone through-out will be marked with K followed by a hyphen (K-) for high tone and P followed by a hyphen (P-) for low tone as shown in examples (3)(a) and (b) respectively.

(3)(a) **ubaba** /uβaβa/ K- bz 1a.

(b) **mathunzi** /matʰunzi/ P- isich.

Tone marking will be followed by more grammatical information, given by an abbreviated label in standard lower-case roman letters; a number, in the case of a noun, indicating the class to which it belongs as is shown in examples (2) and (3)(a). The aim of providing grammatical information is to explain the structure of the language itself, not just the meanings of the individual words.

Usage Labels

All dictionaries have a set of labels or style markers to indicate words or senses of words that are restricted in some way with respect to the context in which they may occur. The contextual restrictions may be geographical, i.e. depicting regional usage, as well as with regard to time (e.g. archaic), style (e.g. informal), and topic (e.g. medicine). In this section, the types of usage labels proposed for the AND are reviewed.

Dialect labels refer to geographical restrictions. The word *inopi* was particularly resisted as a legitimate entry in the ISN. It was argued that *inopi* was restricted to the Plumtree region in Zimbabwe where it was borrowed from the Kalanga language predominately spoken there. Unfortunately, because the ISN does not have usage labels, the region could not be given to indicate the restricted use of the word, although it appears in the Ndebele corpus as a loan word. It will be entered as follows in the AND.

(4) **inopi** /inʰopi/ KPP bz 9. *PLUMTREE*

Plumtree is the region where this word is believed to be predominantly used. The use of the regional or dialect label would have been sufficient to assuage criticism. Related to this entry is also the inclusion in the ISN of words which were argued to be informal and should hence not be part of the dictionary. Again, the headwords in question had sufficient evidence of usage as gleaned from the corpus. But schoolteachers argued that, because the words are used only informally, they would not allow their pupils to use them in class or in their essay or composition writing. One such word in the ISN is *ishamari* which will be entered as follows in AND.

- (5) **ishamari** /iʃamari/ [itʰamali] KKPP bz 5. *KILIZA*

The label *kiliza* (informal) will indicate that the word is characteristic of conversational and casual rather than official or formal speech or writing.

The presentation of offensive words is very difficult to deal with in most cultures. A large body of offensive words in Ndebele consists of those that may be called obscene. Read (1934), quoted by Landau (1991: 229), describes obscenity as follows: "Any reference to the bodily functions that gives to anyone a certain emotional reaction, that of 'fearful thrill' in seeing, doing or speaking the forbidden. Thus, it is the existence of the ban or taboo that creates the obscenity." Offensive or taboo words refer to a variety of words that may be considered obscene, vulgar, impolite, insulting and derogatory, those referring to physical deformity or mental impairment and all those usually avoided or normally found unacceptable in the Ndebele culture. Their use goes against the norms of good behaviour because they are either injurious to one's dignity or hurt the feelings of individuals or social groups (Chabata and Mavhu 2007: 257).

Obscene words usually refer to private body parts connected with sexual activities and excretion processes and whose exposure in public is prohibited or censured. Such words are socially marked. They may not be spoken or used in public without causing embarrassment either to the speaker or the listener. Examples of offensive Ndebele words to be included in the AND are listed under (6).

- (6) *ubolo* (penis), *igolo* (vagina), *umsunu* (clitoris), *isende* (testicle), *umdididi* (anus), *-bhotsha* (defecate), *-chema* (urinate/ejaculate).

Obscene words also include terms generally considered insulting. Examples of such words are derogatory terms referring to certain ethnic groups, e.g. the Ndebele terms: *itshabi* (a foreigner or a person not of Ndebele origin), *iswina* (a Shona person) and *ihole* (a person not of Nguni origin). In Ndebele the use of offensive words in ordinary speech is prohibited and viewed as socially deviant.

Although Jackson (2002: 112) contends that there is little left in our society that is forbidden, and that some modern dictionaries no longer include usage labels, the AND will adopt the following usage labels for obscene words.

- (7)(a) **ubolo** /uβolo/ KKK bz 11. *HLONIPHISA*
 (b) **igolo** /igolo/ KKK bz 5. *HLONIPHISA*
 (c) **ihole** /irole/ KKP bz 5. *THUKA*
 (d) **itshabi** /itʃaβi/ KKK bz 5. *THUKA*

Commenting on the use of such words in the first monolingual Shona dictionary, Kahari (1996–1997: 38) states that "the degree of specificity of most of the

items goes beyond what the reader who is a native speaker wants to know". The inclusion of obscene words in the dictionary, he argues, lacks "traditionally accepted taste and decorum". The AND, following the DGC, will adopt the usage labels or stylistic markers shown in example (7) (although these slightly deviate from those used in the DGC). As in previous examples, (7) presents typical entries planned for the AND.

With reference to words referring to private body parts, the AND will use the label *hloniphisa*, meaning 'embarrassing', as can be seen in examples (7)(a) and (b). Words that carry a derogatory connotation will be marked *thuka* meaning 'insulting' or 'offensive' as shown in examples (7)(c) and (d). The labels or markers will be a warning to the user that the terms are of restricted or context sensitive use. Giving such a warning is not only informative to the users but also courteous to those who might take offence.

The components of a typical major entry in the dictionary will therefore be the headword, in bold lower-case roman letters. This will be followed by the IPA transcription between slashes. Subsequently tone marking is noted, which in turn precedes further grammatical information. In case of a noun, this is rendered by an abbreviated label in standard lower-case roman letters and a number, indicating the respective noun class. The last part of the lexical entry consists of the usage label, indicating contextual restrictions such as geographical (regional), historical, and stylistic (offensive) registers. Hadebe (2007: 295) correctly observes that dictionary users usually assume and expect the dictionary to give them the 'correct' spelling, pronunciation and usage. The AND will, to a certain extent, meet some of the expectations of the users in this regard.

Modern dictionaries also often add labels indicating technical vocabulary. Svensén (1993: 49) notes that technical language arises as a consequence of constant development and specialisation in the fields of science, technology and sociology. New concepts are constantly being defined, and in order to exchange information about them, new linguistic expressions have to be coined. There are many specialised terms used in technical subjects in education, economics, sport, law, medicine and others. However, only terms that are commonly used in the Ndebele-speaking community will be candidates for selection. The prime targets are words derived from specialised fields, referred to as "international". Chitauro-Mawema (2000: 212) defines international words as those "technical words which carry specific unchanging and unambiguous senses in the contexts in which they occur and are used internationally". These are terms for respective concepts usually acquired by borrowing from other languages in which they occur. Dictionaries usually add an appropriate label to mark that a word is used in a particular field or belongs to a specific topic. The examples under (8) are typical labels that would be used in the AND.

(8)(a) **ibhanditshi** /ibæn.ditʰi/ KPPP bz 1a *KWEZEMPILAKAHLE* (bandage – in medicine)

(b) **ibhokiseni** /ibokiseni/ K- bz 5. *KWEZEMIDLALO* (boxing – in sport)

- (c) **ithena** /ithɛna/ KKP bz 5. *KWEZOMCULO* (tenor – in music)
- (d) **ijesi** /idʒɛsi/ KPP bz 5. *KUZEMBATHO* (jersey – clothing)
- (e) **iyabhakhasi** /ijæb^hkasi/ KKPPP bz 5. *KUSAMU* (abacus – in mathematics)

By encouraging greater use of the language in various specialised fields, the dictionary is also contributing to the conducting of all kinds of education in the mother tongue. The mother tongue is the medium through which education is most effective and personalised, and through which it can best pervade society, raising levels of skill and ability with a resultant higher standard of life (Fortune 2005: 5). The AND will provide a basis on which to build specialised terminology and, while building and enriching it, be put to use in all the affairs of life from the first steps in education up to the realms of law and government, in the economy, in sport, and in the media. As Fortune (2005: 6) puts it:

The aim [...] is to see the Zimbabwean languages reach that level of development, earning and enjoying the status of a country with its own character and idiom. To do so will need a concerted commitment and effort on the part of all its speakers. But, given that, there is no reason why the African languages should not serve and deliver to their societies all that human language has done and does elsewhere.

The Back Matter

According to Gouws (2006: 87), modern-day lexicography emphasises the need for a frame structure with front and back matter texts complementing the central list and adding venues that can accommodate lexicographic data. Users must be made aware that the dictionary as a carrier of text types presents much more than the data included in the central list. While the front matter normally carries instructions of use, describing how the dictionary is organised, how it was compiled and most importantly, how it can effectively be used, the back matter can contain practical information about language and culture that can be useful to the user's daily life.

Since the AND is a more comprehensive dictionary, its back matter will expand on the information provided in that of the ISN. For instance, the list of common first names will include information on the derivation of these names, e.g. *Musawenkosi* (God's grace) is clearly motivated by a Christian background. Further, the list of the months will also be complemented with etymological information to reveal the derivation of these names, e.g. *uNhlolanja* (literally, "dog's taboos", February) is the time for the mating of dogs, or *uNtulikazi* (literally, "a lot of dust", July) is the time characterised by much wind and dust.

Moreover, the back matter will also contain a comprehensive list of Ndebele vowels and consonants and orthographical rules. Additionally it is envisaged that it will comprise common mathematical and scientific symbols gener-

ally used in higher and tertiary institutions to introduce the language to areas where it has not been previously used. The final appendix will therefore consist of the most useful and comprehensive list applicable to the areas of linguistic analysis and literary appreciation and criticism.

Conclusion

The article has highlighted some of the key considerations of each lexical entry in the AND, all of which go beyond just giving the meaning. The style manual of the AND states that one of the crucial information types the dictionary will provide is grammatical information, pronunciation, tone marking, usage labels and a comprehensive back matter to enhance the coverage and usefulness of the dictionary. The aim of the article has been to justify the inclusion of this information, presenting some proposed trial entries for the AND. The dictionary will hopefully encourage a wider use of the Ndebele language and by so doing also provide the basis for ultimately conducting education of all kinds in the mother tongue.

Endnotes

1. The Ndebele language discussed here is that spoken in Zimbabwe. It is very closely related to the Zulu language. There is also the Ndebele language spoken in South Africa. There are no known studies comparing Zimbabwean and South African Ndebele.
2. The ALLEX Ndebele corpus is a systematic, well-designed and selective collection of written material and transcribed speech (Hadebe 2002: 47). The size of the corpus is currently over two million tokens (Khumalo 2007: 52).

References

- Chabata, E.** 2007. The African Languages Research Institute: A Milestone in the Development of Zimbabwean Languages. *Lexikos* 17: 278-291.
- Chabata, E. and W. Mavhu.** 2005. To Call or Not to Call a Spade a Spade: The Dilemma of Treating 'Offensive' Terms in *Duramazwi Guru reChiShona*. *Lexikos* 15: 253-264.
- Chimhundu, H.** 1997. Language Standardisation without Policy or Planning: Zimbabwe as a Case Study. Røyneland, U. (Ed.). 1977. *Language Contact and Language Conflict*: 129-150. Volda: Volda College.
- Chimhundu, H. (Ed.).** 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. (Ed.).** 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Chitauro-Mawema, M.** 2000. Grammatical Terminology for the Teaching of African Languages at Tertiary Level. *Lexikos* 10: 209-224.
- Fortune, G.** 2005. Book Review, *Duramazwi Guru reChiShona*. Unpublished.
- Gouws, R.H.** 2006. Discipline, Disciplines, Dilemmas and Data Distribution in the Planning and Compilation of Monolingual Dictionaries. *Lexikos* 16: 84-94.

- Hadebe, S.** 2002. *The Standardisation of the Ndebele Language through Dictionary Making*. Ph.D. Dissertation. Oslo: University of Oslo/Harare: University of Zimbabwe.
- Hadebe, S.** 2007. Language Planning and Monolingual Dictionaries: With Special Reference to Ndebele. *Lexikos* 17: 292-306.
- Hadebe, S. et al. (Eds.)**. 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Jackson, H.** 2002. *Lexicography: An Introduction*. London/New York: Routledge.
- Kahari, G.P.** 1996–1997. Review of *Duramazwi reChiShona*. *Southern African Political and Economic Monthly* 10(3–4): 33-38.
- Khumalo, L.** 2003. *From a General to an Advanced Ndebele Dictionary: An Outline*. Unpublished Monograph. African Languages Research Institute.
- Khumalo, L.** 2007. *An Analysis of the Ndebele Passive Construction*. Unpublished Ph.D. Dissertation. Oslo: University of Oslo.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. New York: Cambridge University Press.
- Read, A.W.** 1934. An Obscenity Symbol. *American Speech* 9(4): 264-278.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford/New York: Oxford University Press.

The Compilation of a Shona Children's Dictionary: Challenges and Solutions

Peniah Mabaso, *African Languages Research Institute (ALRI), University of Zimbabwe, Harare, Zimbabwe (penmabaso@arts.uz.ac.zw)*

Abstract: This article outlines the challenges encountered by the African Languages Research Institute (ALRI) team members in the compilation of the monolingual Shona Children's Dictionary. The focus is mainly on the problems met in headword selection. Solutions by the team members when dealing with these problems are also presented.

Keywords: SHONA CHILDREN'S DICTIONARY, LOANWORDS, TABOO, HEADWORD SELECTION, SLANG, TONE MARKING, ILLUSTRATIONS, SYNONYMS, VARIANTS

Opsomming: Die samestelling van 'n Sjonakinderwoordeboek: Uitdagings en oplossings. Hierdie artikel skets die uitdagings waarvoor die spanlede van die African Languages Research Institute (ALRI) te staan gekom het by die samestelling van die eentalige Sjonakinderwoordeboek. Die fokus is hoofsaaklik op die probleme ondervind by die keuse van trefwoorde. Oplossings deur die spanlede by die benadering van hierdie probleme word ook aangebied.

Slutelwoorde: SJONAKINDERWOORDEBOEK, LEENWOORDE, TABOE, TREFWOORD-KEUSE, SLENG, TOONAANDUIDING, ILLUSTRASIES, SINONIEME, VARIANTE

Introduction

This article outlines the challenges encountered by ALRI researchers in the compilation of the monolingual Shona Children's Dictionary (henceforth SCD) which is the first of its kind in Shona. The compilers adopted some of the techniques used in compiling the advanced Shona dictionary *Duramazwi Guru reChiShona* (henceforth DGC). Two of the SCD compilers and editors were also part of the compilers and editors of DGC. However, the dictionary is different from the DGC in many respects such as the target users, the structure and the contents. The SCD is mainly meant for primary school children, the seven to thirteen years age group and their teachers. Student teachers at teacher training colleges and their lecturers are also some of the potential target users. On the contrary, the DGC targeted high school learners and students at tertiary institutions like colleges and universities.

The main challenge of this project was what exactly such a specified dictionary should contain. Another challenge the compilers faced was the issue of

language dynamism. The question was if slang, colloquial and loanwords should be included in the dictionary. Should taboo or obscene words be given as well? Is it important to mark tone? It had to be decided which illustrations should be included in the dictionary, considering the different urban and rural backgrounds of the main target users. The compilers agreed on possible solutions to some of the challenges for the dictionary to be user-friendly, satisfying its users after publication.

Challenges encountered in headword selection

Headword selection is an important stage in the compilation of any dictionary as it determines the lexical items to include in and exclude from a dictionary. Svensén (1993: 40) argues that the selection of headwords must be preceded by a number of decisions regarding the size and aim of the dictionary. For the SCD compilers, headword selection was difficult mainly because of the specified nature of the dictionary, which is meant for school children. As a result not all Shona words could find their way into the dictionary. Only words which the compilers and reference groups agreed were suitable for the targeted age group were entered as headwords.

In dictionary compilation, headwords have to be carefully selected and defined in a special way to satisfy the needs of the target users. This can only be achieved when the users can access lexical items which they anticipate and which are helpful to them. The selection process for the SCD headwords was designed according to well-informed principles. The compilers selected headwords specifically suited for children. In this case, the problem was how to determine whether a headword was commonly used by children.

The problem was partially solved when the compilers agreed that Shona primary school textbooks should be used as sources for headwords. However, from all the Shona textbooks, the compilers could not collect enough headwords for the dictionary. As a result, they decided that words from other primary school textbooks for subjects such as mathematics and general publications encompassing subjects such as science, biology, geography, religious studies and agriculture should also be used as headword sources. The headwords were then adopted in and adapted to Shona. This also implies that the dictionary incorporated loanwords to make provision for new concepts which previously did not exist in Shona.

Loanwords in Shona: Why are there so many?

The use of loanwords is not unique to Shona speakers. Shona has not been immune to language contact and its resultant effect of cultural and language borrowing. In Zimbabwe, English has always enjoyed a prestigious position in both urban and rural areas, mainly because Zimbabwe is a former British col-

ony. Despite the fact that Zimbabwe has enjoyed political independence for 28 years, English has remained the official language. It is still obligatory for social promotion, being associated with prestige. Many children from well-to-do families are sent to Group A schools or former white-only schools where English is used as medium of instruction. In Zimbabwe where many Shona and Ndebele speakers find themselves in a diglossic situation with English in the most prominent position, borrowing therefore mainly occurs from English. Chitiga (1995: 6) who discusses the politics of language in Zimbabwe, asserts that English has assumed the most important position in the life of Zimbabweans in all formal sectors like education, commerce, media and technology. Owing to the prevailing socio-linguistic situation in Zimbabwe, the Shona lexicon largely contains loanwords from English.

Basing themselves on the above historical, social, technological and economic factors, the compilers included loanwords as headwords in the SCD. To exclude loanwords from the SCD was to ignore the actual language situation, as some loanwords have in fact become part of the Shona lexicon. All languages develop through borrowing, as Ndlovu (1998: 33) argues:

Language change is a universally accepted and attested principle that has long been established by historical linguistics. No language is fixed; all languages undergo constant change over time. New words are constantly coming into use and at the same time old words are gradually dropping out of use.

This assertion accentuates that the lexicons of all languages are continuously changing. These language changes are due to a number of factors, among others, multilingual interactions as well as the continued spreading of English and other languages inherited from former colonial powers. In Zimbabwe, literate Shona speakers use both English and Shona in general conversation causing English sounds, grammar and vocabulary to exercise a strong influence on Shona.

Criteria used for the selection of loanwords

Kipfer (1984: 35) argues that lexicographers have to keep up with new words and meanings as well as the development of language in general. The SCD compilers saw it unwise to exclude loanwords since they now constitute a significant percentage of the Shona lexicon. They considered the historical, social, technological and economic factors to justify their selection of loanwords in the SCD. From the July 2005 outreach programme held by the SCD compilers, different views were received from target users on the issue of loanwords. There were some Shona language conservatives who were against the inclusion of loanwords, while 'modernists' argued for their inclusion. Svensén (1993: 47) asserts that one of the principal recommendations of publishers is that their new dictionaries contain the latest material. Taking Svensén's point of view

into consideration for the SCD to appeal to its target users, the latest material in the form of loanwords relating to science education and technology were included as headwords. Where there is an original Shona term for some words, that Shona term was used. However, there were some cases where it was found that loanwords were more popular, even though there were indigenous Shona terms. For example, it is now rare to hear speakers refer to mathematics/maths as *masvomhu* rather than *metsi*, bicycle as *bhizautare* rather than *bhasikoro*, and sugar as *tsvigiri* rather than *shuga*.

Faced with such a problem, the SCD compilers (and other ALRI team members) agreed that not every loanword should find its way into the dictionary. Systematic and well-defined ways were followed in selecting loanwords. The SCD compilers chose loanwords from the viewpoint of their stability and acceptance within the language. It was agreed that loanwords incorporated into the language and used in formal speech, showing that they have stood the test of time, should be included as headwords.

Another problem for the SCD compilers was their dealing with words containing letters or digraphs not recognised in Shona orthography. Some letters and digraphs like *q*, *x* and *hl* are prevalent in Shona speech, especially in the Karanga and Ndaou dialects. Although they are found in the major Shona dialects, they are not part of the standard Shona alphabet. Magwa (2002) notes:

From 1967 onwards, speakers of different dialects were experiencing certain difficulties arising from the defective alphabet and the spelling and word division system. The current orthography is linguistically constricting in a number of ways. For example, the standard alphabet does not have symbols representing the sounds [l] and [x], which are found in ChiKaranga, ChiNdaou and ChiKorekore dialects.

The compilers agreed that English loanwords should be adapted to conform to the Shona CV syllable structure. Most English loanwords are written as they are pronounced in Shona. Many technical terms such as *fax*, *diskette* and *computer* were entered respectively as **fakisi**, **dhisiketi** and **kombiyuta** in the SCD. Most Shona writers make sure that loanwords from English adhere to the Shona phonetic system. As has been mentioned earlier, in the Shona phonological alphabet, there is no lateral /l/. The closest equivalent of /l/ in Shona is the trill /r/. Loanwords like *loan* and *leave* were respectively rendered as **roni** and **rivhi** in Shona. However, the use of /r/ in place of /l/ does not apply to all English loanwords. There are exceptions such as **kukala** (to colour) and **yelo** (yellow) (Chitauro-Mawema 2000: 211). When *r* replaces *l*, the meaning of words changes, for example **kukala** (to colour) will become **kukara** (to be greedy). This shows that there is a limit to replacing loanwords with the letter *l* with those with the letter *r*. It seems that nowadays new words with *l* are incorporated into Shona. At present it is rare to hear someone using the accepted trill *r* in these loanwords. Shona speakers no longer say **raki** (luck), **rika** (leak) or **roni** (loan). In the SCD, all loanwords were included and represented

as they are pronounced using the letters, digraphs and trigraphs not recognised in standard Shona orthography. The compilers hope that these considerations will compel those responsible for changes in the Shona orthography, like the Shona Language and Culture Association (SLCA) and the Ministry of Education, to allow the use of some of the letters which are not recognised in the current Shona orthography.

