Lexikos 15

Lexikos 15

Redakteur

Editor

J.C.M.D. du Plessis

Resensieredakteur

Review Editor

T. Harteveld

African Association for Lexicography

Adviesraad / Advisory Board

Redaksiekomitee / Editorial Committee

Inhoud / Contents

Artikels / Articles

Projekte / Projects

Lexikonotas / Lexiconotes

Lexikoprogrammatuur / Lexicosoftware

Lexikovaria / Lexicovaria

Resensies / Reviews

Voorwoord

Vanjaar is dit tien jaar sedert AFRILEX tot stand gekom en *Lexikos* die amptelike mondstuk van AFRILEX geword het. In hierdie nommer verskyn 'n kort oorsig deur dr. Mariëtta Alberts, die huidige President van AFRILEX, oor die aanloop tot die totstandkoming, en die samestelling en aktiwiteite van AFRILEX. Ook word die beleidsrede van prof. R.R.K. Hartmann ingesluit soos vanjaar gedurende Julie gelewer tydens die Tiende Internasionale Konferensie van AFRILEX in Bloemfontein. Dit is 'n gepaste insluiting aangesien prof. Hartmann as voorsitter en fasiliteerder opgetree het by die totstandkoming van AFRILEX tydens die Agste Internasionale Konferensie van ALASA gedurende Julie 1995 op Stellenbosch.

Na verlede jaar se lywige nommer waarin referate opgeneem is wat gedurende November 2002 in Pretoria by AFRILEX se seminaar oor aanleerderwoordeboeke gelewer is, is *Lexikos* vanjaar weer terug tot sy normale grootte. Die gevolg hiervan is die oordra na volgende jaar se nommer van baie artikels wat ontvang, voorberei en gekeur is vir publikasie, maar wat nie in vanjaar se nommer ingesluit kon word nie vanweë 'n gebrek aan ruimte. Die redakteur wil al die bydraers bedank vir hul begrip en hul bereidwilligheid om vir 'n aansienlike tyd te wag vir hul artikels om gepubliseer te word.

'n Woord van dank gaan aan die keurders wat so geredelik en toegewyd die artikels vir hierdie en komende uitgawes van *Lexikos* beoordeel het. Behalwe dat sommige keurders meer as twee artikels gekeur het, is hulle dikwels ook op baie kort kennisgewing genader om artikels te beoordeel. Hul deeglike deurgaan van bydraes verseker grootliks die gehalte van *Lexikos*.

Daar was verskeie veranderinge in die Redaksiekomitee van *Lexikos*. Die redakteur wil sy dank uitspreek aan die volgende uittredende lede vir hul bydrae tot die sukses van *Lexikos* gedurende 'n aansienlike aantal jare: prof. L.G. de Stadler, me. K.P. Kavanagh, dr. M. Lisimba, mnr. T.X. Mfaxa, prof. F.A. Ponelis, prof. R. Sitaram, dr. J. Tsonope en prof. M.W. Visser. Die redakteur verwelkom die volgende lede wat almal reeds in die verlede 'n waardevolle diens aan *Lexikos* gelewer het deur artikels by te dra en te keur: mnr. H.L. Beyer, dr. S. Hadebe, dr. P.A. Mavoungou en mnr. M.C. Mphahlele.

Soos in die verlede het me. Riette Ruthven al die artikels met professionaliteit en bedrewenheid geset, en ook geduldig en noukeurig die veranderinge en verbeteringe aangebring aan die verskeie weergawes waardeur baie van die artikels gegaan het. Me. Tanja Harteveld het met bekwaamheid talle tegniese rekenaarprobleme en -haakplekke opgelos. Hulle altwee verdien dank vir die suksesvolle voltooiing van hierdie nommer van *Lexikos*.