Taboo, obscene and offensive words

In Shona, it is a cultural taboo to use derogatory words, especially those describing excretory and sexual organs and acts publicly in spoken or written form. The compilers had problems in deciding whether to include these obscene and offensive words in the dictionary. A few examples of such words are **gongo** (clitoris), **jende** (testicle) and **beche** (vagina). Since it is a taboo in Shona culture to use these words in all public communication, the Shona people have developed euphemistic terms to refer to these organs or acts. The anus is referred to as **kushure** (the back) and the penis as **nhengo yomurume** (a man's private part). The use of such euphemistic terms and phrases when referring to human anatomy and sexual behaviour shows how sensitive the Shona people are to these obscene and offensive words. The biggest problem likely to be caused by the use of these euphemistic terms in the SCD is that the children would not understand them since they appear ambiguous. Including obscene and offensive words in a dictionary meant for such a young age group would be a matter of licensing their use in everyday speech. The dictionary would be the first-ever Shona publication meant for children to use obscene and offensive words. However, excluding them from the dictionary would have been to ignore the actual language situation. It would also have been to overlook the educational value that the correct reference to these human anatomical parts can have for teaching children about sexual diseases and abuse.

The SCD compilers also faced many problems in dealing with offensive words describing certain negative features of people, for instance cripples and albinos. Shona words like **zvirema** for cripples, **mukawu/musupe/murungudunhu** for albinos, **hure** for a prostitute and **ngomwa** for a barren woman are often avoided, especially in the presence of the affected people themselves, who might feel offended. The use of such terms is against the Shona culture that protects the socially disadvantaged. Nevertheless, some of these offensive words were entered into the dictionary.

Slang and colloquial words

Shona speakers use many slang and colloquial words. According to Tullock (1994: 1454), slang refers to words, phrases and language use regarded as informal and often restricted to special contexts or peculiar to specific classes of

users. The use of slang and colloquial words varies with age groups. Among Shona speakers, children and youths seem to be the main age groups that often use slang. Some slang and colloquial words are only used for a short period while others are well established in the language. In Shona, slang words like **bhoo** (okay) are now part of everyday speech, especially among children. Some slang words like **chibhanzi** (money) and **shuzura** (to go) are transitory. Chimhundu (2001: xxxv) likens such transitory slang words to fashion which comes and goes, and therefore they were not included in the SCD.

The SCD compilers agreed that slang words that have become permanent in the Shona vocabulary should be treated as headwords and given the style marker "manje" (an abbreviation of "chimanjemanje") to show that they are colloquial words. Transitory slang words should not be included in the dictionary. The challenge for the SCD compilers was to distinguish between slang words that are established in Shona and those that are transitory.

Tone marking

In the DGC, tone was marked. Tone marking in Shona is very important as it distinguishes between words spelt the same but pronounced differently. For example, if tone is not marked, there will be no difference between **nzara** *low tone* (hunger) and **nzara** *low, high tone* (finger nail). However, primary school teachers and lecturers consulted during outreach programmes argued that tone should not be marked in such a dictionary, as it will confuse children. The dictionary should be very simple and concentrate on the meaning of words.

Linguistic details

The compilers agreed upon the marking of all parts of speech in the dictionary to give linguistic detail which was thought necessary for such a dictionary. It is indicated whether a headword is a noun, pronoun, adverb, adjective, preposition and so on.

Synonyms and variants

Synonyms are lexical items which have the same meaning or are so close in meaning that they can be used interchangeably in certain contexts. In Shona the word *baboon* can be translated with **bveni**, **dede**, **diro**, **gudo** and **mutiro**, depending on one's dialect. According to Crystal (1997: 408), a variant can be defined as a set of alternatives in a given context. Shona has many variants, mainly because of its different dialects. For example, the Karanga people in some instances use *zh* and *gw* where the other dialects like Zezuru and Manyika use *nz* and *rw* respectively as can be seen in words like **zhara** and **nzara** (hunger), **-gwara** and **-rwara** (to be sick). The problem regarding synonyms and

variants is their treatment in the dictionary, whether to define them separately or not.

In this connection, the compilers followed the procedure of the DGC compilers. The synonym or variant that is commonly used, carries the definition and all the others that are less common are cross-referenced to it. For example, for the synonyms **bveni**, **dede**, **diro**, **gudo** and **mutiro**, **gudo** carries the definition to which the others are cross-referenced, as shown below:

bveni D- z 5. ONA gudo 5.

gudo D- z 5. Mhuka yemusango ine mahobi neuso hwakada kufanana nehwetsoko. FAN bveni 5, dede 5, diro 5, mutiro 3.

Variants are defined as follows:

zhara D- z 9. ONA nzara 9.

nzara [zhara] D- z 9. Kunzwa kuda kudya nokuti mudumbu munenge musisina chikafu.

The SCD like the DGC contains as much information on the grammar in the front matter as is necessary for the target users to understand the abbreviations used.

Idioms

Idioms are also important in the SCD since they reflect the culture of the Shona speech community, so just like the DGC, the SCD includes and explains them. However, not all Shona idioms are included and explained in the SCD but only those suitable for children. To this purpose, the compilers used Shona primary school textbooks as sources for the idioms.

Illustrations

Since the dictionary is meant for children in both rural and urban areas, it has to include illustrations to explain some unfamiliar headwords in the dictionary. The major problem was to determine which headwords are unfamiliar to the target users, especially considering their different backgrounds. Some things familiar to urban children might not be familiar to rural children. Trying to satisfy the expectations of both the rural and the urban target users might have caused the dictionary to become nothing more than a book of illustrations. The compilers therefore agreed to include group illustrations as part of the back matter of the dictionary.

Conclusion

The article discussed some of the major problems unique to the SCD, when compared to other general dictionaries ALRI had compiled and published. Headword selection has proved to be the greatest challenge for the SCD compilers. Different word categories found in Shona like loanwords, offensive words, taboo or obscene words and slang words made headword selection difficult. Shona culture and the inadequacy of the Shona orthography caused a dilemma for the SCD compilers. Adhering to the current Shona orthography would have meant excluding some important headwords common in speech. The compilers explicitly state the challenges in the front matter and the compromises and solutions decided and agreed on.

References

- ALRI. 2002. *Shona Children's Dictionary Style Manual*. Unpublished Document. Harare: University of Zimbabwe.
- Chabata, E. 1998. Using the Predictability Criteria for Selecting Extended Verbs for Shona Dictionaries. *Lexikos* 8: 140-153.
- Chimhundu, H. (Ed.). 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Chitauro-Mawema, M.B. 2002. Challenges Encountered in the Compilation of an Advanced Shona Dictionary. *Lexikos* 10: 209-224.
- Chitiga, M. 1995. *Codeswitching in the Classroom*. Unpublished M.A. Thesis. Harare: University of Zimbabwe.
- Crystal, D. 1997. *The Cambridge Encyclopedia of Language*. Second Edition. Cambridge/New York: Cambridge University Press.
- Kipfer, B.A. 1984. *Workbook on Lexicography: A Course for Dictionary Users with a Glossary of English Lexicographical Terms*. Exeter: University of Exeter.
- Magwa, W. 2002. The Shona Writing System: An Analysis of its Problems and Possible Solutions. *Zambezia* 29(1): 1-11.
- Ndlovu, F. 1998. *Considering Factors that are Relevant to Adopting and Adapting COBUILD Defining Formats for the General Ndebele Dictionary*. Unpublished M.A. Thesis. Harare: University of Zimbabwe.
- Svensén, B. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. New York: Oxford University Press.
- Tullock, S. 1994. *Reader's Digest: Complete Word Finder*. Oxford: Reader's Digest.

Target Users' Expectations versus the Actual Compilation of a Shona Children's Dictionary

Esau Mangoya, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe (emangoya@arts.uz.ac.zw)*

Abstract: The article discusses the challenges that confronted the team of compilers working on the monolingual Shona Children's Dictionary (henceforth SCD). It looks at the active involvement of the target users in shaping the project and discusses the considerations for the implementation of their recommendations. Matters of concern include issues of headword selection, especially problems of dialect representation in the dictionary. The article also discusses the inclusion of grammatical information such as tone marking. The SCD is one of the many efforts to promote and raise the functions of the indigenous languages in Zimbabwe. Of special interest in this article is therefore the discussion on the inclusion of international words like those denoting weight, distance and mathematical and scientific concepts. The article also considers the treatment of taboo words in the SCD, which takes into account that society has to open up culturally to be able to unequivocally view some of the aspects of life with which present-day children are confronted. The article lastly discusses how a balance was reached for accommodating the target users' proposal for the inclusion of an elaborate back matter in view of the limited available scope and space of the SCD.

Keywords: TARGET USERS, DIALECT REPRESENTATION, INTERNATIONAL WORDS, TABOO WORDS, TONE MARKING, BACK MATTER, DICTIONARY, LEXICOGRAPHY

Opsomming: Teikengebruikers se verwagtinge teenoor die werklike samestelling van 'n Sjonakinderwoordeboek. Die artikel bespreek die uitdagings waarvoor die span samestellers te staan gekom het tydens werk aan die eentalige Sjonakinderwoordeboek (voortaan SKW). Dit kyk na die aktiewe betrokkenheid van die teikengebruikers by die ontwikkeling van die projek en bespreek die oorwegings vir die uitvoering van hulle aanbevelings. Sake van belang sluit in kwessies van trefwoordkeuse, veral probleme van dialekverteenvoerdiging in die woordeboek. Die artikel bespreek ook die insluiting van grammatiese inligting soos toonaanduiding. Die SKW is een van die baie pogings om die funksies van die inheemse tale in Zimbabwe te bevorder en te vergroot. Van spesiale belang in hierdie artikel is dus die bespreking oor die insluiting van internasionale woorde soos dié wat gewig, afstand en wiskundige en wetenskaplike konsepte benoem. Die artikel beskou ook die behandeling van taboewoorde in die SKW wat in aanmerking neem dat die gemeenskap kultureel oper moet word om in staat te wees om sommige aspekte van die lewe waarvoor hedendaagse kinders te staan kom, onomwonde te beskou. Laastens bespreek die artikel hoe 'n balans verkry is om die teikengebruikers se voorstel vir die insluiting van 'n uitvoerige agterwerk in te pas in ooreenstemming met die beperkte beskikbare omvang en ruimte van die SKW.

Sleutelwoorde: TEIKENGEBRUIKERS, DIALEKVERTEENWOORDIGING, INTERNASIONALE WOORDE, TABOEWOORDE, TOONAANDUIDING, AGTERWERK, WOORDEBOEK, LEKSIKOGRAFIE

Introduction

A team of three people at the University of Zimbabwe's African Languages Research Institute (ALRI) embarked on a pioneering project to compile a monolingual Shona Children's Dictionary (henceforth SCD) whose main target users are school children at primary level. This team worked closely with teachers and parents who automatically became an interested party in the dictionary project, consequently constituting the majority of the target users. The project started in 2002 and the dictionary was supposed to be published in 2006. However, there were technical hitches which saw the draft of the dictionary being submitted to the publishers only in 2009.

In Zimbabwe, a culture of dictionary compilation and use has not been thoroughly developed. Consequently, most of the lexicographic products in the Zimbabwean context come as pioneering projects. In societies with a long history of lexicography, it is easy to identify the gaps that a newly initiated dictionary project is meant to fill within the context of other earlier and previously published dictionary products. In such situations, both the compiler and the target users have a common goal and expectation. The target users' needs are easy to identify and satisfy since the compiler will not be starting a completely new process. However, despite the varying needs of the target users and the difference in the challenges faced by compilers working on different dictionary projects, the target users' needs have to remain a priority. Hartmann (2001: 80) says: "Ultimately all dictionaries are motivated by and judged against the lexical needs of those who consult them."

Emphasis is always laid on the need to address the target users' expectations. Gouws (1993: 29) says: "It is an accepted fact that every dictionary has to be compiled in accordance with specific needs and demands of a well-defined target user." Here the task for the compiler is clearly defined. The parameters for the production of a "perfect" dictionary are delineated. The expectations are that a dictionary must have authority, scholarship and precision (Landau 2001: 6). At the same time, it has to be considered that their expectations differ for various reasons, including the linguistic and the socio-cultural nature of the communities for which the dictionaries are compiled. As a result, each dictionary project may face problems that remain peculiar to it. The SCD team conducted outreach programmes and workshops with the aim of determining the actual needs and expectations of the target users. Primary school teachers, primary school children and interested and knowledgeable parents were consulted during these programmes and workshops. This article discusses the challenges that the team compiling the SCD faced in view of the recommendations given by the target users. The problematic areas that the team had to deal

with are mainly dialectal representation, tone marking for the purpose of pronunciation, inclusion of international words, the treatment of words deemed culturally taboo and the setting of the back matter.

Dialectal Representation

Shona has five major dialects namely: Karanga, Zezuru, Ndau, Manyika and Korekore. What the research team gathered from the outreach programmes and workshops was that in the past there has been preferential treatment of the Zezuru and Karanga dialects in the Shona resource books used at primary school level. These two dialects have minimal differences compared to the rest of the dialects. This tended to disadvantage pupils who speak other dialects and it was alleged that there was a high failure rate in Shona examinations among them. Zezuru and Karanga which are more central to the whole Shona-speaking community are inclined to dominate and the rest of the dialects play a peripheral role. The feeling was that children particularly from Ndau- and Korekore-speaking areas fail to comprehend the vocabulary from the central dialects. Manyika figured less prominently in the debate as it forms a buffer dialect that is in contact with the central dialects and Ndau and Korekore.

Doke (1931: 81) laid down the basic principle that was to be followed in the compilation of Shona dictionaries. He recommended that if a Shona dictionary should be compiled, it has to be as inclusive as possible of words from Zezuru, Karanga, Manyika and Ndau, but for the time being Korekore words have to be admitted sparingly. On the contrary, earlier in his report, on page 9, he shows reservations about the inclusion of Ndau words when he says: "Into their vocabulary is quite a noticeable infiltration of Zulu following the conquest of Gungunhana." This gives the impression that from the time of the defeat of the Ndau, their dialect is no longer pure. As a result Ndau and Korekore cannot without reservations be seriously considered as dialects of Shona. Today the speakers of these dialects are fighting to have their dialects recognized as belonging to the Shona language. A new environmental situation has given them reason to clamour for recognition of their dialects. This is in view of the recommendations by the National Language Advisory Panel of 1998. In the report, Chimhundu et al. (1998: 36) say: "Use of indigenous/local languages should be made mandatory in schools firstly as subjects and later as the medium of instruction." Although Korekore and Ndau are dialects of Shona, their clamour for recognition is similar to that of some Zimbabwean indigenous languages which ought also, according to the report, to be granted the same right as the other languages.

The SCD team had to do strategic research in order to cover the so-called neglected dialects as much as possible. After going through the usual headword selection process in which student research assistants were involved, the team printed a list of the collected headwords. This also saw a substantial number of headwords from these neglected dialects appearing. The next stage

was to engage the target users, comprising teachers, children and parents. The group of teachers that was engaged firstly had to be speakers of one of these neglected dialects, and secondly had to be teaching in schools in areas where these dialects are spoken. These teachers had the advantage of being speakers and practitioners who were knowledgeable about the problems their pupils face when learning Shona as a subject.

In this way, an additional list of headwords was collected. However, at the same time this strategy also encountered its own problems. As is always the case with groups that feel excluded, there is a tendency to be emotional about their dialects and languages, trying to be as inclusive and exhaustive as possible. Therefore, because of this, a problem surfaced with regard to the many synonymous headwords from the Ndaou and Korekore dialects in comparison with the single headwords from the central dialects as shown in (1) and (2) below.

- | | | |
|-----|--|---|
| (1) | central dialects headword
chikuvauro

guvhu | headwords from Ndaou
<i>chikwa, chichayo chiroo</i> (wooden tool used for flattening the floor)
<i>chikuwuhu, chikumvu, mukuwuhu</i> (navel) |
| (2) | central dialects headword
gangamwiwe

gangemukange | headwords from Korekore
<i>fengu, topo, mupagada</i> (type of inedible pumpkin)
<i>madzimba, majochwe</i> (type of ant) |

There came plenty of headwords from the excited group of target users who had received the opportunity to show the linguistic ability of their dialects. This resulted in a variety of synonyms and variants to single headwords from the central dialects.

This obviously was a strenuous effort by the speakers to try to prove as much as possible that their dialects have plenty of words that are always omitted. This great number of words posed yet another problem to the dictionary compilers because not all the headwords that were given would be of complete equivalence to the headwords from the central dialects and even amongst the presumed peripheral dialects themselves. This would result in a situation whereby words from the presumed peripheral dialects dominate at the expense of the central dialects. The other problem was how to determine the level of synonymy the terms from the peripheral dialects have among themselves and to those from the central dialects. It seems as if these dialects have a higher level of term productivity than the central dialects.

This touches on the sociolinguistic situation of the people in Zimbabwe. There is a reaction to the colonial history in which the administrative demarcations have become tribal demarcations as well. Dialects become a feature of speech that serves to identify the speaker as a member of a particular speech community, past or present (Landau 2001: 220). This provides the speakers of

dialects with arguments as they protest at being deliberately excluded where the rule and policy is that the local and indigenous languages have to be empowered without discrimination. The empowerment of languages, as given in the Advisory Panel Report, also means all the dialects have to enjoy equal status. After independence, the policy has been set as a goal to empower the indigenous languages which have always come second to English. The policy sounds all-embracing, implying that there has to be equal treatment of the indigenous languages and implicitly their dialectal varieties. As pointed out by Landau (2001: 219), forms of expression and pronunciation limited to a particular region have long been regarded as objects of contempt and ridicule. This automatically implies that lexical items from those regions are discriminated against when it comes to entering them into the dictionary. This is a feature not peculiar to the Zimbabwean situation only. In general lexicographic works elsewhere, it has also been pointed out that dialects are considered substandard if not plainly incorrect (Landau 2001: 220). The assertion is showing that dialect issues are always problematic. Wherever they feature, they present emotional sentiments. In the Korekore and Ndaou situation, their exclusion or selective inclusion leads to protest by speakers who see it as an injustice.

To omit or partly omit some dialect forms such as those from Korekore and Ndaou as suggested by Doke, should serve as a warning not to disregard peripheral dialects. However, sidelined dialects will always have their headwords not well recognised in other Shona-speaking areas. As a result, words from these dialects are not quite as familiar to many of the compilers compared to items from the central dialects. The forms of these recognised dialects are always the first to be included in the dictionaries. The central varieties also enjoy being used as the standard, their vocabulary featuring prominently in literary publications which are normally the source for headwords in the dictionaries. Consequently, Korekore and Ndaou dialects have always been sidelined in Shona publications including dictionaries that are purported to be comprehensive.

There has to be a policy for handling the dialectal items. As one way of standardising Shona, Doke (1931: 81) recommended that a dictionary has to be a collection of headwords gathered from the different dialects. This has however been found to be a too simplistic way of dealing with the persistent question of unfair treatment of dialects in Shona dictionaries. With every publication there have been allegations by speakers of other dialects that there were no headwords included from their dialects. These groups that maintain that their dialects are not being considered in written Shona also claim that the pass rate in Shona examinations in those areas where Ndaou and Korekore are spoken are perpetually low because pupils are always made to study texts representing the central dialects. This is the major reason why target users from those areas were engaged as a reference group. However, with the abundance of headwords coming from these target users, some of them had to be eliminated to balance the dialect ratio in the dictionary. The target users were asked to rank

the Ndaou and Korekore words they had provided in order of preference based on their knowledge of the precise equivalence of these words with the headwords from the central dialects. Eventually the ordering of the above examples was reached as shown under (3) and (4).

- | | | |
|-----|--|---|
| (3) | central dialects headword
chikuvauro

guvhu | headwords from Ndaou
(2) <i>chikwa</i> , (3) <i>chichayo</i> , (1) <i>chiroo</i> (wooden tool used for flattening the floor),
(1) <i>chikuvhu</i> , (3) <i>chikumvu</i> , (2) <i>mukuvhu</i> (navel) |
| (4) | central dialects headword
gangamwiwe

gangemukange | headwords from Korekore
(3) <i>fengu</i> , (2) <i>opo</i> , (1) <i>mupagada</i> (type of inedible pumpkin)
(1) <i>madzimba</i> , (2) <i>jochwe</i> (type of ant) |

This ranking was helpful in eliminating headwords from Korekore and Ndaou as all those indicated as first were also considered for entry into the dictionary.

Tone marking

Tone is a fundamental aspect of language. It refers to the distinctive pitch level of a syllable (Crystal 1983: 353). In Shona, the tone contrasts of syllables in words are vital in bringing out the meaning of words. Indicating the tone of words in a dictionary is also important in demonstrating the way words are pronounced. The issue of tone marking became problematic when a recommendation against its inclusion was suggested by the prospective target users during the workshops and outreach programmes. As compilers become well acquainted with dictionary making, some of the practices become routine, usually laid down as standard. ALRI's first two monolingual dictionaries *Duramazwi reChiShona* (1996) and the advanced *Duramazwi Guru reChiShona* (2001) did not only include tone markings but went a step further by coining the Shona terms for high and low tone. These are K from *kwidza* (raise) for high tone and D from *dzikisa* (lower) for low tone, as shown in example (5) of the headword *akapero* (acapella music).

- (5) **akapero** DDKD z 9

The headword has four syllables **a+ka+pe+ro**. D stands for low tone and K for high tone. So the first two syllables **a** and **ka** marked by DD are low tones, the third syllable **pe** marked by K is high tone and the last one **ro** marked again by D is low tone. After the tone comes other vital information like **z** from *zita* (noun) which is of class 9, then followed by the definition. In Shona lexicography, tone markings are relatively new grammatical coinages that are being

standardised not only in the SCD, but also in other ongoing projects that deal with Shona language and grammar.

The target users' position can be understood. The introduction of new elements such as tone marking which gives specialised grammatical information vital to the pronunciation of words is mostly regarded with reservation. The main question arose whether the SCD team had to include information which did not interest the target users. On the other hand it can also be asked if tone marking should be omitted simply because some target users are not fully acquainted with its use. The problem was complicated by the fact that the speakers of the central dialects needed assistance in the pronunciation of the many new headwords from Ndau and Korekore as much as the speakers of these peripheral dialects needed to acquaint themselves with the tone of the headwords from the central dialects.

As a result, the inclusion of tone was considered helpful to users when pronouncing words. It may be argued that children have not yet been exposed to these complicated grammatical aspects such as tone. However, this information may assist the teacher working with the children. A dictionary that is inclusive has an added advantage to those who might need that extra information. The teacher can help pupils extract what is relevant to them. At the same time the availability of a dictionary which has extra details leaves the teacher assisting the pupils with additional information although the pupils may not use it themselves. This does not mean that the suggestions by the target users had to be disregarded completely, but the merit of such suggestions had to be reviewed as well.

As a reference work, a dictionary has many functions. The SCD team agreed that there was need to have the tone marked in the data base and again recommended the idea of indicating tone in the dictionary to the working groups consisting of target users. The admissibility of tone markings in the dictionary were again assessed in the sample headwords and definitions which were taken to the target users during the outreach and seminar programmes. When the same aspects are again brought up for discussion and viewed in print the target users are normally forthcoming with suggestions for improvement. Tone marking was eventually accepted as an essential tool in the dictionary.

International Words

What have come to be regarded as international words are those ones that have originated in particular languages but are now used across the world in almost every language. At this stage, they have not been adopted in the Shona language yet but because of their specialised sphere of use are always mentioned in Shona texts. Examples of these are words denoting weight, length and scientific and mathematical terms, as demonstrated in the examples under (6).

- (6) **weight:** kilogramme, tone
length: centimetre, metre, kilometre
scientific and mathematical terms: atom, equation, equator

These and many others are recognised standard terms in many languages. The initial view on the possible inclusion of these terms in the SCD was purist in nature. As they were considered foreign, they were seen as unsuitable for a dictionary meant to strengthen children's communication skills in the Shona language. However, because the purpose of compiling the dictionary was to empower the Shona language, brought some change in perception. The language has to be developed so that it would be able to describe scientific concepts. The target users did not accept the idea of outright adoption but opted for the coinage of "pure" Shona terms. A precedence has already been set before when coinage was attempted in Hannan's (1981) *Standard Shona Dictionary* which yielded words like *dutavanhu* (bus), *hambautare* (motorcar) and *bhizautare* (bicycle) in an attempt to avoid using English terms. Naturally, when people speak they say *bhazi* (bus), *motokari* (motorcar) and *bhasikoro* (bicycle) which are all adoptives. People hardly use the once prescribed coinages.

Being aware of the fate of some coinages in Hannan's dictionary, the SCD team of compilers argued against their inclusion. These coinages would end up as redundant entries in the dictionary and the terminological gap they were meant to fill would remain open. If its natural behaviour is considered, a language gives preference to adoptives that are close to the source language in spelling and pronunciation. However, there was the difficulty that some of the borrowed headwords had to be adapted to suit the Shona spelling system. The examples of the international words above were entered as shown under (7):

- (7) **weight:** *kirogiramu, tani*
length: *sendimita, mita, kiromita*
scientific and mathematical terms: *atomu, ikwezheni, ikweta*

The Shona alphabet does not have the letter *l* but the trend has been that it can be substituted by *r*, hence the change from the English term *kilogramme* to *kirogiramu*. In Shona, the vowel *o* cannot be realised as /a/ as in the case of the English word *tone* in which the first syllable sounds like /ta/. The new adoptive for *tone* is now /tani/. Similarly the initial vowel *e* in the word *equator* is realised as /i/, therefore the change to /ikweta/. All syllables in Shona are marked by a vowel, hence the change of the word *mitre* to /mita/ where it now has the syllables /mi/ and /ta/.

The adoptives had to be adapted to the Shona writing system. This was found to be an important step towards empowering the Shona language, particularly in a book targeted at a group that is at the foundation level of the education system. This will give the target users the confidence to use the language in fields where other languages such as English have always dominated. The target users' creative abilities will be enhanced. This will broaden their view of

the world in which they, using their mother tongue, can be active participants. The belief is that what has already been initiated will continue and eventually the Shona language will not be void of such important terminology.

Taboo words

Another category of words that needed special attention was taboo words. Taboo words are known to speakers but avoided in some, most or all forms or contexts of speech (Matthews 1997: 371). Considerations regarding taboo words can be culture-bound. Culture, however, changes, causing the taboo tag given to some lexical elements to be removed partly or wholly. It is important for dictionary compilers to fully know and understand the feelings of the target users on some important aspects of the dictionary they compile. The issue becomes much more complicated when questions arise on whether some aspects are culturally acceptable to be entered into the dictionary. The dictionary compiler has to have a full understanding of the users whom the compilers have sometimes ignored so that there was little appreciation for some dictionary products. The position of target users has to be ascertained in order to avoid criticism of the dictionary product. About the dictionary user, Hartmann (2001: 80) says: "We still know very little about the dictionary user whom several authorities have called the familiar stranger, presumably because we take him for granted without bothering further."

Hartmann's assertion assumes that the participation of the target users in making their recommendations towards the contents of a dictionary may help enhance the lexicographer's insights leading to a lexicographic product that answers the user's needs. The question that has to be answered is whether the shape of a dictionary should be completely determined by the target users. Does the compiler have some authority and power to incorporate new challenges that can help change the society's view of the world? Such challenges were encountered in the compilation of the SCD with regard to taboo words. The target users, particularly the teachers and parents, were clear and adamant that human private parts and some physiological terms would not be acceptable in a dictionary meant for children. According to their view, these children are at the critical stage where there has to be a total induction into their culture. Part of the culture forbids having access to knowledge of these taboo words, particularly by primary school children.

The refusal to enter headwords deemed to be breaching the ethical code of the society is not peculiar to the Zimbabwean situation. All communities at the initial stages of their lexicographic tradition seem to follow the trend of reservations on the inclusion of what is deemed taboo (see Landau 2001 : 228). There is the feeling that the exclusion of taboo words contributes towards the upholding of ethical values.

Because the situation of the SCD was found to be rather unique, it brought

a number of challenges. Firstly, by aiming primary school children as target users, the dictionary belongs to a type of its own in comparison to the earlier publications *Duramazwi reChiShona* (1996), targeted at ordinary level students and *Duramazwi Guru reChiShona* (2001), targeted at college and university level students. Secondly, there are new challenges not only to the Zimbabwean society but to societies worldwide that need to be addressed to this age group as a matter of urgency. One channel of doing so is through books such as the SCD. Because the world is changing fast, what society once might have strongly considered taboo might need to be reconsidered in today's world. One of the biggest challenges that have besieged the world is the advent of HIV and AIDS. There is also the problem of the early indulgence in sexual intercourse by today's children and the dangers that go with such activities. There is the rampant problem of child abuse and rape. These and other challenges have resulted in the need to teach and inform children as a precautionary measure about the dangers of the various sexually transmitted diseases and the eventual consequences. All these problems cannot be fully understood if discussed in unclear and veiled language.

In Zimbabwe, the language situation is such that English has been accorded the official status and Shona has always come second to English. This has led to a culture in which anything presumed to contain difficult and foreign concepts should be explained in English. Everything is left to feature in the English texts to which the majority of school-going children eventually have access. However, since English in Zimbabwe is learnt as a second language, there is little consideration at what time or stage in the children's educational development, they will be able to fully comprehend English. This has consequently led to the exclusion and avoidance of vital terms in Shona textbooks. The SCD was caught up in a similar dilemma. As a result, the teachers who have become accustomed to this culture over the years oppose the inclusion of these taboo words in a children's dictionary, their argument being that it will be antisocial to expose children to such terminology.

After serious discussions, it was agreed that some of these presumed taboo words had to be included as long as their role was to give information found to be vital to the children. These are terms specifying the body parts, describing the signs and symptoms of venereal diseases and naming some sexual activities which cannot be avoided mentioning when teaching the children. Examples of such headwords are listed under (8).

- (8) *mhata* (anus)
mheche (vagina)
mboro (penis)
maronyo (sperm/semen)
siki (venereal disease)
njovhera (type of venereal disease)
-svira (have sexual intercourse)

These were allowed into the dictionary because of needs and challenges. However, despite their relatedness to the words under (8), there were some reservations on the inclusion of another group of words which were deemed unnecessary for the level of primary school children. These are common taboo words describing the nature of certain sexual organs or the manner of some sexual activities. Examples of words about which teachers had reservations are given under (9).

- (9) *beche* (big vagina)
zimboro (big penis)
nzvonyo (circumcised penis)
mukodo (sexual intercourse from the back)
mufongo (sexual intercourse from the back)

It was agreed that specimens of normal use should be given for nouns. It was felt that including words meant to give an obscure or indistinct explanation of the mentioned organs and activities would actually destroy the main aim of teaching, and could create confusion and uncertainty in the children that will in reality be contrary to the aim and purpose of the words' inclusion. So the taboo words would have to be selectively entered into the dictionary. Inclusion of some of the vital words will assist in the newly introduced subjects on HIV and AIDS. The children are also being taught issues of human rights in which child sexual abuse would be discussed. Children should know what abuse is and should be able to describe it if the need arises. Because teachers and parents should be able to talk to the children, the vocabulary that was considered to be taboo has to be made normal and familiar by including it in a general dictionary such as the SCD.

Back Matter

One of the main concerns of a dictionary is to give comprehensive information. A dictionary of the SCD type has to contain information on issues children might not have had a true experience of and direct exposure to in their lives. For example, a child born and bred in town might not be fully aware of aspects that are predominantly rural and vice versa. As a result the dictionary is expected to play an informative role with many aspects reserved for the back matter. As has been pointed out, the SCD project generated much interest among teachers who played a major role as a reference group advising the compiling team. There was also much interest from the community at large who also felt the dictionary would be a useful resource for their children, particularly when writing the Grade 7 Shona examinations. These are vital examinations which need to be passed before children can proceed to secondary level. There have always been complaints from teachers and parents about the

lack of resource materials for the Shona examinations at primary school level. The idea was to try as much as possible to provide for the different sections of the examinations usually written at this level. So the SCD project was considered to be a major step in compensating for this lack. This resulted in a number of suggestions on additions to be made to the dictionary over and above the general headwords and their definitions. All the additional information could be placed in the dictionary as back matter. It was suggested that this back matter could include a list of people and their professions, a list of animals and their young ones, illustrations of kitchen utensils, a list of common proverbs and idioms and a list of kinship terms. While this would generally be a recommendable idea, it had its own limitations. An example of how this was to be done is shown under (10).

- (10) *Tete ihanzvadzi yababa* (One's aunt is one's father's sister)
Tezvava ndibaba vomudzimai (One's father-in-law is one's wife's father)

From the first example, the words *tete* (aunt), *hanzvadzi* (sister) and *baba* (father) would have been entered as headwords in the dictionary. Such is also the case with words like *tezvava* (father-in-law), *baba* (father) and *mudzimai* (wife) from the second example. They again appear as repeat entries in the back matter where they, despite being combined in a new context, are in fact a duplication within the same dictionary. Such is also the case with illustrations and all the aspects that were suggested for inclusion in the back matter. A careful selection had to be made to include only those that would bring extra information beyond what could already be found in the headword section. An example of a topographical illustration that was chosen, is a mountain. The illustration would then show *gomo* (mountain), *mawere* (cliff), *mupata* (valley) and *bako* (cave).

The lexicographer is always contending with the issue of space in the dictionary. At the same time, it has to be understood that the dictionary has its limitations as a reference book. It cannot be the ultimate solution to everything related to examinations. Moreover, as this is a general and not an encyclopaedic dictionary, it was found to be difficult to include all the suggested aspects in the back matter limited by size and detail. Eventually it was resolved that the dictionary could include trimmed lists. However, the compilers had to be careful not to outrightly refuse special requests from the reference group which forms a major component of the target users.

As a compromise it was agreed that some of the requests would actually be better incorporated in school text books. A dictionary was only meant to address problems that usually arise out of the subject matter of text books. It was not meant to provide solutions for these problems and at the same time function as a text book which addresses all aspects of Shona as a school subject. It was hoped that the compilation of the SCD would help generate interest by other relevant participants to write Shona books for primary school level.

Conclusion

It is every dictionary compiler's wish to produce a dictionary that fully makes provision for the needs of target users. As a result, it is necessary to consult its prospective target users. The SCD project involved them to a great extent. However, each dictionary project faces challenges of a general nature and at times some which are peculiar to it. This article has discussed those which are peculiar to the SCD, starting with the nature of the target user group which mostly comprised teachers and parents, despite the fact that the real target users are children. Children as a group have to be guided, hence they are mostly represented by the teachers and parents so that their own contribution is minimal.

The article has discussed the discrepancies that in the past occurred in the treatment of Shona dialects. It has looked at the pressure exerted by the speakers of the neglected dialects, showing how the new scenario has impacted on the SCD project, including the problem of excessive headwords that had to be reduced to bring about a balance with other dialects. Dictionary making is also important in introducing and standardising new aspects for the benefit of target users. As a result, the SCD tried to uphold the new phenomenon of marking the tone, which is vital for the pronunciation of headwords. Although target users are not yet comfortable with tone marking, it is hoped that once introduced and maintained as a lexicographic tradition, target users will soon be familiar with it.

This being a pioneering project of a unique nature, it had to confront new challenges. As a result it included some adopted international words. This is an attempt to bring Shona on the same level of communicative ability with other developed languages so that foreign languages need not to be used in discussions dealing with foreign concepts. As the world is changing, this also demands cultural change. Therefore, the once considered taboo words found their way into the SCD. The aim is to enable communication with the children, so resources such as the SCD must have a high information level. For this reason, some additional carefully selected material has been added as back matter to the general dictionary, at the same time taking limitations of space and detail into account. Challenges occur in dictionary making but they can be discussed and overcome in some way or other.

References

- Chimhundu, H. et al.** 1998. *Report on the Formulation of a National Language Policy*. Harare: Ministry of Sport, Recreation and Culture.
- Chimhundu, H. (Ed.)**. 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. (Ed.)**. 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Doke, C.** 1931. *Report on the Unification of Shona Dialects Carried Out under the Auspices of the Government of Southern Rhodesia and the Carnegie Corporation*. Hertford: Stephen Austin.

- Crystal, D.** 1991. *A Dictionary of Linguistics and Phonetics*. Third Edition. Cambridge: Basil Blackwell.
- Gouws, R.H.** 1993. Afrikaans Learner's Dictionaries for a Multilingual South Africa. *Lexikos* 3: 29-66.
- Hannan, M.** 1959. *Standard Shona Dictionary*. Salisbury: The College Press.
- Hartmann, R.R.K.** 2001. *Teaching and Researching Lexicography*. London: Longman.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Matthews, P.H.** 1997. *Oxford Concise Dictionary of Linguistics*. Oxford: Oxford University Press.
- Mpofu, N. et al. (Eds.).** 2004. *Duramazwi reUrapu neUtano*. Gweru: Mambo Press.

Duramazwi reDudziramutauro neUvaranomwe against the Background of Shona Lexicography

Gift Mheta, *Department of Linguistics, University of the Western Cape,
Bellville, Republic of South Africa (2971565@uwc.ac.za)*

and

Maxwell Kadenge, *Department of Linguistics, School of Literature and
Language Studies, University of the Witwatersrand, Johannesburg, Republic of
South Africa (maxwell.kadenge@wits.ac.za)*

Abstract: This article considers and analyses *Duramazwi reDudziramutauro neUvaranomwe* (henceforth DRU), a 2007 product of the University of Zimbabwe's African Languages Research Institute (ALRI). It evaluates DRU which is the first terminological dictionary dealing with Shona literary and linguistic terms. Furthermore it traces the compilation process of this dictionary, describing the stages of collection, selection, entry and definition of terms. It also examines the microstructure and the macrostructure of DRU aimed at target users who are both students and teachers in high schools and tertiary institutions offering Shona as a subject or course.

Keywords: BILINGUAL, COLLECTION, COMPILATION, COMPUTATIONAL, CORPUS, DEFINITION, DEVELOPMENT, ENTRY, HEADWORD, LEXICOGRAPHY, LINGUISTIC, LITERARY, MACROSTRUCTURE, MICROSTRUCTURE, MONOLINGUAL, SELECTION, TERMINOLOGY

Opsomming: *Duramazwi reDudziramutauro neUvaranomwe teen die agtergrond van Sjonaleksikografie.* Hierdie artikel beskou en ontleed *Duramazwi reDudziramutauro neUvaranomwe* (voortaan DRU), 'n 2007 produk van die Universiteit van Zimbabwe se African Languages Research Institute (ALRI). Dit beoordeel DRU wat die eerste terminologiese woordeboek is wat Sjona- literêre en linguïstiese terme behandel. Verder gaan dit die samestellingsproses na deur die stadiums van versameling, keuse, inskrywing en definiëring van terme te beskryf. Dit ondersoek ook die mikrostruktuur en die makrostruktuur van DRU, gerig op teikengebruikers wat sowel studente as opvoeders in hoërskole en tersiêre instellings is wat Sjona as 'n vak of kursus aanbied.

Sleutelwoorde: TWEETALIG, VERSAMELING, SAMESTELLING, REKENAAR-, KORPUS, DEFINISIE, ONTWIKKELING, INSKRYWING, TREFWOORD, LEKSIKOGRAFIE, LINGUISTIES, LITERÊR, MAKROSTRUKTUUR, MIKROSTRUKTUUR, EENTALIG, KEUSE, TERMINOLOGIE

Introduction

According to Kipfer (1984), lexicography is the study and practice of dictionary-making. A broad definition that transcends the stress placed on dictionary-making to include other related activities is well expressed by Svensén (1993: 1) when he says:

Lexicography is a branch of applied linguistics which consists in observing, collecting, selecting and describing units from the stock of words and word combinations in one or more languages ... Lexicography also includes the development and description of theories and methods which are to be the basis of this activity.

Lexicography contributes significantly to terminology development in languages (Mheta 2007: 11). As noted by the National Language Policy Advisory Panel (1998: 44), one area of linguistic research that is now recognized internationally as being important for language standardization and development is lexicography or dictionary-making. Traditionally, standardization is viewed as involving the twin process of codification, that is, making rules about the writing system, and grammar and elaboration, which refers to expanding and intellectualizing the vocabulary and diversifying the functions of the language so that it becomes an effective tool for communication even in specialized fields such as literature and linguistics, with which this article deals. The National Language Policy Advisory Panel further mentions that in both these processes, a standard dictionary has an important normative influence, as its users will accept and apply the lexicographers' descriptions of word-forms and their statements on their meanings. Having noted the importance of language standardization, which includes terminology development, it is necessary to trace the history of Shona lexicography that has positively contributed to Shona terminology development.

Background

The Shona language has a lexicographic history of nearly one and half centuries. It stretches back to the period preceding the arrival of the Pioneer Column and the establishment of Southern Rhodesia by the British South Africa Company when writings on the Shona language in English and compilations of Shona vocabulary are already found, such as the first bilingual dictionary in Shona and English, *The Languages of Mosambique* (1856) by Bleek, which was followed by Elliot's *Dictionary of the Ndebele and Shona Languages* (1887) and Hartmann's *English–Mashona Dictionary* (1894) (see Doke 2005). Ever since then, the trend has been that of compiling bilingual linguistic dictionaries. The most popular of these dictionaries is Hannan's *Standard Shona Dictionary* published in 1959 and subsequently revised in 1984. The trend of compiling bilingual dictionaries continued until 1996 when the African Languages Lexical (ALLEX) Project, then housed in the Department of African Languages and Literature at

the University of Zimbabwe, produced the first-ever monolingual Shona dictionary, *Duramazwi reChiShona*. This was followed in 2001 by *Duramazwi Guru reChiShona* (hereafter DGR) which is an advanced Shona monolingual dictionary. Since the publication of this dictionary, there has been a paradigm shift in Shona lexicography, from the production of general linguistic dictionaries to the compilation of specialized dictionaries. Examples of Shona specialized dictionaries are *Duramazwi reUrapi neUtano*, a dictionary of biomedical terms published in 2004, and *Duramazwi reMimhanzi*, a dictionary of musical terms published in 2005. The most recent terminological dictionary in Shona is *Duramazwi reDudziramutauro noUvaranomwe* (hereafter DRU), dealing with Shona literary and linguistic terms and published in 2007. Currently there are also running projects on dictionaries in specialized fields such as *Duramazwi Guru reMimhanzi* (an advanced Shona musical terms dictionary) and a Shona medical terms dictionary. The two are sub-projects of the NUFU-funded Cross-border Languages Project that is a network project that involves researchers at the African Languages Research Institute (ALRI), the Eduardo Mondlane University in Mozambique and the University of Oslo in Norway. It is in the context of these lexicographic activities that Shona terminology development will be discussed.