J.C.M.D. du Plessis
Buro van die Woordeboek van die Afrikaanse Taal

Foreword

This year it is ten years since AFRILEX has been established and *Lexikos* has become the official mouthpiece of AFRILEX. In this issue, appears a short overview by Dr Mariëtta Alberts, the present President of AFRILEX, about the preamble to the establishment and the structure and activities of AFRILEX. Also included is the keynote address of Prof. R.R.K. Hartmann as presented this year during July at the Tenth International Conference of AFRILEX in Bloemfontein. This is a fitting inclusion as Prof. Hartmann acted as chairman and facilitator at the establishment of AFRILEX at the Eighth International Conference of ALASA during July 1995 in Stellenbosch.

After last year's bulky issue in which papers presented during November 2002 in Pretoria at the seminar of AFRILEX on learners' dictionaries were featured, *Lexikos* is back to its normal size again this year. This resulted in the carrying-over to next year's issue of many articles which had been received, prepared and refereed for publication, but could not be included in this year's issue because of a lack of space. The editor wishes to thank all contributors for their understanding and willingness to wait a considerable time for their articles to be published.

A word of thanks goes to the referees who so readily and unstintingly evaluated the articles for this and coming issues of *Lexikos*. In addition to refereeing more than two articles, some referees were also called upon on very short notice to judge articles. Their thorough scouring of contributions greatly assures the quality of *Lexikos*.

There have been several changes to the Editorial Board of *Lexikos*. The editor wishes to express his thanks to the following outgoing members for their contribution to the success of *Lexikos* during a considerable number of years: Prof. L.G. de Stadler, Ms K.P. Kavanagh, Dr M. Lisimba, Mr T.X. Mfaxa, Prof. F.A. Ponelis, Prof. R. Sitaram, Dr J. Tsonope and Prof. M.W. Visser. The editor welcomes the following new members, all of whom have already rendered valuable service to *Lexikos* in the past by contributing and refereeing articles: Mr H.L. Beyer, Dr S. Hadebe, Dr P.A. Mavoungou and Mr M.C. Mphahlele.

As in the past, Ms Riette Ruthven typeset all the articles professionally and competently, and also patiently and carefully introduced the changes and corrections to the several versions through which many articles went. Ms Tanja Harteveld efficiently solved many technical computational problems and hitches. Both of them deserve thanks for the successful completion of this issue of *Lexikos*.

J.C.M.D. du Plessis
Bureau of the Woordeboek van die Afrikaanse Taal

'n Woord van AFRILEX

Taalpraktisyns wat gemoeid is met leksikografie en terminologie/terminografie werk met standaarde — óf die standaardvariëteit van 'n taal óf standaarde as sodanig.

Leksikograwe stel woordeboeke saam deur (gewoonlik) die woordeskat van die standaardtaal te dokumenteer. Sou hulle 'n woordeboek saamstel wat nie die standaardvariëteit van die taal behels nie, word sodanige woordeboek gewoonlik 'n dialekwoordeboek genoem.

Terminoloë aan die ander kant dokumenteer terminologie met die doel om vakspesialiste en leke van standaardterme te voorsien wat die ooreenstemmende begrippe verteenwoordig. In vakgerigte werk is dit baie belangrik om gestandaardiseerde terme te gebruik om eksakte kommunikasie nie alleen tussen vakkundiges onderling nie, maar ook tussen vakkundiges en leke te bewerkstellig.

Lede van die taalkundige gemeenskap het onlangs 'n werkbyeenkoms bygewoon om die status van standaardisering van die verskillende amptelike Afrikatale te bepaal. Die lede van die Tegniese Komitees vir Standaardisering van die Nasionale Taalliggame het reeds begin met die proses om die huidige spel- en skryfreëls van die amptelike tale te hersien. Die Khoe en San Nasionale Taalliggaam en die Suid-Afrikaanse Gebaretaal Nasionale Taalliggaam het ook belangstelling getoon om dié onderskeie tale te standaardiseer en om, waar van toepassing, spel- en skryfreëls te skep.