Despite the relatively long lexicographic history, the Shona language is still very much like other African languages in lacking adequate terminology in specialized fields. Like many indigenous languages in Africa, there are many problems that are un conducive to the effective use of Shona in specialized fields, one of the major ones being the absence of standard terminology in technical areas. Discussing the problem of terminology development in Ndebele, Hadebe (2006) states that the language situation in Zimbabwe, where African languages are confined to the unofficial domains, has led to an underdevelopment of terminology in almost all fields. This is a true reflection of the current status of the Shona language as well, and the general situation of indigenous languages in most African countries. For example, Shona is still taught in English at the University of Zimbabwe, the oldest and largest academic institution in Zimbabwe. The reason for this is that terminology for teaching Shona has not been standardized. Tertiary institutions have been using different terms to refer to the same literary and linguistic concepts. The rectification of this anomaly necessitated the compilation of DRU.

To appreciate fully why DRU may be considered important in the development of Shona lexicography, it is necessary to look at the procedures followed by the ALRI lexicographic team that culminated in the successful compilation of this dictionary. How the different stages of collection, selection, entry and definition took place will be described in the following sections.

Collection

Although DRU is the first dictionary of its kind, previous Shona dictionaries were used as sources for collecting terms for it. These dictionaries include Dale

(1981), Hannan (1984), Chimhundu (1996, 2001), Mpofu et al. (2004), Mheta (2005) and ZIMSEC/CDU (2006). They provided the bulk of the headwords found in DRU.

Apart from dictionaries, headwords were also collected from books used in teaching Shona literature and grammar in secondary and high schools, teachers' colleges and universities. Such books include Doke (1931, 2005), Fortune (1955, 1972), Chimhundu (1995, 1998, 2000) and Magwa (1999).

Other headwords were drawn from the African Languages Lexical Project (ALLEX) Shona corpus. According to Crystal (1997: 95), a corpus refers to a collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting-point for linguistic description or as a means of verifying hypotheses about a language. A similar definition is given by Kennedy (1998: 1), who describes a corpus as "a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description". The ALLEX Shona corpus is more than just a written text. It is a systematic, well-designed and -selected collection of written texts and transcribed speech in electronic form. In corpus linguistics terms, it is a dynamic or monitor corpus, i.e. an open-ended machine-readable corpus (see Kennedy 1998: 22.) It is an example of what Sinclair (1992: 382) calls "huge, changing bodies of no finite size, flowing across a set of filters which extract linguistic evidence". It was collected by members of the ALLEX project when it compiled the first corpus-aided monolingual dictionary, *Duramazwi reChiShona* (1996). The collection of the data for the corpus started in 1993 and continues to the present day for purposes of producing more lexicographic projects at ALRI and facilitating other research activities in the Shona language. The Shona corpus, which has 2 962 412 tokens or running words is composed of both written and oral materials from various genres collected from all the Shona-speaking districts of Zimbabwe. The Shona corpus includes Shona literary and linguistic terms. Below is an excerpt adapted from the ALLEX Shona corpus.

ALRI SHONA corpus (2 962 412)

Search for ".*uvaranomwe" in SHONA

12 hits (limited to 1 000)

yora rumwe rudzi rwemhinduro refu **dzouvaranomwe** paya pavanenge vozopindura mibvu mbo Kachirongwa aka kakarerekera **kuuvaranomwe**. Mhinduro dzinoti: (a) Mutambo unor ipa muenzaniso handiti, kuti pangava **neuvvaranomwe**, pangave neronedzero. Ko zvidzid ipa muenzaniso handiti, kuti pangava **neuvvaranomwe**, pangave neronedzero. Ko zvidzid dzokuzvivaraidza sezvakaite marudzi **ouvaranomwe** anoti ngano, nganonyorwa, mitamb kuti svondo rino tichava neronedzero, **ruvaranomwe**, nzwisiso, tsika namagariro. Asi hat kuti svondo rino tichava neronedzero, **ruvaranomwe**, nzwisiso, tsika namagariro. Asi hat inotaridza kuti inzwisiso here kana kuti **ruvaranomwe**. Ongororo zvinotaridza kuti basa rac inotaridza kuti inzwisiso here kana kuti **ruvaranomwe**. Ongororo zvinotaridza kuti basa rac yaya Kachirongwa aka ngakatorwe **sekeuvaranomwe**. Gakava Apa pfungwa dzinosiyana, ana tava kuongorora nezvemutauro uye **uvaranomwe**. Vakabatsirana naProfessor Geor oziva marudzi akasiyana ezvinyorwa **zveuvaranomwe** vasati vaongorora mabhuku aya av

Apart from relying on written sources, both hard copies and electronic versions, the ALRI lexicographic team had to assemble a reference working group (hereafter RWG) from which it obtained more Shona literary and linguistic terms. The RWG comprised Shona language experts from universities. In this regard, experts were drawn from the University of Zimbabwe in Harare, Great Zimbabwe State University in Masvingo and Midlands State University in Gweru. Lecturers were also drawn from teachers colleges such as Hillside in Bulawayo, Bondolfi in Masvingo, Marymount in Mutare and Seke in Chitungwiza. Shona teachers from different provinces in Zimbabwe furthermore formed part of the RWG. The ALRI lexicographic team benefited from this group of Shona language experts because various members of this group represented different institutions and every institution had its own set of Shona literary and linguistic terms. From the RWG it emerged that different terms were used by different institutions to refer to the same literary and linguistic concepts. For example, some institutions used *mapishapisha* while others preferred *makakatanwa* to refer to 'conflict'. Other examples are given in the following table:

Shona Terms at Institution A	Shona Terms at Institution B	English Gloss
<i>dendavabatsiri</i>	<i>nhendo</i>	acknowledgements
<i>hweretwa</i>	<i>izwi rokutapa</i>	adoptive
<i>chidudzirachiito chewadzano</i>	<i>chidudzirachiito chekuita pamwe</i>	adverb of association
<i>nhoredzera yenyaudzira</i>	<i>mutinhirosiri</i>	allophone
<i>pikisadudziro</i>	<i>izwi rokupikisa</i>	antonym
<i>wirirano</i>	<i>sungawirirano</i>	concordial agreement
<i>tsoropodzo</i>	<i>ongororo</i>	critique
<i>chiitomwana</i>	<i>chiitogama</i>	defective verb
<i>nganonyorwa yoruferefeto</i>	<i>nganonyorwa yorufefeto</i>	detective novel
<i>izwi guru</i>	<i>izwi simboti</i>	head word

It should be noted that in some instances a single literary or linguistic concept could be signified by five or even more different signifiers. For example, *turabefu*, *chifema*, *chiturabefu*, *chituramafemo*, *koma* and *nhurabefu* are different words used by different academic institutions to refer to 'comma'. Another term that yielded different representations from different academic institutions is 'auxiliary verb', as shown, for example, by the following words: *chiitobatsiri*, *chiitomhini*, *chiitogasva*, *chiitogama* and *chiitomwana*.

Different institutions represented in the RWG meant that the ALRI lexicographic team ended up collecting many headwords from the group of Shona language experts. At this stage, diversity was an advantage and not a challenge because the target for the lexicographic team was to collect as many literary and linguistic terms as possible. The challenge of dealing with diversity emerged at the selection stage.

Selection

With a set target of 1 500 Shona headwords, it meant that the ALRI lexicographic team had to select the most relevant literary and linguistic terms from the collection they had brought together. The selection was done with the assistance of the RWG. It was through discussions with members of the RWG that the ALRI lexicographic team managed to select the main headwords. As already mentioned in the previous section, the lexicographers had to contend with the challenge of different headwords used to refer to the same literary and linguistic concepts by different institutions. In response to this challenge the lexicographers together with the RWG selected the main headwords on the basis of frequency and acceptability. For example, *demberezeze* referring to the literary term 'suspense' was more frequent in the various institutions represented in the RWG than *dokwaidzo*. For this reason, the former was selected as the main headword that carries the definition and the latter as a synonym as is demonstrated in the excerpt from DRU (p. 51) below:

demberezeze D- z 5 (suspense). Aya manyorerwo enyaya kana maburitsirwo echionwa anoita kuti vaverengi, vaoni kana vateereri vave nechidokwadokwa kana havi yokuda kuziva kuti chii chakazoitika muchinyorwa, chionwa kana nhaurwa. FAN *dokwaidzo* 9.

demberezeze L- n cl. 5 (suspense). This is a way of writing a story or showing a film that retains the interest of the readers, listeners or viewers. SYN *dokwaidzo* 9.

In instances where Shona literary or linguistic terms were found in the ALLEX corpus the frequency of terms in the corpus was compared to that which became apparent from discussions with the RWG. A consensus was always reached as members of the RWG were constantly reminded not to be regional defenders but to base their selections on facts and figures.

Entry

Of interest in this section is how grammatical information for headwords was entered in DRU. The ALRI lexicographic team agreed to mark tone, and indicate grammatical category and class of headwords. For example, the headword **demberezeze** given above in the previous section is a class 5 noun as indicated by z 5 that has a low tone throughout as shown by D-. To this end, the grammatical information for the headwords was shown using the format found in general linguistic dictionaries such as DGR. What comes as an innovation in this dictionary is the morphological description of each formative that is part of the headword. For example, the formatives that compose the headword **chiratidzakure** are presented as follows in DRU (p. 28):

chiratidzakure [chitaridzakure] D- z 7 (far demonstrative) <-ratidz-a + kure.

In this case, the reader is informed that the headword **chiratidzakure** comes from *-ratidz*, a verb radical that means 'show' and *kure* an adverb of place that means 'far'.

Therefore, the added advantage DRU has over general linguistic dictionaries is that it offers a more complete description of entries. Even before the users come to the definition, they will have an idea of the meaning of the headword. In a way, the additional morphological information complements the definitions.

Definition

The ALRI lexicographic team encountered a number of challenges at the definition stage. The most difficult task was how to construct a framework that would assist in composing definitions of the terms with more encyclopedic information than those found in general linguistic dictionaries. Since DRU is encyclopedic in nature, the team had to frame definition guidelines that would ensure the inclusion of most of the encyclopedic information. For instance, definitions for terms on oral literature such as types of dances had to indicate

- (a) whether the dance is traditional or modern,
- (b) the origin and region where it (especially when traditional) is found,
- (c) the performer(s) of the dance as well as its context and the time of its performance,
- (d) the costumes worn, equipment used or instruments played, and
- (e) the type of music involved.

With clearly spelt out definition formats it became easy for the team to draft exhaustive definitions contrasting with the concise ones found in general linguistic dictionaries. A comparison of two definitions, one from DGR and the other from DRU clearly illustrates this point.

Definition from DGR (p. 714):

ngano D- z 9 > 10. *Ngano* inyaya inoronedzera zvakaitika kare, zvingave zvechokwadi kana zvisingakwanisi kuitika. FAN *rungano* 11. *Ngano* inyaya isingabvumwi nemunhu anenge achiudzwa. *Dai wangonditaurirawo chokwadi. Kwete ngano dzauri kundiudza idzi.*

ngano L- n cl. 9 > 10. A *folktale* is a story that relates what happened in the past, which might be fantasy or fact. SYN *rungano* 11. A *folktale* is a story that the listener does not believe. *Just tell me the truth and not folktales.*

Definition from DRU (p. 116)

ngano D- z 9 > 10. *Ngano* inyaya inotaurwa **nasarungano** achironedzera zvakaitika kare, zvingave zvechokwadi kana zvisiri zvechokwadi, kana kutoti

zvisingakwanisi kuitika. Ngano dzinowanzotaurirwa kuvana vadiki nechina-ngwa chokuvadzidzisa nokuvatandadza. Makarekare ngano dzaiitwa muchirimo munguva dzemanheru, vanhu vapedza chisvusvuro chemanheru. Zvidzidzo zvomungano zvinobva muzvinhu zvinenge zvakaitika kune mumwe munhu, asi vatambi vacho vanogona kunge dziri mhuka dzinogona kutaura nokuita zvose zvinoitwa nevanhu. Kune ngano dzine vatambi vanenge vachingova vanhu, dzimwewo dzine vatambi vanenge vari mhuka, kozoti dzimwewo dzinosanganisa vatambi vevanhu nevemhuka, nedzimwe dzimwezve dzine vanhu vanoshanduka kuita mhuka; asi vatambi vese ivava vanenge vachikwanisa kutaurirana sevanhu. Muzviitiko netsaona dzinopinda vatambi vakasiyana-siyana ndimo munobuda zvidzidzo pamusoro pezvinofanira kuitwa nezvisingafaniri kuitwa muupenyu. Saka vateereri vengano kana kuti vana vadiki vanenge vachitodzidziswa pamusoro petsika nokugarisana kwakanaka. FAN *rungano*.

ngano L- n cl. 9 > 10. A *folktale* is a story told by a story-teller, which may be fantasy or fact. Folktales are usually told to children with the intention of teaching and entertaining them. Long ago, folktales were told during the dry season, in the evening, after supper. Lessons in folktales come from what happens to people but the characters in folktales can either be human beings or animals or a combination of both. In folktales, animal characters speak like human beings. In some cases, human beings change into animals. Lessons are drawn from the events that take place in the folktales. In this regard, children are taught their culture through folktales.

The abundance of information on literary and linguistic terms shows that the DRU lexicographic team successfully framed definition guidelines which have helped in presenting as much encyclopedic information as possible. For example, the definition for **ngano** in DGR is concise but the one in DRU is elaborate as it gives almost every finer detail about the headword. It should also be noted that while DGR only defines the headword **ngano** DRU goes a step further by defining the various types of **ngano**, such as **ngano dzakanyorwa** and **ngano dzinotaurwa** given as separate headwords.

Additional Information in the Dictionary

Unlike other general literary and linguistic dictionaries, DRU has an index at the end of the Aa–Zz dictionary section. The index contains English literary and linguistic terms and their Shona equivalents. This section is of great value to readers who wish to quickly look up Shona equivalent terms for English literary and linguistic terms they know without having to refer to the definitions. In addition, it has a grammar section, which comprises explanations of grammatical categories such as nouns, verbs, ideophones and orthographic rules for the Shona language. To this end, the dictionary contains almost every important aspect of the Shona language.

Conclusion

The discussion of DRU shows that the dictionary is indeed an asset to the educational sector in Zimbabwe, which had hitherto lacked adequate and relevant Shona written materials with standardized terminology on literature and linguistics. It is a dictionary that testifies that any specialized area can be taught in the Shona language. It also confirms the fact that all languages, being equally complex, can be used to express any idea. Like any other venture, terminological development in African languages is a big but not insurmountable challenge. The objective of teaching any specialized field in an indigenous language is attainable as evidenced by the publication of DRU which is indeed of importance in the development of the Shona language. It will go a long way in assisting both students and teachers in the mastery of the intricacies of the Shona language.

References

- Bleek, W.H.I. 1856. *The Languages of Mosambique*. London: Harrison.
- Chimhundu, H. 1995. *Zambuko 3*. Harare: College Press.
- Chimhundu, H. 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. 1998. *Zambuko 4*. Harare: College Press.
- Chimhundu, H. 2000. *Zambuko 3 & 4: Bhuku roMudzidzisi*. Harare: College Press.
- Chimhundu, H. 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Chimhundu, H. 2005. *Lexicography and Language Raising: Dictionaries in Zimbabwean Languages*. Public lecture presented as part of the Golden Jubilee Celebrations of the University of Zimbabwe, Harare, 13 April 2005.
- Chimhundu, H. and E. Chabata. 2007. *Duramazwi reDudziramutauro noUvaranomwe*. Gweru: Mambo Press.
- Crystal, D. 1997. *A Dictionary of Linguistics and Phonetics*. Fourth Edition. Oxford: Basil Blackwell.
- Dale, D. 1981. *Duramazwi: A Shona-English Dictionary*. Gweru: Mambo Press.
- Doke, C.M. 2005. *Report on the Unification of Shona Dialects. A Photographic Reprint with an Introduction by H. Chimhundu*. Oslo: ALLEX Project.
- Elliot, W.E. (Ed.). 1887. *Dictionary of the Ndebele and Shona Languages*. London: David Nutt.
- Fortune, G. 1955. *An Analytical Grammar of Shona*. Cape Town/New York/London: Longmans, Green.
- Fortune, G. 1972. *A Guide to Shona Spelling*. Salisbury: Longmans.
- Hadebe, S. 2006. *The Standardisation of the Ndebele Language through Dictionary-Making*. Harare/Oslo: The ALLEX Project.
- Hannan, M. 1959. *Standard Shona Dictionary*. Harare: Literature Bureau.
- Hannan, M. 1984. *Standard Shona Dictionary*. Revised Edition with Addendum. Harare: College Press.
- Hartmann, A.M. 1894. *English-Mashona Dictionary*. Cape Town: Juta.
- Kennedy, G. 1998. *An Introduction to Corpus Linguistics*. London/New York: Longman.

- Kipfer, B.A.** 1984. *Workbook on Lexicography: A Course for Dictionary Users with a Glossary of English Lexicographical Terms*. Exeter: University of Exeter.
- Magwa, W.** 1999. *Manyorerwo eChiShona*. Gweru: Mambo Press.
- Mheta, G.** 2005. *Duramazwi reMimhanzi*. Gweru: Mambo Press.
- Mheta, G.** 2007. *Strategies Employed in the Development of Shona Musical Terms*. Unpublished M.Phil. Thesis. Harare: University of Zimbabwe.
- Mpofu, N.** 2004. *Duramazwi reUrapi neUtano*. Gweru: Mambo Press.
- National Language Policy Advisory Panel.** 1998. *Report on the Formulation of a National Language Policy*. Harare: Ministry of Education, Sport and Culture.
- Sinclair, J.M.** 1992. The Automatic Analysis of Corpora. Svartvik, J. (Ed.). 1992. *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82, Stockholm, August 4–8, 1991*. Berlin: Mouton de Gruyter.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- ZIMSEC/CDU.** 2006. *Duramazwi reZvidzidzo zveChiShona*. Harare: ZIMSEC.

Websites

The African Languages Research Institute. Available at: <http://www.uz.ac.zw/alri/>.

The Treatment of Encyclopedic Information in *Duramazwi reMimhanzi*

Gift Mheta, *Department of Linguistics, University of the Western Cape, Bellville, Republic of South Africa (2971565@uwc.ac.)*

Abstract: This article presents the problems and challenges of dealing with encyclopedic information in Shona lexicography. According to Zgusta (1971: 198), encyclopedic dictionaries are primarily concerned with the denotata of the lexical units. They give information about the extralinguistic world, physical or non-physical. In other words, encyclopedic dictionaries contain extensive information on lexical entries. The article focuses on how such encyclopedic information is collected, selected, entered and treated. Emphasis is on identifying the problems and challenges experienced by lexicographers at the four main stages of dictionary making. It also presents solutions to the encountered difficulties. It brings out discernible patterns in the treatment of encyclopedic information. The analysis, from the collection stage up to the treatment stage, is done with reference to the relatively long history of compiling linguistic dictionaries in Shona.

Keywords: ENCYCLOPEDIA, COLLECTION, TREATMENT, ENTRY, ETYMOLOGY, ILLUSTRATION, LEXICOGRAPHY, LINGUISTIC DICTIONARY, MUSIC, SELECTION

Opsomming: Die behandeling van ensiklopediese inligting in die samestelling van *Duramazwi reMimhanzi*. Hierdie artikel behandel die probleme en uitdagings by die aanbieding van ensiklopediese inligting in Sjonaleksikografie. Volgens Zgusta (1971: 198) is ensiklopediese woordeboeke primêr gemoeid met die denotata van die leksikale eenhede. Hulle verskaf inligting oor die ekstralinguistiese wêreld, fisies en niefisies. Met ander woorde, ensiklopediese woordeboeke bevat uitvoerige inligting oor leksikale inskrywings. Die artikel fokus op hoe sulke ensiklopediese inligting versamel, gekeur, aangebied en behandel word. Klem is op die identifisering van probleme en uitdagings wat deur leksikograwe ondervind word by die vier hoofstadiums van woordeboekmaak. Dit bied ook oplossings vir die teëgekome probleme. Dit bring onderskeibare patrone in die behandeling van ensiklopediese inligting na vore. Die ontleding, van die versamelingstadium tot by die behandelingstadium, word gedoen met verwysing na die relatiewe lang geskiedenis van die samestelling van linguistiese woordeboeke in Sjona.

Sleutelwoorde: ENSIKLOPEDIE, VERSAMELING, BEHANDELING, INSKRYWING, ETIMOLOGIE, ILLUSTRASIE, LEKSIKOGRAFIE, LINGUISTIESE WOORDEBOEK, MUSIEK, KEUSE

Introduction

After the publication of *Duramazwi reChiShona* (1996) and *Duramazwi Guru re-ChiShona* (2001), there has been a paradigm shift in Shona lexicography, from

the production of general linguistic dictionaries to the compilation of specialized dictionaries as evidenced by the publication of the first two of these, *Duramazwi reUrapi NeUtano* (2004) and *Duramazwi reMimhanzi* (2005).

This article highlights the experiences and challenges faced by the African Languages Research Institute (ALRI) in the compilation of an encyclopedic music dictionary, *Duramazwi reMimhanzi* (henceforth the DRM). The article is based on experiences and observations as editor-in-chief of the DRM. The dictionary, published in 2005, is the first of its kind in the lexicographic history of Zimbabwe. The DRM is trilingual, presenting three languages, namely Shona, Ndebele and English, the official languages in Zimbabwe. However, the dictionary has one alphabetizing language, i.e. Shona, spoken by 75 percent of the country's population of 11.6 million people. In this dictionary, Shona headwords are followed by equivalent terms in Ndebele and English and then by definitions, all in the Shona language. At the end of the Aa–Zz section of the dictionary, there is a section with illustrations followed by two indices, the one comprising Ndebele and English equivalent terms and the other made up of Ndebele headwords and Shona glosses. The dictionary is targeted at tertiary institutions and some secondary schools offering music as a course or subject.

Collection Stage

Challenges of compiling the DRM were encountered at the initial collection stage. The main problem was how to collect terms for such a specialized dictionary when there were no other Shona musical terms dictionaries. The DRM editorial team ended up relying on general linguistic dictionaries such as *Duramazwi reChiShona* and *Duramazwi Guru reChiShona*. However, the editorial team only managed to collect about 700 Shona musical terms. With a target of between 1 000 and 1 500 musical terms, ways of reaching the set target had to be explored. The DRM editorial team resolved to work with different experts in the music field who included music lecturers and teachers, musicians and performers. Despite contributing greatly to the compilation of the DRM, the reference working groups (RWGs) presented a number of problems to the lexicographic team. The main problem was that the groups had to consist of music lecturers and teachers together with linguists from all the provinces in Zimbabwe so that whatever the editorial team together with the RWGs agreed upon, especially newly created terms, would be accepted nationwide.

Assembling RWGs required an initial survey aimed at identifying personnel with relevant credentials throughout Zimbabwe. This survey was carried out during 2002, establishing contacts nationwide. It required much travelling which had serious financial implications. A substantial amount was required for planning and carrying out the trips in various parts of the country. At least three experts were identified in every province. The RWGs worked closely with the DRM team from the collection right up to the definition stage. However, the other problem was that the RWGs were not as big as desired since

members of these groups had to be paid for their professional services.

Notwithstanding the challenges encountered in working with these experts, the team managed to collect and define the 1 200 musical terms found in the DRM today.

Selection Stage

Selection of indigenous musical terms was not a problem because there are not so many of them. As already mentioned above, musical terms collected from Shona dictionaries constituted less than half of the set target. Selection problems only appeared when dealing with Western musical terms to be included in the dictionary. The problem lay in choosing from the abundance of Western musical terms. The *New Grove Dictionary of Music and Musicians*, for instance, consists of twenty volumes, each having approximately 900 pages. With the upper limit of the target for music headwords fixed at 1 500, the selection process had inevitably to be strict. The team ended up formulating stringent selection guide-lines that would yield the most salient music headwords from Western music literature. All elements of music, for instance, had to be incorporated into the dictionary since they constitute the core of the music discipline. In this manner terms such as *beat, harmony, melody, pitch, pulse, tempo* and *rhythm* were prioritized over other musical terms dealing with musical concepts alien to Zimbabwean music culture.

Entry Stage

With the experience of compiling linguistic dictionaries, the editorial team at ALRI was faced with the challenge of selecting what to enter into the database. It agreed to enter no grammatical information other than tone and etymology, which add to encyclopedic information. An encyclopedic dictionary offers a more complete description of entries than a linguistic dictionary. In order to have comprehensive information about entries, the editorial team had to enter etymological information for newly created musical terms especially those concepts borrowed from Western music. The histories of the newly adopted terms were included to complement the treatments. The target users are likely to be informed of the meanings of new terms even before reading the treatments. An example demonstrating how etymological information is included is the following (Mheta 2005: 7-8):

bhesi [bhasi] KD *ibhesi* (ND) *bass* (EN). <ChiRatini *bassus* 'kudzama/kudzika'>. *Bhesi* izwi gobvu rakadzikira repasi pamazwi mana anowanzoshandiswa pakuimba. Izwi rerudzi urwu rinowanzoshandiswa pakuimba nevanhurume asi kune vamwe vanhukadzi vanorikwanisa. Kune zviridzwa zvakasiyanasiyana zvinobudisa *bhesi*.

(**bhesi** [bhasi] HL **ibhesi** (ND) **bass** (EN). <Latin *bassus* 'deep'>. *Bass* is a voice that is deep. It is the lowest when compared to the other three voice registers that are mainly used when singing. Bass is usually sung by men, but there are some women who can also sing it. There are numerous musical instruments that also produce *bass*.)

Treatment Stage

Defining of musical terms presented a number of challenges. The most difficult task was how to construct a framework for treating terms in a manner that would yield definitions with more encyclopedic information than those found in linguistic dictionaries. Since the dictionary is encyclopedic in nature, the editorial team had to frame treatment guidelines ensuring the inclusion of most of the encyclopedic information. Treatment guidelines for all major entries or concept categories such as ceremonies, costumes, dances, instruments, performers and songs were framed for use in the treatment of all headwords in the DRM. For instance:

Treatment of ceremonies had to consider:

- (a) whether the ceremony is traditional or modern as well as where it (especially the traditional) is performed,
- (b) the purpose or context of the ceremony,
- (c) the participants involved,
- (d) the costumes worn, and
- (e) the type of songs/instruments played.

Treatment of costumes had to consider:

- (a) whether the costume is traditional or modern,
- (b) the type of material used in making the costume,
- (c) the people wearing the costume and how it is worn,
- (d) the importance/significance of the costume, and
- (e) the ceremonies and geographical regions in which the costume is worn.

Treatment of dances had to consider:

- (a) whether the dance is traditional or modern,
- (b) the origin and region where it (especially the traditional) is found,
- (c) the performer(s) of the dance as well as the context of the dance and time when it is performed,
- (d) the costumes worn, equipment used or instruments played, and
- (e) the type of music involved.

Treatment of instruments had to consider:

- (a) the type of material used in the making of the instrument, e.g. reed, wood, clay, metal or fibre glass,

- (b) the parts which constitute the instrument,
- (c) the size of the instrument, e.g. small, medium, large,
- (d) the way of playing the instrument, e.g. plucked, shaken, beaten, stamped, scraped and blown,
- (e) the character of the sound produced by the instrument, e.g. sharp, pitched, heavy, light, piercing, etc.,
- (f) the classification system, i.e. Nketia's or Mararire's,
- (g) the period to which the instrument belongs, i.e. whether it is modern or traditional,
- (h) the source of origin of the instrument, and
- (i) the player of the instrument, the context in which it is played, and also its music and non-music functions.

With clearly spelt out definition formats it became easy for the editorial team to draft exhaustive treatments contrasting with the concise ones found in linguistic dictionaries. A comparison of the two treatments, one from the DRM and the other from *Duramazwi Guru reChiShona* clearly illustrates this point.

Treatment from the DRM (Mheta 2005: 87):

nyungwenyungwe [nyunganyunga] KDKD inyungwenyungwe (ND) mbira (EN).
Nyungwenyungwe nderumwe rudzi rwembira rune makiyi anokwana gumi neshanu. Inoridzwa kazhinji muzvikoro nemakoreji anodzidzisa nezvemimhanzi. Mukuiridza muridzi anoshandisa *gamburabota* kuridza makiyi epasi okurudyi kwake uye *gamburabota* rekuruboshwe kuridza makiyi epasi okuruboshwe kwake. Anoshandisawo munongedzo wekurudyi kuridza makiyi epamusoro ese. *Imbira* inokwanisa kuridzwa chero naani zvake nguva ipi zvayo. Yakabva kudunhu reNyungwe kuMozambiki naJeke Taperu uyo akauya nayo muZimbabwe kuma1960 akatanga kudzidzisa maridzirwe ayo paKwanongoma College of Music yava kunzi United College of Education kuBulawayo. Inowanikwawo zvakare kuAmerika uko yakaendeswa nemushakabvu Dumisani Abraham Maraire aiiridza nekuidzidzisa ikoko. FAN *karimba*.

nyungwenyungwe [nyunganyunga] HLHL inyungwenyungwe (ND) mbira (EN).
Nyungwenyungwe is a type of *mbira* that has 15 keys. It is usually played in schools and tertiary institutions that teach music. When playing it, the instrumentalist uses the left thumb to pluck the bottom placed keys on his/her right side and uses the right index finger to pluck the upper keys. It is a type of *mbira* that can be played by anyone at any given time. The instrument originated from the Nyungwe District in Mozambique where it was discovered by Jeke Taperu who brought it to Zimbabwe in the 1960s and started teaching students at Kwanongoma College of Music, now known as the United College of Education, in Bulawayo how to play it. *Nyungwenyungwe* is also found in the United States where it was introduced at various universities by the late musician and ethnomusicologist Dumisani Abraham Maraire. SYN *karimba*.

Treatment from *Duramazwi Guru reChiShona* (Chimhundu 2001: 772):

nyunganyunga KDKD z 9>10. *Nyunganyunga* nderumwe rudzi rwembira rwunoridzwa. FAN *karimba* 12.

(**nyunganyunga** HLHL n 9>10 *Nyunganyunga* is a type of mbira that is played. SYN *karimba* 12.)

The abundance of information on musical terms shows that the DRM editorial team successfully framed treatment guidelines which have helped in the tapping of as much encyclopedic information as possible. Having dealt with the problem of how to define terms in an encyclopedic dictionary, the editorial team was left with the challenge of how to complement the definitions in the dictionary.

Illustrations

One of the decisions made by the DRM editorial team was to include illustrations between the dictionary and indices sections. The illustrations mainly depict musical instruments, costumes and music notational symbols. It was, however, difficult to make a choice on the type of illustrations to be included. Photographs would have been preferred since they are user-friendly, fully capturing most features of the items to be illustrated. They enhance a dictionary. However, good quality paper which is essential for the inclusion of photographs is very costly. This is the main reason why the DRM team opted for line drawings. The only problem is that line drawings are not as easily accessible as photographs. They require users to have prior knowledge of the concepts being illustrated. Research shows that they can easily be misunderstood by users unfamiliar with certain drawing techniques and conventions such as cross-sections.

Conclusion

Compiling encyclopedic dictionaries is not an easy task as demonstrated by the numerous challenges met by the ALRI lexicographic team in the compilation of the DRM. Such challenges were only solved by close co-operation with experts in the music field. Working closely with experts in specialized fields is the key to success in the production of encyclopedic dictionaries.

Bibliography

- Chimhundu, H.** 1996. *Duramazwi reChiShona*. Harare: College Press.
Chimhundu, H. 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
Mheta, G. 2005. *Duramazwi reMimhanzi*. Gweru: Mambo Press.
Mpofu, N. 2004. *Duramazwi reUrapu neUtano*. Gweru: Mambo Press.
Zgusta, L. 1971. *Manual of Lexicography*. Prague: Academia / The Hague/Paris: Mouton.

The Use of Loan Translation as a Term-Creation Strategy in *Duramazwi reMimhanzi*

Gift Mheta, *Department of Linguistics, University of the Western Cape,
Bellville, Republic of South Africa (2971565@uwc.ac.za)*

and

Itai Muhwati, *Department of African Languages and Literature, University of
Zimbabwe, Harare, Zimbabwe (itaimuhwati@yahoo.co.uk)*

Abstract: The article presents loan translation as one of the key strategies that lexicographers at the African Languages Research Institute (ALRI) employed in the creation of new terminology during the compilation of the music dictionary *Duramazwi reMimhanzi*. Inspired by the authors' keen interest in lexicography and active involvement in the compilation of *Duramazwi reMimhanzi*, it discusses the advantages and disadvantages of loan translation. Analysis of this term-creation strategy is done using Chimhundu's (1996) translation and terminology development theory known as the scan and balance theory.

Keywords: COINAGE, CULTURE, DECODING, ENCODING, STRATEGY, TERM-CREATION, TRANSLATION, LOAN TRANSLATION

Opsomming: Die gebruik van leenvertaling as 'n termskeppingstrategie in *Duramazwi reMimhanzi*. Die artikel bied leenvertaling aan as een van die sleutelstrategieë wat leksikograwe by die African Languages Research Institute (ALRI) gebruik het vir die skepping van nuwe terminologie gedurende die samestelling van die musiekwoordeboek *Duramazwi reMimhanzi*. Geïnspireer deur die skrywers se sterk belangstelling in die leksikografie en aktiewe betrokkenheid by die samestelling van *Duramazwi reMimhanzi*, bespreek dit die voordele en nadele van leenvertaling. Ontleding van hierdie termskeppingstrategie word gedoen deur gebruik van Chimhundu (1996) se vertalings- en terminologieontwikkelingsteorie bekend as die skandeer- en balanseerteorie.

Sleutelwoorde: VINDING, KULTUUR, DEKODERING, ENKODERING, STRATEGIE, TERMSKEPPING, VERTALING, LEENVERTALING

Introduction

One strategy employed in the creation of musical terms in *Duramazwi reMimhanzi* is loan translation, calquing, paraphrasing or circumlocution and the resultant terms are called loan translations, calques, paraphrases or circumlocutions. In loan translation, according to Crystal (1997), the morphemes in the borrowed word are translated literally or item by item. Similarly, music con-

cepts in English were translated into Shona and the resultant loan translations denote the same musical concepts. Examples of such loan translations are:

- (1) lyrics > *mazwi erwiyo* (literally: *words of a song*);
- (2) genre > *mhando yemumhanzi* (literally: *type of music*);
- (3) ensemble > *mubatanidzwa wevaimbi nezviridzwa* (literally: *combination of musicians and their instruments*);
- (4) composer > *munyori werwiyo* (literally: *writer of a song*);
- (5) piano body > *mutumbi wepiyano* (literally: *body of a piano*);
- (6) minor scale > *neramhanzi diki* (literally: *small music ladder*);
- (7) major scale > *neramhanzi guru* (literally: *big music ladder*);
- (8) maestro > *nyanzvi yemimhanzi* (literally: *expert in music*);
- (9) instrumental music > *nziyo dzemagitare* (literally: *songs with guitars*);
- (10) staff notation > *runyoro rwezvitsvimbo* (literally: *writing of knobkerries*);
- (11) national anthem > *rwiyo rwenyika* (literally: *song for the country*);
- (12) music appreciation > *yemuro yemumhanzi* (literally: *appreciation of music*).

Creation of the given loan translations is not a random but rather a systematic process which involves understanding both the English and Shona languages and cultures. The process of creating the above loan translations can be represented by the rule notations that follow.

The basic rule is that the term-creator should understand his role as that of constantly moving in and out of the source language (SL) and target language (TL) to find or create equivalent terms and then match or balance these in terms of meaning and form.

To arrive at each of the loan translations from (1) to (12) given above, the following procedure should be followed. The term-creator is faced with the lexical item *lyrics*, which does not have an equivalent term in Shona. The term-creator constantly scans, for example, the English language and culture as he searches or attempts to create the Shona equivalent for *lyrics*. The term-creator also tries to strike a balance between meaning and form throughout the term-creating process. The two major processes, scanning and balancing, are done in six main stages of terminology creation. How the output *mazwi erwiyo*, for instance, is finally reached, can now be considered. The decoding process can be diagrammatically represented as follows:

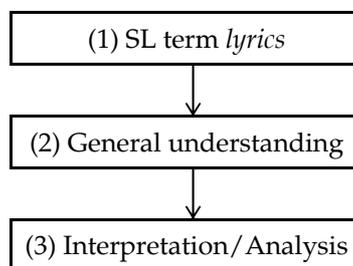


Figure 1: An Example of the Decoding Process

In this case, the term-creator is presented with the SL text *lyrics* at stage (1). At stage (2), he searches for the general meaning of the term *lyrics* which he decodes as the text of a song. However, as the term-creator continues searching for the meaning of the SL text, he begins to register more information about the SL text at stage (3). At this stage the term-creator decodes different forms of lyrics, the importance of lyrics in English music culture and how the lyrics are transcribed and performed. In essence, he searches and decodes everything related to the term *lyrics*. It should be noted that the most crucial analysis of the term *lyrics* is done at stage (3), in the context of English culture.

Stages (4) and (5) which constitute the encoding process taking place in the TL culture can be diagrammatically represented as follows:

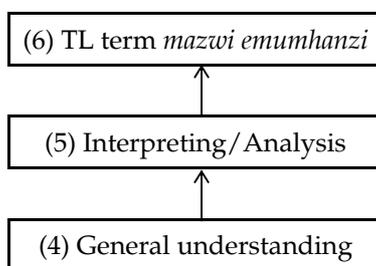


Figure 2: An Example of the Encoding Process

At stage (4), the term-creator may have a general idea of the options existing in the Shona language for the English term *lyrics*. The options include both Shona equivalent terms that might exist for the term *lyrics* and the strategies that can be used in the recreation of the SL text in the TL text. In this case, the term-creator searches for the Shona lexical equivalent for the term *lyrics* but does not find it because it does not exist in the Shona lexicon. There being no equivalent term, the term-creator proceeds to stage (5) where he recreates the concept of *lyrics* through a loan translation, *mazwi emumhanzi*, which is the output that appears at stage (6). The meaning intended by the lexical item *lyrics* at stage (1) is the same meaning signified by the output *mazwi emumhanzi* at stage (6). In other words, there is matching at the semantic level. However, the form is changed in order to achieve the intended SL meaning. What is presented as a one-word form at stage (1), becomes a two-word form at stage (6). The compromising of form is a balancing act intended to achieve the SL meaning in the TL. The same process takes place when the English term *maestro* yields through loan translation the output *nyanzvi yemimhanzi*. What is a single noun in the SL (specifically English) becomes a noun phrase consisting of *nyanzvi*, which is qualified by another noun made up of the possessive inflection *ye-* prefixed to the noun *mimhanzi*.

From the preceding discussion, it becomes clear that the encoding process involves literal translation of the sense of the SL text into the TL. The term *lyr-*

ics, for instance, is rendered into Shona by describing what it is, hence the resultant loan translation *mazwi emumhanzi*. Similarly, the English term *maestro* is translated into Shona on the basis of function, that is, through decoding and encoding what a *maestro* does and how he does it. As stated earlier, the resultant loan translation for *maestro* is *nyanzvi yemimhanzi*, which, when back-translated into English, means "one who is very good at music". The back translation proves that the Shona loan translation fully represents the original meaning of the SL text *maestro*. The Shona loan translation perfectly captures the trade of a *maestro*, which is music. It also brings out the qualities of a *maestro*, namely a person who is really good at music.

From the above discussion, it can be concluded that the term-creator should ensure that the sense carried by the SL term is retained as much as possible in the TL. There might be slight adjustments at the level of form to ensure effective rendition of the SL meaning in the TL. The inevitability of such adjustments is also noted by Gentzler (1993: 87) who states:

Every conception of translation of any real significance and consistency finds its principal manifestation in the shifts of expression, the choice of aesthetic means, and the semantic aspects of the work. Thus in a translation we can, as a rule, expect certain changes because the question of identity and difference in relation to the original can never be solved without some residue.

Slight changes at the level of form are only permissible if they ensure the effective rendition of the intended SL meaning in the TL. The discussed examples have demonstrated that if a balance is struck between form and meaning, the resultant loan translation is usually an acceptable one. What is important is that the loan translation should carry "the invariant core" (Bassnett-McGuire 1980: 26-27) or "the relevant substance for the two texts" (Catford 1965: 25). In other words, the loan translation should carry what Nida (1982) calls the functional or dynamic equivalence. Form is, therefore, sacrificed for expressive expediency.

However, the discussed examples should not create the misleading impression that all loan translations are effective. There are cases of ineffective loan translations involving Shona musical terms which fail to capture the original meaning of the SL terms. A well-known Shona musical term that other cultures have tried to translate using the loan translation strategy is *mbira*. Attempts were made to translate *mbira* into English. The resultant loan translation *hand piano* failed to gain currency, which could be attributed to the loss of too much information during the translation process. The operational efficiency or expressive power of the English loan *hand piano* falls short of the SL term *mbira*. If the scan and balance theory had been applied in the creation of this English loan translation, the term-creator who was presented with the Shona musical term *mbira* at stage (1) should have decoded its meaning from stage (2) to (3) in the SL (specifically Shona) culture. This decoding process can be diagrammatically represented as follows:

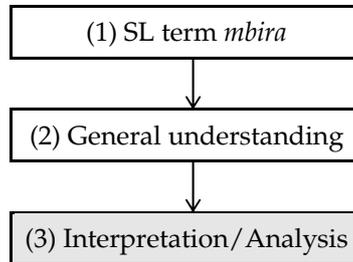


Figure 3: An Example of the Decoding Process

The first two stages are represented by two clear rectangles because they are free of errors in the decoding process. At stage (1) no error of judgement occurs when the term-creator is presented with the term *mbira*. Like the first, the second rectangle is also clear because the general decoding of the SL text, specifically *mbira*, is done correctly. At stage (2) the term-creator correctly registers the general knowledge that a *mbira* is a musical instrument. However, the last rectangle representing stage (3) is shaded to show that this is where the term-creator's errors might have occurred. The interpretation and analysis of the musical term *mbira* is ineffectively carried out in the SL because of the term-creator's initial failure to do a thorough decoding process. However, at stages (4) and (5), the term-creator encoded, producing the loan translation *hand piano* based on drawing parallels between the SL (Shona) and TL (English) cultures. The encoding process can be diagrammatically represented as follows:

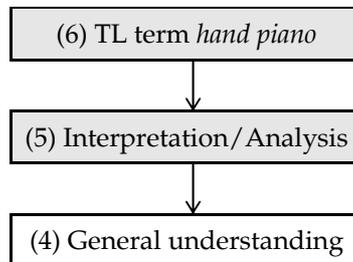


Figure 4: An Example of the Encoding Process

The English loan creation was probably produced after searching a musical instrument which looks like a *mbira*. At the general level, that is, at stage (2), it is true that the *mbira* is a musical instrument. However, the problem seems to have occurred at stage (3) when the term-creator failed successfully to complete the decoding process at the analysis or interpretation stage, by coming to the conclusion that a *mbira* functions like a piano. While there are in fact a few similarities especially in the way the keys for these two musical instruments are arranged, it should be emphasised that the differences between the two are many and as such far outweigh the presumed similarities. For instance, in con-

trast with piano keys, mbira keys are made of metal and even differ in terms of size, the manner in which they are played and, above all, the quality of the sound they produce. Another difference is that the body of a piano contains many parts not found on a mbira (the musical instrument), which consists of a "gwariva" (wooden plank), the "mbira" (metal keys), a "mutanda" (bridge), the "majeke" (rattles) and a "deze" (resonator) constituting the optional part of the instrument. Except for these differences between a piano and a mbira, the major difference is not that of appearance but of the diverse social functions these instruments have in the respective cultures from which they originate. On the one hand, the mbira is a revered musical instrument that holds a special position in traditional Shona culture. In some instances, it is even used as a "gokoro" (a musical instrument or anything used to evoke the spirit of the deceased). It plays a crucial role in facilitating ancestral spirit possession. The mbira is a musical instrument that is of great significance in Shona culture, used in most Shona traditional ceremonies such as the "kurova guva" (a family ritual for bringing back the spirit of the deceased). It is central to traditional ceremonies. The mbira is therefore seen in most circles as a symbol of traditional Shona culture. In terms of meaning *mbira* is not just an ordinary, but a culturally bound musical term that can only be translated into another language if the translator has full knowledge about its cultural significance. On the other hand, the piano does not have a similar symbolic value in English and other Western cultures.

The cultural differences between the two musical instruments, the mbira and the piano, are too great to be reconciled. Because of this, it is inappropriate to create a loan word based on a comparison of the one instrument with the other. The English word *hand piano* partially compares the physical structure of the mbira with a piano but fails to capture the cultural significance of the instrument for the Shona people. The two musical instruments are therefore incomparable at the level of culture. For the same reason, an attempt to translate *mbira* with *lamellaphone* also failed. The translations *hand piano* and *lamellaphone* never gained currency. It is the borrowed term *mbira* that is more popular than either of the coined English terms.

During the compilation of *Duramazwi reMimhanzi*, the lexicographers at ALRI rendered *mbira* as *mbira* in English. This decision was reached after noting that any attempt to translate it in any way would not be very effective because the mbira is quite unique in both structure and social significance. The lexicographers were quite aware that repeating the term in the TL would not make much sense to speaker-hearers of English who have not seen the mbira before. This was overruled, however, by the fact that the mbira is now very popular world-wide owing to Zimbabwean musicians who have been playing it for decades. The mbira has become such a popular instrument that it is even taught at some American universities, including the University of California. With such popularity, the retention of the term in the TL became justified. For the same reason, some Western musical terms were rendered through bor-

rowing. An example of such terms is *piano* itself, which is rendered in Shona as *piyano*. An attempt to create a Shona loan translation for *piano* was considered unnecessary because it was almost certain that such a translation would not gain currency against *piyano*, which is already an integrated loan word.

Conclusion

The article has identified loan translation as a productive term-creation strategy. It has shown that there is always need for term-creation in Shona lexicography. The task of creating new terminology is thus inescapable. It is mandatory that the lexicographer acquaints him-/herself with term-creation strategies such as loan translation. In addition, the article has demonstrated that the scan and balance theory, consisting of six stages, could assist the translator to find direct and accurate translation equivalents. The theory has been suggested as a better method of dealing with the lack of equivalents. However, it can only assist the translator if he/she has full information about the referent.

References

- Bassnett-McGuire, S.** 1980. *Translation Studies*. London: Routledge.
- Catford, J.C.** 1965. *A Linguistic Theory of Translation*. London: Oxford University Press.
- Chimhundu, H.** 1996. Translation in African LLDs and Communication for Development. *Proceedings of the XIV World Congress of the Fédération Internationale des Traducteurs (FIT)*: 448-59. Ausit: The Australian Institute of Interpreters and Translators.
- Crystal, D.** 1997. *A Dictionary of Linguistics and Phonetics*. Fourth Edition. Oxford: Basil Blackwell.
- Gentzler, E.** 1993. *Contemporary Translation Theories*. London: Routledge.
- Nida, E.A.** 1982. *The Theory and Practice of Translation*. Leyden: Brill.

Phonological Adaptation of Borrowed Terms in *Duramazwi reMimhanzi*

Gift Mheta, *Department of Linguistics, University of the Western Cape,
Bellville, Republic of South Africa (2971565@uwc.ac.za)*

and

William Zivenge, *African Languages Research Institute, University of
Zimbabwe, Harare, Zimbabwe (williamzivenge@yahoo.com)*

Abstract: This article analyses the phonological characteristics of Shona musical terms borrowed from English. It discusses the phonological processes that take place when words are borrowed directly or indirectly from English. Essentially, the article analyses the adoption and adaptation of Shona loan-words at phonological level. It draws examples from the dictionary of Shona musical terms *Duramazwi reMimhanzi* (2005). This exploration of loan-word adaptation enhances the understanding of the phonological changes that the musical terms undergo during the borrowing process.

Keywords: ADOPTION, ADAPTATION, EPENTHESIS, CONSONANT, LEXICOGRAPHY, MUSIC, PHONEME, PHONOLOGY, SUBSTITUTION, VOWEL

Opsomming: Fonologiese aanpassing van ontleende terme in *Duramazwi reMimhanzi*. Hierdie artikel ontleed die fonologiese kenmerke van Sjonamusieterme wat aan Engels ontleen is. Dit bespreek die fonologiese prosesse wat plaasvind wanneer woorde direk en indirek aan Engels ontleen word. Die artikel ontleed hoofsaaklik die oorname en aanpassing van Sjonaleenwoorde op fonologiese vlak. Dit neem voorbeelde uit die woordeboek van Sjonamusieterme *Duramazwi reMimhanzi* (2005). Die ondersoek na leenwoordaanpassing dra by tot die verstaan van die fonologiese veranderinge wat die musiekterme tydens die ontleningsproses ondergaan.

Sleutelwoorde: OORNAME, AANPASSING, EPENTESE, KONSONANT, LEKSIKOGRAFIE, MUSIEK, FONEEM, FONOLOGIE, VERVANGING, VOKAAL

Introduction

An analysis of data drawn from *Duramazwi reMimhanzi* shows that English words are modified phonologically during the process of borrowing. This modification is a result of the different phonological structures existing between Shona and English. The two languages have different phoneme invento-

ries, syllable structures and phonotactic constraints. Therefore, the rephonologization of English terms is necessary to make them fit into the phonological structure of Shona. One of the common phonological processes found in loanword adaptation is vowel substitution. To understand this process fully Shona and English vowels should be briefly described.

Shona has a simple five vowel system comprising /a, e, i, o, u/. Generally, these vowels occur in consonant vowel (CV) syllables. Shona is typically a CV language, but there are situations where some of these Shona vowels occur alone in a V syllable structure, for instance, 'who' /ani/ which is VCV in sequence. The five Shona vowels can be described as follows:

- /a/ low central vowel;
- /e/ mid front vowel;
- /i/ high front vowel;
- /o/ mid back vowel;
- /u/ high back vowel.

The situation in English is much more complex. English vowels are divided into two major types, monophthongs and diphthongs. According to their manner of articulation, monophthongs refer to vowels (pure vowels) where there are no detectable change in quality during a syllable (Crystal 1997). Diphthongs are vowels that exhibit a change in quality within a single syllable (O'Grady et al. 1987). Some examples of monophthongs are [ɪ], [e], [ɔ:], [ɒ], [æ], [ʌ] and [ɑ]. Examples of diphthongs are [aɪ], [eɪ], [eə], [ɪə], [ɔɪ], [ɔə], [ʊə], [aʊ] and [əʊ].

Another set of English complex vowels is called triphthongs. They refer to a type of vowel where there are two noticeable changes in quality during their production (Crystal 1997). Examples of English triphthongs are [aɪə] and [aɪə].

The next section looks at vowel substitution as one of the rephonologization processes occurring during the borrowing of English words.

Vowel Substitution

Substitution is a term used in linguistics to refer to the process or result of replacing one item by another at a particular place in a structure (Crystal 1997). Therefore, vowel substitution is the process whereby English vowels are replaced by Shona vowels during the borrowing process (Chimhundu 2002, Zivenge 2006). This process occurs to replace English vowels that do not exist in Shona. The data below show how English vowels were substituted by Shona vowels during the borrowing process and how English monophthongs are realised in Shona as a result of rephonologization. Table 1 presents English vowels and their substitutes in Shona.

Table 1: Vowel Substitution

English Form	English Vowel(s)	Shona Rephono- logized Form	Shona Vowel Substitute(s)
/bænd/	/æ/	[bendi]	[e]
/bə'su:n/	/ə/ and /u/	[basuni]	[a] and [u]
/'kʌntri/	/ʌ/ and /ɪ/	[kandiri]	[a] and [i]
/'ɔ:gən/	/ɔ/ and /ə/	[ogani]	[o] and [a]
/ɒpərə/	/ɒ/ and /ə/	[opera]	[o] and [a]
/tju:bə/	/u/ and /ə/	[tjubə]	[u] and [a]
/'rɪ'kɔ:dəʃ/	/ɪ/, /ɔ/ and /ə/	[rekoda]	[e], [o] and [a]

Table 1 shows that the English vowels, /ʌ/ and /ə/, are both substituted by the Shona vowel [a], which has distinctive features close to these two English vowels. They share phonetic features such as [- round], [- back] and [- front], but, in contradistinction, the Shona [a] is [+ low]. It is noteworthy that [a] is the only vowel in Shona that is [+ low]. As such, the words 'country' /'kʌntri/ and 'tuba' /tju:bə/ are realised in Shona as [kandiri] and [tjubə] respectively. Similarly, /ɒ/ as in 'opera' /ɒpərə/ and /ɔ/ as in 'organ' /'ɔ:gən/ are realized in Shona as [opera] and [ogani] respectively. The reason for this disparity is that there is no one-to-one match between English and Shona vowels. English has twenty-five vowels including pure vowels and diphthongs and Shona has five (Zivenge 2006). English vowels /æ/, /ɪ/, /ɔ/ and /u/ are substituted by [e], [i], [o] and [u] in words such as 'band' /bænd/, 'violin' [vajorini], 'recorder' /'rɪ'kɔ:dəʃ/ and 'bassoon' /bə'su:n/ respectively. In all cases, the same principle of substitution takes place by using the vowels with the closest distinctive features.

This section has shown that all the English vowels not found in Shona have been replaced by Shona equivalents. The following subsection looks at vowel insertion as one of the phonological processes used to make English words fit into the phonological structure of Shona.

Vowel Epenthesis / Insertion

Epenthesis refers to a type of intrusion where an extra sound has been inserted in a word (Crystal 1997). There are two major reasons why vowels are inserted during the rephonologization process of English words into Shona. These include epenthesis to break up consonant clusters and epenthesis to open syllables. The reason for this is that Shona has no closed syllables.

Vowels in English musical terms borrowed by Shona were epenthesized in order to make them conform to the syllable structure requirements of Shona phonology. English naturally has closed syllables, a characteristic that is not shared by Shona. When exposed to the Shona phonological environment which

has closed syllables, vowel epenthesis had to be applied to English musical terms in order to produce a phonologically acceptable output.

The next section analyses the insertion performed to break up English consonant clusters and create open syllables.

Word Medial Vowel Epenthesis

This type of epenthesis involved the insertion of vowels into the middle of words. It is also known as mid-word vowel epenthesis. It was carried out to break the cluster consonants not acceptable in Shona. Table 2 shows the words in which the vowels were epenthesized to break clusters not permissible in Shona.

Table 2: The Use of Word Medial Vowel Epenthesis in Breaking Consonant Clusters

English Form	Shona Rephonologized Form	Epenthetic Vowel(s)
/əlegro/	[aregiro]	[i]
/flu:t/	[fureti]	[u]
/træns'fɔ:məʔ/	[tiranzifoma]	[i] + [i]

The data in Table 2 show that consonant clusters such as /gr/, /fl/, /tr/ and /ns/ are not allowed in Shona phonology. This can be observed in words such as [aregiro], [fureti] and [tiranzifoma] where the clusters are broken up. The data given below demonstrate the environments in which vowels are epenthesized in order to open closed syllables.

Word Final Vowel Epenthesis

Word final vowel epenthesis is vowel insertion at the end of the word. This process is also known as paragogic vowel epenthesis (Hock 1991). English words have closed syllables, while all Shona syllables are open. Therefore, word final vowel epenthesis was applied to open up the closed syllables. This is illustrated by the following examples:

Table 3: Word Final Vowel Epenthesis

English Form	Shona Rephonologized Form
/dɪsk/	[dɪsiki]
/'kɒnet/	[koneti]
/klef/	[kirefu]
/bæs drʌm/	[besi dɪramu]

The examples in Table 3 demonstrate word final vowel epenthesis. As the data show, the epenthesis vowels are [i] and [u]. The vowel [u] is followed by a labial fricative, for instance, /f/ in /klef/ that becomes [kirefu]. It can also follow a labial nasal, for instance, /m/ in /bæs drʌm/ 'bass drum', which becomes [bɛsi dɪramu] 'bhesi dhiramu' in Shona. An attempt to randomly epenthesize the vowels will yield unacceptable forms. Examples to this effect are [*bɛsi dɪramo], [*disiku] and [*kirefo] for the English forms /bæs drʌm/, /disk/ and [klef] respectively.¹ Therefore, word final vowel epenthesis is a systematic process regulated by the language that would have borrowed words from another, which in this case is Shona. As mentioned earlier, all vowel epenthesis processes are carried out to satisfy Shona phonotactic constraints, that is, to avoid consonant clusters that are not permissible in Shona. Another phonological process commonly used in adaptation of borrowed words is consonant voicing.

Consonant Voicing

Voicing is the change of a sound from a voiceless to a voiced state. Every language has phonological sequences that it allows or disallows, that is, phonotactic constraints. This section discusses the phonotactic constraints governing the sequencing of segments in words which result from the borrowing process. Table 4 exemplifies the phonotactic constraints which manifest themselves when borrowing takes place.

Table 4: Phonotactic Constraints in Shona

English Form	Shona Rephonologized Form
/æn'tenə/	[andena]
/pen'tætɒnɪk/	[pendatoniki]
/'tɪmpənɪ/	[timbani]
/tʃɑ:nt/	[tʃandi]
/'trʌmpət/	[tirambeti]

Table 4 shows that Shona does not allow prenasalization of voiceless obstruents² such as /*mp/ and /*nt/. These consonants are realized as voiced prenasalized stops, namely [mb] and [nd]. They are voiced, because in Shona all prenasalized consonants are voiced. Shona only allows voiced consonants to be prenasalized as in the examples [mb] and [nd] that are respectively realised as [timbani] and [tirambeti] on the one hand and [andena], [pendatoniki] and [tʃandi] on the other.

Glide Epenthesis

This is a phonological process which entails the insertion of two glides, namely

the palatal approximant [j] and the labiovelar glide [w]. Hock (1991) refers to these glides as semi-vowels. The insertion of these glides was a way of dealing with diphthongs and triphthongs in Shona.

Epenthesis of the Palatal Approximant [j]

The palatal approximant [j] was epenthesized in order to break diphthongs and triphthongs, which do not exist in the phonetic inventory of Central Shona. Table 5 shows the epenthesis of the palatal approximant /j/.

Table 5: Epenthesis of the Palatal [j] Approximant

English Form	Borrowed Form (Glide Insertion)
/æmplə'faɪəʔ/	[amburifaja]
/stɑɪl/	[tʃitajera]
/vɪ:'əʊlə/	[vijora]
/'wʌɪələs/	[wajiresi]

The examples in Table 5 capture the epenthesis of the palatal approximant [j] which was inserted to break the following /aɪ/ and /əʊ/ in the English words 'style' /stɑɪl/ and 'viola' /vɪ:'əʊlə/ respectively. It was also inserted to break the triphthongs /aɪə/ and /aɪə/ in the English words 'amplifier' /æmplə'faɪəʔ/ and 'wireless' /'wʌɪələs/ respectively. The resultant forms of the insertion of the palatal approximant [j] to break both diphthongs and triphthongs are [tʃitajera] 'chitayera', [vijora] 'vhiyora', [amburifaja] 'amburifaya' and [wajiresi] 'wairesi' respectively.

Epenthesis of the Labiovelar Glide [w]

The labiovelar glide [w] was also epenthesized to break diphthongs in the same manner as the palatal approximant [j]. It should however be noted that the epenthetic glide [w] appeared in a phonological environment different from that of its counterpart [j]. Table 6 shows the phonological environments in which the epenthetic glide [w] appeared.

Table 6: Epenthesis of the Labiovelar Glide [w]

English Form	Shona Rephonologized Form
/kaʊntə'pɔɪnt/	[kawundapojindi] 'counterpoint'
/saʊnd/	[sawundi] 'sound'
/saʊndtræk/	[sawunditireki] 'soundtrack'
/saʊndpru:f/	[sawundipurufu] 'soundproof'
/səʊl/	[sowuru] 'soul' (music type)

The table shows that the insertion of the labiovelar glide [w] was carried out to break the diphthongs /aʊ/ and /əʊ/ in the English words 'counterpoint' /kaʊntəˈpɔɪnt/, 'sound' /saʊnd/, 'soundtrack' /saʊndtræk/, 'soundproof' /saʊndpruːf/ and 'soul' /səʊl/. The resultant forms of the epenthetic labiovelar glide [w] are [kawundapojindi], [sawundi], [sawunditireki], [sawundipurufu] and [sowuru] respectively. This epenthesis was carried out to break the diphthong /aʊ/ and create another syllable [wʊ].

From the given examples, it is clear that there is a pattern of how the labiovelar glide was epenthesized in the rephonologized Shona musical terms during the borrowing process. It was epenthesized when preceding the back vowel [u].³ The glide [w] was epenthesized when the vowel [u] was preceding it, because the two have similar distinctive features which can be presented as follows:

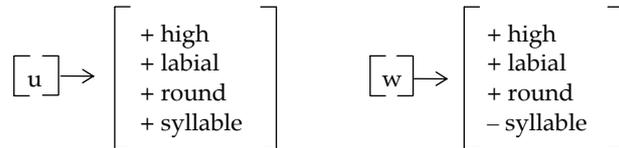


Figure 1: Distinctive Features of [u] and [w]

Figure 1 shows the similar distinctive features of the labiovelar [w] and the vowel [u] enabled epenthesis of [w] to be effected, as in 'sound' /saʊnd/ which becomes [sawundi]. The only notable difference between the two segments is that the labiovelar [w] is [-syllabic], while [u] is [+syllabic]. Glide epenthesis is therefore systematic in Shona. As shown, the insertion of the glide depends on the phonological environment.

Consonant Substitution

Consonants were substituted in the same manner as vowels. See the section on vowel substitution. Substitution was carried out to replace English consonants not existing in Shona. There are basically two English consonants not existing in the phonetic inventory of Central Shona. These are the lateral /l/ and the click /q/. The closest phonetic approximations of these two consonants are the trill or roll [r] and the velar [k]. Examples of Shona musical terms showing these consonant substitutions are [b̥araḍi] and [koteti], in which [r] and [k] substitute the English consonants /l/ and /q/, as in /ˈbæləd/ and /kwɔːˈtɛt/, respectively.

The given examples show that substitution is done systematically. It is only possible between consonants sharing a similar place of articulation. The English consonant /l/ is substituted by [r] in Shona because both consonants

are liquids. They are both apico-alveolar sounds (Crystal 1997: 23). In other words, the two consonants are articulated when the tip or end of the tongue hits the alveolar ridge. Similarly, the English consonant /q/ is substituted by /k/ because both are alveolar consonants.

This section has shown that there are certain phonological changes which take place when borrowed words are assimilated into Shona. In other words, occurring phonological changes serve to adapt the borrowed musical terms to the Shona language. In this process, English sounds are rendered using the closest phonological equivalents.

The examples discussed in this section are Shona musical terms borrowed from English and consequently adapted into Shona. Since English has a sound inventory different from that of Shona, the borrowed Shona musical terms had to be rephonologized to suit the phonological rules of Shona. The process of borrowing and the consequent adaptation of the Shona musical terms can diagrammatically be represented as follows:



Figure 2: Adoption and Adaptation of Borrowed Terms

Figure 2 summarizes the borrowing process. It shows that when terms are borrowed from another language, which in this case is English, they are adapted to the other language, which in this case is Shona. In other words, they are made to appear like original Shona words through processes such as phonological adaptation, which includes vowel substitution, vowel epenthesis, consonant voicing, glide epenthesis and consonant substitution.

Conclusion

This article has shown how borrowing is used as a method of creating new terminology in Shona lexicography. It has illustrated the phonological processes and rules lexicographers have to be acquainted with if they are to create new terms using the borrowing strategy. The importance of borrowing in the expansion of the Shona lexicon cannot be overemphasized. It is mandatory that lexicographers acquaint themselves with the phonological rules governing the languages in which they will be carrying out their lexicographic activities.

Endnotes

1. The asterisk marks a form that is unacceptable in Shona phonology.
2. A term used in the phonetic classification of speech sounds to refer to sounds involving a constriction which impedes the flow of air through nose or mouth, as in plosives, fricatives and affricates (Crystal 1997).
3. The authors are aware that epenthesis of the glide [w] is only possible in Shona when the vowel [u] is preceding or following it. However, the data collected for this article can only exemplify environments in which [u] is preceding [w].

References

- Chimhundu, H.** 2002. *Adoption and Adaptation in Shona*. Oslo: Allex Project.
- Crystal, D.** 1997. *A Dictionary of Linguistics and Phonetics*. Fourth Edition. Oxford: Basil Blackwell.
- Hock, H.H.** 1991. *Principles of Historical Linguistics*. New York: Mouton de Gruyter.
- Mheta, G. (Ed.)**. 2005. *Duramazwi reMimhanzi*. Gweru: Mambo Press.
- O'Grady, W. et al.** 1987. *Contemporary Linguistics: An Introduction*. London: Longman.
- Zivenge, W.** 2005. *An Analysis of Phonological Nativisation of English Loans in Tonga*. Unpublished M.A. Dissertation. Harare: University of Zimbabwe.

Affirming Verb Lexemes in *A Practical Ndebele Dictionary* and in *Isichazamazwi SesiNdebele*: The Case of **-wa** Verbs

Jerry Zondo, *Department of African Languages and Literature, University of Zimbabwe, Harare, Zimbabwe (jerryzondo@arts.uz.ac.zw)*

Abstract: Both *A Practical Ndebele Dictionary* (1971) by Pelling and *Isichazamazwi SesiNdebele* (2001) by Hadebe et al. lemmatize a number of **-wa**-ending verbs. Each one of the verbs is easily recognizable as a canonical or base form from which different word forms can be derived. Indeed they should all be seen as legitimate headwords that have been carefully selected for entry until a keener observation as presented in this article, questions the validity of the inclusion. It would be understood from both compilers that all the verbs entered are by virtue of their being legitimate verb bases. Since the verbs ending with **-wa** are canonical forms, they should be inflected with **-iw-/-w-**, the passive extension morpheme, and by any other extension morphemes. When **-iw-/-w-** cannot inflect directly, the usual initial inflection to causative or applicative and then to passive should then apply. When **-iw-/-w-** cannot apply, it is likely that the headword in question is in fact a passive form already derived from another base form. The article argues that these words cannot claim to be legitimate headwords. The article further observes that other verb endings treated can be similarly tested and treated accordingly. Possibly headword selection will then be more consistent and characterized by a more considered and reasoned approach.

Keywords: BASE FORM, CANONICAL FORM, DICTIONARY ENTRY, DERIVATION, HEADWORD, INFLECTION, PREDICTABILITY, SELECTION, STYLIZED VERBAL EXTENSION

Opsomming: Die bevestiging van werkwoordlekseme in *A Practical Ndebele Dictionary* en in *Isichazamazwi SesiNdebele*: Die geval van **-wa-werkwoorde**. Sowel *A Practical Ndebele Dictionary* (1971) deur Pelling en *Isichazamazwi SesiNdebele* (2001) deur Hadebe et al. lemmatiseer 'n aantal **-wa**-eindigende werkwoorde. Elkeen van die werkwoorde is maklik herkenbaar as 'n kanonieke of basisvorm waarvan verskillende woordvorme afgelei kan word. Hulle moet inderdaad almal gesien word as egte trefwoorde wat noukeurig gekies is vir opname totdat 'n skerper beskouing soos aangebied in hierdie artikel die geldigheid van insluiting bevraagteken. Dit word deur altwee samestellers te kenne gegee dat al die werkwoorde opgeneem is op grond daarvan dat hulle egte werkwoordbasse is. Aangesien werkwoorde wat op **-wa** eindig kanonieke vorme is, behoort hulle verbuig te word met **-iw-/-w-**, die passiewe uitbreidingsmorfeme, en deur enige ander uitbreidingsmorfeme. Wanneer **-iw-/-w-** nie direk kan verbuig nie, behoort die gewone beginverbuiging na kousatief of applikatief en dan na passief van toepassing te wees. Wanneer **-iw-/-w-** nie van toepassing kan wees nie, is dit waar-

skynlik dat die betrokke trefwoord in werklikheid 'n passiewe vorm is wat alreeds afgelei is van 'n ander basisvorm. Die artikel redeneer dat hierdie woorde nie aanspraak kan maak om egte trefwoorde te wees nie. Die artikel merk verder op dat ander werkwoorduitgange op dieselfde manier getoets en dienoreenkomstig behandel kan word. Moontlik sal die keuse van werkwoorde dan meer konsekwent wees en deur 'n meer deurdagte en beredeneerde benadering gekenmerk word.

Sleutelwoorde: BASISVORM, KANONIEKE VORM, WOORDEBOEKINSKRYWING, AFLEIDING, TREFWOORD, VERBUIGING, VOORSPELBAARHEID, KEURING, GESTILEERDE VERBALE UITBREIDING

1. Introduction

Two dictionaries in Ndebele that are widely used by students from secondary to university education and by teachers and lecturers in various institutions have been compiled by Pelling (1971) and Hadebe et al. (2001). Pelling's work is appropriately called *A Practical Ndebele Dictionary* (forthwith PND), because it is a bilingual dictionary most usable in practical situations of translation, especially by first and second year secondary students and by practitioners in language service institutions writing and translating tracts into Ndebele–English, English–Ndebele. The second dictionary with Hadebe as chief editor is called *Isichazamazwi SesiNdebele* (forthwith ISN), the first monolingual dictionary published in Ndebele meant for high school, college and university students. The Ndebele section of the African Languages Research Institute (ALRI) at the University of Zimbabwe, Harare, is planning an advanced Ndebele dictionary whose compilation will gain from reactions to and analyses of the PND and the ISN. The following article aims to make a contribution toward headword selection of verbs, attempting to be representative of any type of dictionary. It refers to the appropriateness of an entry, its legitimacy and the linguistic justification for its choice. The compilation of an advanced Ndebele dictionary will indeed be a more complex but more geared activity central around considered and reasoned approaches to dictionary making.

The selection of headwords for any dictionary is a key stage in dictionary compilation. It impacts on the user-friendliness of a dictionary and on its suitability to its various target users. Each lexicographer has her own considerations on selection but a target user will sometimes seek the selection criteria for certain canonical forms that a lexicographer has worked with in order to make sense of the dictionary under consideration. A more consistent and informed selection system is likely to appeal and be quite usable to the target user (cf. Chabata 1998: 141). This article seeks to make a contribution toward a scientific selection of entries for future dictionary work in Ndebele by making certain presumptions on the selection criteria for the PND and ISN. There is another Ndebele dictionary, *Isichazamazwi SezoMculo* that, because of its specialization in music terminology in Ndebele, does not apply to this investigation.

This article looks at the **-wa**-ending verbs in the PND and the ISN to establish if the two dictionaries have been consistent and accurate in using an appropriate canonical set in all cases. The principles of selection the lexicographers used are considered and evaluated. The aim is to establish a legitimate **-wa** verb entry and consequently, legitimate entries for all the other verbs which might be confused with a verb extension inflection. An overriding criterion is that a legitimate **-wa** entry should be inflected by **-iw/-w-**, a passive morpheme, to be a valid base form. If it cannot, then it would already be a passive form of an appropriate original base form and should be excluded from the headword list. Furthermore, the **-wa**-ending verb canonical forms should be freely inflected by other stylized extension morphemes.

Since the PND and the ISN are used in schools, colleges and universities, the basic assumption is that all the users have a sufficient knowledge and training to easily predict, or to predict with some consideration, the necessary and appropriate inflections and derivations of new words from the given base forms (cf. Chabata 1998: 142-143). The same users can easily do the same with **-wa**-ending verbs as bases. The research however reveals that compilers of both dictionaries have included forms that should be eliminated. Main entry selection depends on space, size, the frequency of the usage of a term, the specialization of the dictionary, a comprehensive and balanced coverage, the appropriateness of the vocabulary to the target group, legitimate entries or entry lexical validity, and a sense of what is wanted (Landau 2001: 356-359).

2. **-wa verb forms in PND and ISN**

Pelling has 14 entries of **-wa** verbs. The entire PND has 154 pages including the front and back matter, but the dictionary does not state the actual number of entries. Hadebe has 37 **-wa** verb entries in the ISN. Hadebe (2001: xxxiv) declares "zilutshwane impambosi zesenzo ezethuliweyo lapha" (there are only a few verb extension forms that have been entered here (as headwords)). The entire ISN has 556 pages including back matter. The ISN has 47 pages of front matter. The ISN fittingly has more information on how the dictionary has been compiled and is to be used. It provides more details on Ndebele grammar than does PND.

Pelling's entries are the following:

dakwa, v.i.: be drunk; **ehlelwa**, v.i.: experience; **felwa**, v.i.: be bereaved of; **fuywa**, v.i.: 1, be addicted to, 2, be domesticated; **gwegwa**, v.t.: hook, catch on a hook; **helelwa**, v.i. be in labour, have labour pains; **hlulukelwa**, v.i.: be sad; **khohlwa**, v.t.: forget; **kholwa**, v.t.: believe, trust, have faith; **khulelwa**, v.i.: become pregnant; **lwa**, v.t.: fight, contend; **selwa**, v.i.: 1, wane (of moon); 2, be late; **wa**, v.i.: fall, drop down; **zwa**, v.t.: 1, hear, understand; 2, feel; 3, taste.

The 37 entries of Hadebe are as follows:

aziwa (get known), **banjwa** (get caught), **bekwa** (get placed at), **bhajwa** (get stuck), **bulawa** (get killed), **buthwa** (get collected), **dakwa** (get drunk), **dinwa** (get tired), **dletshwa** (get scratched), **dliwa** (get eaten), **ehlelwa** (experience), **felwa** (have death in the family), **fuywa** (get addicted to), **gwegwa** (hook), **gqwagqwa** (get cooked), **gqwegqwa** (skirt an area), **helelwa** (feel birth pains), **hlelwa** (get organized), **hitshwa** (get choked), **hlulukelwa** (be sad), **hlwa** (turn to dusk), **jutshwa** (get impaled), **khohlwa** (forget), **kholwa** (believe), **khulelwa** (be overpowered), **lwa** (fight), **nethwa** (be drenched), **phiswa** (be pressed by urine), **qalekiswa** (be condemned by the divine), **selwa** (wane), **sikiswa** (accept a meat dish (be a bride for the first time)), **thwithwa** (drink noisily), **tshonelwa** (be delayed into evening), **vukelwa** (get possessed by an evil spirit), **vuthwa** (ripen), **zalelwa** (be born at), **zikhohlwa** (be extremely drunk).

The translations will in some cases lean towards the passive form. Pelling provides his own translations. As Hadebe's ISN is a monolingual dictionary, the glossing is by the author of this article.

3. Proof of the base form status

The selection criterion for most dictionary entries starts with a base form. A base "refers to any part of the word seen as a unit to which an operation can be applied, as when one adds an affix to a root or stem" (Crystal 1991: 35). The base form is sometimes called the lexeme, "the abstract unit which underlies some of the variants ... observed in connection with words" (Carter 1998: 7). The lexemes will underlie different grammatical variants which Carter (1998) calls word-forms. Lexemes will then be the basic contrasting units of vocabulary. Dictionaries enter lexemes rather than words. Words become predictable to language users who will, with assistance from sound linguistic training, derive the necessary words as and when needed in varied contexts. Invariably according to the general language, Pelling (1979: 19) will enter **-DAKWA** as a lexeme, and if he has provided sufficient front matter on verb extensions and the principle of inflection and derivation, the dictionary user will predict the word-forms **dakwela** (get drunk at/for); **dakwiswa** (be made drunk by); **dakwisisa** (be really drunk); in all cases of inflection, the root **-dakw-** is retained, and might appear at a glance to be adequate proof of the base form status of **dakwa**, but Pelling has a wrong base form here; the right one is **-dak-** with **dakela**, **dakiswa** and **dakisisa** as derived forms. Pelling's lexeme should be **-DAKA**. Hadebe (2001: 315) will enter **-LWA** as a lexeme but the user can once again derive its word-forms independently as **lwela** (fight for); **lwisa** (cause to fight), **lwiwa** (fought by); **lweka** (can be fought); **lwisisa** (fight extremely well); **zilwela** (fight for oneself). Luckily for both lexicographers, the two lexemes derive several word-forms each; each lexeme is therefore an ap-

appropriate base, a canonical form and therefore a suitable headword to choose for a general language entry. While the second is a good example of a **-wa**-ending verb, the first is not.

When **-wa**-ending verbs take on the stylized verb extension **-iw-/-w-**, the choice of this base form becomes even more easily and more compelling to the linguist. No further proof would then be needed to establish the status of the base form. The following verbs taken from both Pelling and Hadebe are inflected by the passive extension morpheme:

kholwa-iw-a > **kholwiwa** (be believed by)
gwegwa-iw-a > **gwegwiwa** (be hooked by)
zwa-iw-a > **zwiwa** (be heard by)
lwa-iw-a > **lwiwa** (be fought by)
thwithwa-iw-a > **thwithwiwa** (drunk noisily)
gqwegqwa-iw-a > **gqwegqwiwa** (be skirted)
khohlwa-iw-a > **khohlwiwa** (be forgotten)
gqwagqwa-iw-a > **gqwagqwiwa** (be partially cooked by)

The status of the set of verbs in the above examples as base forms leaves little doubt.

Many verbs in Ndebele will not inflect directly into the passive form. The passive inflection is then realized through another verb extension morpheme. The causative form easily takes the initial inflection on behalf of the passive. The following verbs illustrate this process:

mitha-is-a-w-a > **mithiswa** (+causative +passive: be impregnated by)
funda-is-a-w-a > **fundiswa** (+causative +passive: be taught by)
vika-is-a-w-a > **vikiswa** (+causative +passive: be caused to dodge something by)
vuka-is-a-w-a > **vukiswa** (+causative +passive: be made to awaken by)
sutha-is-a-w-a > **suthiswa** (+causative +passive: be caused to be satiated by)
lala-is-a-w-a > **laliswa** (+causative +passive: be made to sleep by)

It is likely that the six relevant verbs indicate a state, a process, or a condition. Their passivization will not be direct but will depend on an initial causative formation. A set of **-wa** verbs from PND and ISN behave in a similar manner to the six verbs above. The following verbs from Pelling and Hadebe illustrate the point:

bulawa-is-a-w-a > **bulawiswa** (+causative +passive: be caused to die by)
dakwa-is-a-w-a > **dakiswa** (be caused to be drunk by)
helelwa-is-a-w-a > **helelwiswa** (+causative +passive: caused to feel birth pains by)
hlulukelwa-is-a-w-a > **hlulukelwiswa** (+causative +passive: be saddened by)
hlwa-is-a-w-a > **hlwiswa** (+causative +passive: caused to be dark by)
phiswa-is-a-w-a > **phiswiswa** (+causative +passive: caused to be pressed by)
selwa-is-a-w-a > **selwiswa** (+causative +passive: caused to be late by)
vuthwa-is-a-w-a > **vuthwiswa** (+causative +passive: caused to ripen by)

The direct inflection by **-iw-/-w-** of eight of these **-wa** verbs affirms and confirms their categorical status as base forms. They have appropriately been selected as headwords for both dictionaries.

Khumalo (2007) refers to all verbs forming the basis for the discussion in this article as deponent verbs. He states that they are passive in form, but active in meaning. There are very few of these verbs in Ndebele. The two following are given by Canonici (1996):

Sengikhohliwe
Se-ngi-khohl-iw-e
ASP-1SG-forgetVR-PASS TENSE
'I have forgotten'

Uyalwa
U-ya-lw-a
SC-PRES/CONT-fightVRFV
'He is fighting'

Khumalo further points out some verbs in Ndebele that are used more extensively in the passive than in the active form. They will appear in dictionaries as appropriately lemmatized, not being so problematic as **-wa**-ending verbs. In fact, Khumalo (2007: 105-113) presents a summary of all verbal derivations which will be useful for future dictionary making. His examples and annotations cannot be ignored by any dictionary maker.

4. **-wa forms that are not really base**

The PND and the ISN have **-wa** verbs that actually do not belong to this category. The verbs can be eliminated by the test described in Section 3. It is easy to determine their derivation, making it unnecessary to enter them as base forms. They can then be considered as word forms from other appropriate bases. From Pelling's initial list of 14 verbs, three can easily be eliminated as word and not as base forms, using the following evidence:

dinwa < **dina-w-a**: inflection of base form **-din-** (tire)
ehlelwa < **ehlela-w-a**: inflection of base form **-ehlel-** (experience)
fuywa < **fuya-w-a**: inflection of base form **-fuy-** (rear)

Pelling will be left with only 11 genuine **-wa** ending verbs.

Hadebe with 37 entries can have his reduced to only 16. The following 21 of his verbs seem to fail the test:

aziwa < **azi-w-a**: inflection of base form **-az-** (know)
banjwa < **bamba-w-a**: inflection of base form **-bamb-** (hold, detain)
bekwa < **beka-w-a**: inflection of base form **-bek-** (place)
bhajwa < **bhaja-w-a**: inflection of base form **-bhaj-** (stuck)
buthwa < **butha-w-a**: inflection of base form **-buth-** (collect)

dinwa < **dina-w-a**: inflection of base form **-din-** (tire)
dletshwa < **dlepha-w-a**: inflection of base form **-dleph-** NB. Palatal change
dliwa < **dla-iw-a**: inflection of base form **-dl-** (consume)
ehlelwa < **ehlela-w-a**: inflection of base form **-ehlel-** (experience)
felwa < **f-el-w-a**: inflection of base form **-f-** (die at)
fuywa < **fuya-w-a**: inflection of base form **-fuy-** (rear)
jutshwa < **juba-w-a**: inflection of base form **-jub-** (gore)
hitshwa < **hitsha-w-a**: inflection of base form **-hitsh-** (choke)
hlelwa < **hlela-w-a**: inflection of base form **-hlel-** (classify)
khulelwa < **khulela-w-a**: inflection of base form **-khulel-** (over power)
nethwa < **netha-w-a**: inflection of base form **-neth-** (drench)
qalekiswa < **qaleka-isa-w-a**: inflection of base form **-qalek-** (faint)
sikiswa < **sika-is-a-w-a**: inflection of base form **-sik-** (cut)
tshonelwa < **tshona-el-a-w-a**: inflection of base form **-tshon-** (set)
vukelwa < **vuka-el-a-w-a**: inflection of base form **-vuk-** (possess)
zalelwa < **zala-el-a-w-a**: inflection of base form **-zal-** (give birth)

Only 16 genuine **-wa** verbs will remain for Hadebe, five more than those for Pelling. If all base forms shown as inflected here would have already been entered as headwords by Hadebe, then it would be unnecessary to add their passivized forms as entries.

It is clear that a thorough search in Ndebele for **-wa**-ending verbs will not yield more than about 25 in the entire language.

5. The redundancy rule in lexicography

There must be the use of the redundancy rule in lexicography that will direct headword selection. Obviously the stylized verb extension word forms would be redundant when dictionary users already know the principles of derivation. The verb headword will indicate the selection that the radical will make of extension forms after which the users will form the necessary inflections on their own. The strength of the lexicographic entry will be to show the selection as accurately as possible. Just as all plural forms of nouns should not be entered when the singular forms have been lemmatized, so all the verb extension forms should not be entered as base forms. Only copulative and adverbial formatives should be lemmatized, the front matter explaining the use and derivation with general linguistics augmenting and confirming dictionary usage. The PND and the ISN have not fully exploited the redundancy rule with respect to stylized verb extensions.

6. Standardization

What will be considered as a standard entry for **-wa**-ending verbs will obviously be the main entry that can be inflected into various extension forms including **-iw-/-w-**. The process of standardization in the **-wa** verb forms will

consist of consistently choosing as entries only such forms as can beyond doubt behave in the manner of the test established previously. Sager (1990: 115) calls, for example, for a standardization process that involves "the interest of economy ..., the interest of precision ..., the interest of appropriateness". The PND and the ISN show a limited application for each of the contentious verbs. Landau (2001: 357) calls for the selectivity principle, which has a telling effect on the size of any type of dictionary. When word lists for dictionaries are clearly spelt out and the principles of selection well laid out, then all entries in these dictionaries will be characterized by economy, precision and appropriateness. Subsequently, these dictionaries will adhere to the required standards. Hartmann (1983) indeed contends that "all dictionaries wittingly or unwittingly contribute to the codification of language use by their mere existence ..., seem to confer an uncontested authority upon those items listed". They can be undisputed authorities especially if an appropriate choice through adequate testing has been made. The main point is to ensure consistency by avoiding misinterpretations and then arriving at a uniform system of word selection. The standardization process will have an effect on future dictionaries if it is well employed in the compilation of the first or initial Ndebele dictionaries. The lexicographers will adhere firmly to the principles of an appropriate and scientific selection. Further, any present dictionary-making attempts for the Ndebele have inevitably to make a contribution toward the standardization of the Ndebele language in one way or another.

7. Conclusion

The PND and the ISN have entered **-wa**-ending verbs but have in the process confused **-iw/-w-** inflections for **-wa** verbs. Subsequently the selection of base forms for **-wa** verbs has been affected, resulting in false entries. The basic contention is that a more consistent selection of verbs will ensure that only the appropriate base forms for the verbs in question are entered. If well selected, the base forms can be inflected by **-iw/-w-** as well as by the other appropriate stylized extension forms. When a verb is already in the passive form, it cannot be entered as a base form since the passive verb would already be provided for in the original base form. Because the expected scientific and linguistic choice system has not been fully followed, the PND and the ISN have lost out on space and on accuracy by extending the scope of the selection. Proof of base form status is necessary for the right choice of verbs for headword status. The lexicographer has to consider whether stylized extension forms are redundant when entered as base forms. Dictionary users can use the predictability principle in establishing all the possible forms that can be derived from a base form. An accurate selection of **-wa** verbs will indeed help dictionary users in arriving at the right verb types and at advancing their linguistic knowledge. Since dictionaries are essential tools for the codification of a language, due responsibility for accurate and tested selection becomes central in lexicography. It will then

be easier for Ndebele dictionary users, and for language students in general, to pick out or select different types of verbs as base forms, especially if such verb forms end with syllables that might have been seen as stylized verb extensions. Ndebele dictionaries can then be used for several purposes including the effective standardization of the Ndebele language. Compilation of different types of dictionaries in Ndebele is bound to continue since only three Ndebele dictionaries are in use at present. Ndebeles frequently consult Zulu dictionaries and sometimes Xhosa dictionaries too. When further compilation of dictionaries continues, the compilers are likely to take a more positive consideration for the selection of entries which will possibly use a more concerted testing mechanism for the appropriate base forms. Further research is necessary on other verb endings that might on the surface be confused with the rest of the stylized verb radicals, such as **-el-**, the applicative form; **-is-**, the causative form; **-isis-**, the intensive form; **-an-**, the reciprocal form; and **-ek-** the neuter form. This means that there will be **-ela-**, **-isa-**, **-isisa-**, **-ana-**, **-eka-** ending verbs which should not be confused with any of the extensions discussed in this article.

References

- Canonici, N.N. 1996. *Zulu Grammatical Structure*. Third Edition. Durban: University of Natal.
- Carter, R. 1998. *Vocabulary: Applied Linguistic Perspective*. London: Allen and Unwin.
- Chabata, E. 1998. Using the Predictability Criterion for Selecting Extended Verbs for Shona Dictionaries. *Lexikos* 8: 140-153.
- Crystal, D. 1991. *A Dictionary of Linguistics and Phonetics*. Oxford: Basil Blackwell.
- Hadebe, S. et al. (Eds.). 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Hartmann, R.R.K. 1983. *Lexicography: Principles and Practices*. London: Academic Press.
- Khumalo, L. 2007. *An Analysis of the Ndebele Passive Construction*. Unpublished Ph.D. Dissertation. Oslo: University of Oslo.
- Landau, S.I. 2001. *Dictionaries. The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Nkomo, D. and N. Moyo (Eds.). 2006. *Isichazamazwi SezoMculo*. Gweru: Mambo Press.
- Pelling, J.N. 2001. *A Practical Ndebele Dictionary*. Second Edition. Harare: Longman Zimbabwe.
- Sager, J.C. 1990. *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.

VOORSKRIFTE AAN SKRYWERS

(Tree asseblief met die Buro van die WAT in verbinding (wat@sun.ac.za) vir 'n uitvoeriger weergawe van hierdie instruksies of besoek ons webblad: <http://www.wat.co.za>)

A. REDAKSIONELE BELEID

1. Aard en inhoud van artikels

Artikels kan handel oor die suiwer leksikografie of oor implikasies wat aanverwante terreine, bv. linguïstiek, algemene taalwetenskap, rekenaarwetenskap en bestuurskunde vir die leksikografie het.

Bydraes kan onder enigeen van die volgende rubrieke geklassifiseer word:

(1) **Artikels:** Grondige oorspronklike wetenskaplike navorsing wat gedoen en die resultate wat verkry is, of bestaande navorsingsresultate en ander feite wat op 'n oorspronklike wyse oorsigtelik, interpreterend, vergelykend of krities evalueerend aangebied word.

(2) **Resensieartikels:** Navorsingsartikels wat in die vorm van 'n kritiese resensie van een of meer gepubliseerde wetenskaplike bronne aangebied word.

Bydraes in kategorieë (1) en (2) word aan streng anonieme keuring deur onafhanklike akademiese vakgenote onderwerp ten einde die internasionale navorsingsgehalte daarvan te verseker.

(3) **Resensies:** 'n Ontleding en kritiese evaluering van gepubliseerde wetenskaplike bronne en produkte, soos boeke en rekenaarprogramme.

(4) **Projekte:** Besprekings van leksikografiese projekte.

(5) **Leksikonotas:** Enige artikel wat praktykgerigte inligting, voorstelle, probleme, vrae, kommentaar en oplossings betreffende die leksikografie bevat.

(6) **Leksikovaria:** Enigeen van 'n groot verskeidenheid artikels, aankondigings en nuusvrystellings van leksikografiese verenigings wat veral vir die praktiserende leksikograaf van waarde sal wees.

(7) **Verslae:** Verslae van konferensies en werksessies.

Bydraes in kategorieë (3)-(7) moet almal aan die eise van akademiese geskrifte voldoen en word met die oog hierop deur die redaksie gekeur.

2. Wetenskaplike standaard en keuringsprosedure

Lexikos is deur die Departement van Onderwys van die Suid-Afrikaanse Regering as 'n gesubsidieerde d.w.s. inkomstegenererende navorsingstydskrif goedgekeur.

Artikels sal op grond van die volgende aspekte beoordeel word: taal en styl; saaklikheid en verstaanbaarheid; probleemstelling, beredenering en gevolgtrekking; verwysing na die belangrikste en jongste literatuur; wesenlike bydrae tot die spesifieke vakgebied.

3. Taal van bydraes

Afrikaans, Duits, Engels, Frans of Nederlands.

4. Kopiereg

Nóg die Buro van die WAT nóg die African Association for Lexicography (AFRILEX) aanvaar enige aanspreeklikheid vir eise wat uit meewerkende skrywers se gebruik van materiaal uit ander bronne mag spruit.

Outeursreg op alle materiaal wat in *Lexikos* gepu-

bliseer is, berus by die Direksie van die Woordeboek van die Afrikaanse Taal. Dit staan skrywers egter vry om hulle materiaal elders te gebruik mits *Lexikos* (AFRILEX-reeks) erken word as die oorspronklike publikasiebron.

5. Oorspronklikheid

Slegs oorspronklike werk sal vir opname oorweeg word. Skrywers dra die volle verantwoordelikheid vir die oorspronklikheid en feitelike inhoud van hulle publikasies.

6. Gratis oordrukke en eksemplare

Skrywers ontvang vyf gratis oordrukke van elke artikel of resensieartikel van hulle wat gepubliseer is asook een gratis eksemplaar van die uitgawe waarin sodanige artikel(s) verskyn het. Skrywers van suiwer evalueerende resensies en van bydraes tot die rubrieke Leksikonotas, Leksikovaria, Projekte en Verslae ontvang vyf gratis oordrukke van hulle bydraes. In laasgenoemde vier kategorieë kan die redaksie egter, afhangend van die aard en omvang van die bydraes, besluit om ook 'n eksemplaar van die betrokke uitgawe aan 'n skrywer toe te ken.

7. Uitnodiging en redaksionele adres

Alle belangstellende skrywers is welkom om bydraes vir opname in *Lexikos* te lewer en aan die volgende adres te stuur:

Die Redakteur: LEXIKOS

Buro van die WAT

Posbus 245

7599 STELLENBOSCH

Republiek van Suid-Afrika

B. VOORBEREIDING VAN MANUSKRIP

Die manuskrip van artikels moet aan die volgende redaksionele vereistes voldoen:

1. Lengte en formaat van artikels

Bydraes moet verkieslik nie 20 getikte A4-bladsye met teks in dubbelspasiëring en ruim kantlyne (ongeveer 2,5 cm) oorskry nie. Manuskrip moet verkieslik in elektroniese formaat as ASCII-teks, as volledig geformateerde Microsoft Word (DOS of Windows) lêers of as WordPerfect (DOS of Windows) lêers op rekenaarskyf (360 KB tot 1.44 MB) voorgelê word. 'n Rekenaardrukstuk van die artikel moet die skyf vergesel. Elke artikel moet voorsien wees van 'n Engelse opsomming van tussen 150 en 250 woorde, sowel as tussen 10 en 30 Engelse sleutelwoorde.

2. Grafika

Een stel duidelike oorspronklike illustrasies, tabelle, grafieke, diagramme, of kwaliteitsafdrukke daarvan, moet voorgelê word. Die plasing van grafika binne die teks moet duidelik aangedui word.

3. Bibliografiese gegewens en verwysings binne die teks

Kyk na onlangse nommers van *Lexikos* vir meer inligting.

INSTRUCTIONS TO AUTHORS

(For a more detailed version of these instructions, please contact the Bureau of the WAT (wat@sun.ac.za) or refer to our website: <http://www.wat.co.za>)

A. EDITORIAL POLICY

1. Type and content of articles

Articles may deal with pure lexicography or with the implications that related fields such as linguistics, general linguistics, computer science and management have for lexicography.

Contributions may be classified in any one of the following categories:

(1) **Articles:** Fundamentally original scientific research that has been done and the results that have been obtained, or reflecting existing research results and other facts in an original, synoptic, interpretative, comparative or critically evaluative manner.

(2) **Review articles:** Research articles presented in the form of a critical review of one or more published scientific sources.

Contributions in categories (1) and (2) are subjected to strict anonymous evaluation by independent academic peers in order to ensure the international research quality thereof.

(3) **Reviews:** An analysis and critical evaluation of published scientific sources and products, such as books and computer software.

(4) **Projects:** Discussions of lexicographical projects.

(5) **Lexiconotes:** Any article containing practice-oriented information, suggestions, problems, questions, commentary and solutions regarding lexicography.

(6) **Lexicovaria:** Any of a large variety of articles, announcements and press releases by lexicographic societies which are of particular value to the practising lexicographer.

(7) **Reports:** Reports on conferences and workshops.

Contributions in categories (3)-(7) must all meet the requirements of academic writing and are evaluated by the editors with this in mind.

2. Academic standard and evaluation procedure

The Department of Education of the South African Government has approved *Lexikos* as a subsidized, i.e. income-generating research journal.

Articles will be evaluated on the following aspects: language and style; conciseness and comprehensibility; problem formulation, reasoning and conclusion; references to the most important and most recent literature; substantial contribution to the specific discipline.

3. Language of contributions

Afrikaans, Dutch, English, French or German.

4. Copyright

Neither the Bureau of the WAT nor the African Association for Lexicography (AFRILEX) accepts any responsibility for claims which may arise from contributing authors' use of material from other sources.

Copyright of all material published in *Lexikos* will be vested in the Board of Directors of the Woordeboek van die Afrikaanse Taal. Authors are free however to use their material elsewhere provided that *Lexikos* (AFRILEX Series) is acknowledged as the original publication source.

5. Originality

Only original contributions will be considered for publication. Authors bear full responsibility for the originality and factual content of their contributions.

6. Free offprints and copies

Authors will receive five free offprints of each of their articles or review articles published, as well as one complimentary copy of the issue containing such article(s). Authors of purely evaluative reviews and of contributions to the categories Lexiconotes, Lexicovaria, Projects, and Reports receive five free offprints of their contributions. In the case of the latter four categories, the editors may, however, depending on the nature and scope of the contributions, decide to grant the author a copy of the issue concerned.

7. Invitation and editorial address

All interested authors are invited to submit contributions for publication in *Lexikos* to:

The Editor: LEXIKOS
Bureau of the WAT
P.O. Box 245
7599 STELLENBOSCH
Republic of South Africa

B. PREPARATION OF MANUSCRIPTS

Manuscripts of articles must meet the following editorial requirements:

1. Length and format

Contributions should not exceed more than 20 type-written A4 pages with double spacing and ample margins (about 2,5 cms). Manuscript should preferably be in electronic form on a (360 KB to 1.44 MB) floppy disk as either ASCII text, fully-formatted Microsoft Word (DOS or Windows) or WordPerfect (DOS or Windows) files. A computer printout of the article should accompany the disk. Each article must be accompanied by an English abstract of 150 to 250 words, and between 10 and 30 English keywords.

2. Graphics

One set of clear original drawings, tables, graphs, diagrams or quality prints thereof must be submitted. The locations of graphics must be clearly indicated in the text.

3. Bibliographical details and references in the text

Examine recent issues of *Lexikos* for details.

HINWEISE UND RICHTLINIEN FÜR AUTOREN

(Nehmen Sie bitte mit dem Büro des WAT Kontakt auf (wat@sun.ac.za) für eine ausführlichere Wiedergabe dieser Hinweise oder besuchen Sie unsere Webseite: <http://www.wat.co.za>)

A. REDAKTIONELLE ZIELSETZUNGEN

1. Art und Inhalt der Artikel

Es können Artikel aufgenommen werden, die sich mit Themen der Lexikographie befassen oder mit Zusammenhängen, die zwischen der Lexikographie und benachbarten Fachgebieten wie z.B. Linguistik, allgemeiner Sprachwissenschaft, Lexikologie, Computerwissenschaft und Management bestehen.

Die Beiträge sollten einer der folgenden Kategorien entsprechen:

(1) **Artikel**, die grundlegend über neue Forschungsansätze und deren Ergebnisse berichten, oder die bestehende Forschungsergebnisse und andere Informationen selbständig, interpretativ, vergleichend oder kritisch bewertend wiedergeben.

(2) **Rezensionsartikel**, die in der Form eines Forschungsartikels eine oder mehrere veröffentlichten wissenschaftlichen Quellen kritisch rezensieren.

Beiträge in Kategorien (1) und (2) werden streng anonym von unabhängigen wissenschaftlichen Experten begutachtet, um ein internationales fachliches Niveau in *Lexikos* zu gewährleisten.

(3) **Rezensionen**, die veröffentlichte wissenschaftliche Quellen und Produkte, wie z.B. Bücher und Software, analysieren und kritisch bewerten.

(4) **Lexikographische Projekte**, die vorgestellt werden.

(5) **Notizen zum Lexikon**, die praxisbezogene Informationen, Vorschläge, Probleme, Fragen, Kommentare und Lösungen hinsichtlich der Lexikographie enthalten.

(6) **Lexikovaria**, die unterschiedliche Beiträge, Ankündigungen und Pressemitteilungen lexikographischer Vereinigungen, die dem praktischen Lexikographen wichtig sein können, einschließen.

(7) **Berichte** über Konferenzen und Workshops.

Beiträge in Kategorien (3)-(7) müssen im akademischen Stil abgefaßt werden. Sie werden von der Redaktion unter diesem Gesichtspunkt beurteilt.

2. Wissenschaftliche Standards und das Beurteilungsverfahren

Das Erziehungsministerium der südafrikanischen Regierung hat *Lexikos* als eine subventionierte, d.h. einkommenerzeugende Forschungszeitschrift anerkannt.

Artikel werden auf Grund der folgenden Gesichtspunkte bewertet: Sprache und Stil; Sachlichkeit und Verständlichkeit; Problembeschreibung, Argumentation und Schlußfolgerung; Hinweise auf die neueste und wichtigste Literatur; wesentlicher Beitrag zum besonderen Fachgebiet.

3. Sprache der Beiträge

Afrikaans, Deutsch, Englisch, Französisch oder Niederländisch.

4. Das Urheberrecht

Weder das Büro des WAT noch die African Association for Lexicography (AFRILEX) übernehmen Verantwortung für Ansprüche, die daraus entstehen könnten, daß Autoren Material aus anderen Quellen benutzt haben.

Das Urheberrecht aller in *Lexikos* publizierten Artikel wird dem Direktorium unseres Büros übertragen. Es steht Autoren jedoch frei, ihren Beitrag anderweitig zu verwenden, vorausgesetzt, *Lexikos* (AFRILEX-Serie) wird als Originalquelle genannt.

5. Originalität

Nur Originalbeiträge werden begutachtet. Autoren tragen die volle Verantwortung für die Originalität und den sachlichen Inhalt ihrer Beiträge.

6. Sonderdrucke und Freixemplare

Autoren erhalten fünf Sonderdrucke ihrer veröffentlichten Artikel oder Rezensionsartikel gratis sowie ein Freixemplar der betreffenden Ausgabe. Rezensenten und Autoren von Beiträgen zu den Kategorien Notizen zum Lexikon, Lexikovaria, Projekte und Berichte erhalten fünf Sonderdrucke ihrer Beiträge gratis. Die Redaktion kann sich jedoch, abhängig von der Art und dem Umfang der Beiträge der letztgenannten vier Kategorien, vorbehalten, dem Autor ein Freixemplar der Ausgabe zu überlassen.

7. Einladung und redaktionelle Adresse

Alle Autoren, die interessiert sind, Beiträge für *Lexikos* zu liefern, sind herzlich willkommen. Sie werden gebeten, ihre Artikel an die folgende Adresse zu schicken:

Der Redakteur: LEXIKOS
Buro van die WAT
Postfach 245
7599 STELLENBOSCH
Republik Südafrika

B. VORBEREITUNG DES MANUSKRIPTS

Ein Artikelmanuskript muß den folgenden redaktionellen Anforderungen entsprechen:

1. Umfang und Format

Beiträge sollen nicht länger als 20 getippte A4-Seiten in zweizeiligem Abstand und mit Randabständen von ca. 2,5 cm sein. Das Manuskript sollte möglichst als elektronischer Text auf einer (360 KB bis 1.44 MB) Diskette vorgelegt werden, entweder im ASCII-Format, oder in formatiertem Microsoft Word (DOS oder Windows) bzw. WordPerfect (DOS oder Windows). Ein Ausdruck des vollständig formatierten Artikels soll mit der Diskette eingereicht werden. Jedem Artikel ist eine Zusammenfassung im Umfang von 150-250 Wörtern beizufügen. Ferner sollen etwa 10-30 inhaltskennzeichnende Stichwörter zu jedem Artikel angegeben werden.

2. Abbildungen

Ein reproduktionsfähiger Satz der originalen Abbildungen, Illustrationen, Tabellen, Graphiken und Diagramme oder Qualitätsabdrucke muß vorgelegt werden. Der Text selber sollte klare Hinweise auf die Position der Abbildungen enthalten.

3. Bibliographische Einzelheiten und Hinweise im Text

Zu Einzelheiten des bibliographischen Systems sind neuere Ausgaben von *Lexikos* einzusehen.

INSTRUCTIONS AUX AUTEURS

(Pour une version plus détaillée de ces instructions, contacter le Bureau du WAT (wat@sun.ac.za)
ou consulter notre website: <http://www.wat.co.za>)

A. POLITIQUE ÉDITORIALE

1. Caractéristiques et contenu des articles

Les articles seront consacrés à la lexicographie pure, ou aux rapports entre la lexicographie et les disciplines voisines telles que la linguistique, la linguistique générale, l'informatique et le management.

Les contributions pourront appartenir à l'une des catégories suivantes:

(1) **Articles:** Recherches scientifiques originales, avec leurs résultats; ou présentations originales, synoptiques, interprétatives, comparatives, évaluatives et critiques des résultats de recherches en cours;

(2) **Articles bilans:** Articles de recherche présentés sous forme de bilan critique de travaux scientifiques déjà publiés.

Les contributions appartenant aux catégories (1) et (2) seront soumises de manière anonyme à des experts spécialistes indépendants afin d'en assurer la qualité scientifique au niveau international.

(3) **Recensions:** Analyses et évaluations critiques de travaux de recherche et de productions scientifiques, telles que livres ou logiciels;

(4) **Projets:** Présentations de projets lexicographiques;

(5) **'Lexiconotes':** Textes contenant des informations pratiques, ou des suggestions, des problèmes, des questions, des commentaires et des solutions concernant des activités lexicographiques;

(6) **'Lexicovaria':** Articles, annonces, communiqués de presse émanant de centres de lexicographie et qui revêtent un intérêt particulier pour les lexicographes;

(7) **Rapports:** Rapports sur des colloques et ateliers.

Les contributions dans les catégories (3) à (7) devront répondre aux exigences de qualité des publications scientifiques et seront évaluées dans cette optique.

2. Critères et procédures d'évaluation

La revue *Lexikos* est reconnue et subventionnée par le Ministère de l'Éducation du gouvernement Sud-Africain comme revue devant générer des revenus.

Les articles seront évalués selon les critères suivants: langue et style, concision et clarté, formulation de la problématique, raisonnement et conclusion, référence aux travaux les plus importants et les plus récents, contribution substantielle à la discipline.

3. Langue des contributions

Afrikaans, allemand, anglais, français ou néerlandais.

4. Copyright

Le Bureau du WAT ou l'*African Association for Lexicography* (AFRILEX) décline toute responsabilité en cas de réclamations motivées par l'utilisation d'autres sources par les auteurs.

Les droits d'auteurs des documents publiés dans *Lexikos* appartiennent au Conseil d'administration du *Woordboek van die Afrikaanse Taal* (WAT). Cependant,

les auteurs sont libres d'utiliser leurs textes dans d'autres publications, à condition d'indiquer *Lexikos* (collection AFRILEX) comme source de la publication originale.

5. Originalité

Seules les contributions originales seront acceptées pour la publication. Les auteurs conservent l'entière responsabilité de l'originalité et du contenu factuel de leur texte.

6. Tirés-à-part et exemplaires gratuits

Les auteurs recevront gratuitement cinq (5) tirés-à-part de chaque article ou article de bilan, ainsi qu'un exemplaire gratuit de la publication contenant cet article.

Les auteurs des recensions et des publications dans les catégories 'Lexiconotes', 'Lexicovaria', Projets et Rapports recevront cinq (5) tirés-à-part de leur texte. Dans ces quatre dernières catégories, le responsable de la revue pourra néanmoins décider, en fonction de la nature et des dimensions des publications, d'accorder un exemplaire gratuit de la revue à leurs auteurs.

7. Adresse de la revue

Les auteurs intéressés sont invités à soumettre leurs propositions à:

L'éditeur: LEXIKOS
Bureau du WAT
Boîte postale 245
7599 STELLENBOSCH
République d'Afrique du Sud

B. PRÉSENTATION DES MANUSCRITS

Les manuscrits se conformeront aux exigences suivantes:

1. Longueur et format

Les contributions ne devraient pas excéder 20 pages dactylographiées, de format A4, avec double espacement et marges suffisantes (environ 2,5 cm), si possible en format électronique (disquette 360KB à 1,44MB), sous forme de fichier ASCII, complètement formaté sous Microsoft Word ou sous WordPerfect (DOS ou Windows). La disquette sera accompagnée d'un tirage papier. Chaque article sera pourvu d'un résumé en anglais de 150 à 200 mots, et de 10 à 30 mots-clés.

2. Tableaux et graphiques

Les dessins, tableaux, graphiques et diagrammes seront envoyés, soit sous leur forme originale soit sous forme d'une copie de bonne qualité. Leur place dans le texte devra être clairement indiquée.

3. Bibliographie et références dans le texte

Voir les exemplaires récents de *Lexikos*.