Suid-Afrika het op 19 Augustus 2002 'n nasionale ISO/TC 37-komitee op die gebied van terminologie en verwante taalbronne gestig. StanSA TC 37 is die plaaslike standaardiseringskomitee rakende terminologie en verwante taalbronne en is 'n presiese weergawe van die internasionale ISO/TC 37.

Sekere beginsels rakende standaarde is baie belangrik wanneer die ontwikkeling van standaarde oorweeg word. Nasionale standaarde word deur konsensus deur tegniese komitees ontwikkel wat die behoeftes van plaaslike belanghebbendes verteenwoordig, met inagneming van dissipline, moderne tegnologie, programbestuur en kostedoeltreffendheid. Komitees bestaan uit 'n gebalanseerde verteenwoordiging van gebruikergroepe, belangegroepe sowel as die staat. Nuwe standaarde het die voordeel dat hulle op die behoeftes van die gebruikers van die standaarde kan fokus. Die proses is egter baie tydsaam en duur. Om streeks- en internasionale standaarde te aanvaar, het die voordeel dat goed nagevorste standaarde relatief goedkoop verkry word.

Standaarde kan in wetgewing opgeneem word deur na hierdie streeks- en internasionale standaarde te verwys. Verpligte spesifikasies kan as wetgewing herskryf word. Alle standaarde is vrywillig tensy hulle in wetgewing geïnkorporeer is.

Aangesien standaarde net onder geskrewe wetgewing staan, beïnvloed hulle alle aspekte van die menslike bestaan. ISO-gestandaardiseerde terminologie bestaan uit alle terme en definisies wat in ISO-standaarde voorkom. As sodanig vorm standaarde 'n belangrike element van onderrig en opleiding in vakverwante onderwerpe: die beginfase van enige gespesialiseerde gebied beweeg deur die aanleer van die kernbegrippe (en die terme en definisies wat hulle beskryf).

Lede van StanSA TC 37 verteenwoordig Suid-Afrika by die jaarlikse internasionale ISO TC/37-vergaderings. Die StanSA TC 37-span het 'n baie waardevolle rol te speel in gesprekvoering op internasionale vlak wanneer standaarde wat met terminologie, leksikografie en ander taalkundige middele te make het, hersien word. Hulle doen ook waardevolle ondervinding op in die internasionale omgewing wat hulle dan met hul Suid-Afrikaanse kollegas kan kom deel.

Standaardisasie verteenwoordig 'n hele wêreld van aktiwiteite op verskeie gebiede van die lewe wat help om voordele en volhoubare ondersteuning vir almal in die globale samelewing te skep.

Dr. Mariëtta Alberts
President: AFRILEX
PanSAT
Privaat sak X08
ARCADIA 0007
Tel: +27 (0)12 341 9638   Faks: +27 (0)12 341 5938
E-pos: marietta@pansalb.org.za
AFRILEX-tuisblad:
<http://www.up.ac.za/academic/libarts/afrilang/homelex.html>

A Few Words from AFRILEX

Language practitioners concerned with lexicography and terminology/terminography work with standards — either the standard variety of a language or standards as such.

Lexicographers compile dictionaries by (usually) documenting words in the vocabulary of a standard language. Should they compile a dictionary that does not concern the standard variety of a language it is usually called a dialectal dictionary.

Terminologists on the other hand document terminology with the aim to provide subject specialists and laypeople with standard terms denoting the corresponding concepts. In subject-related work, it is very important to use standardised terms to ensure exact communication not only among subject specialists themselves, but also between subject specialists and laypeople.

Members of the linguistic community recently attended a workshop to determine the status of standardisation of the various official African languages. The members of the Standardisation Technical Committees of the National Language Bodies embarked on a process to revise the current spelling and orthography rules of the official languages. The Khoe and San National Language Body and the South African Sign Language National Language Body also indicated interest in standardising the respective languages and to devise spelling and orthography rules where applicable.

On 19 August 2002, South Africa established a national ISO/TC 37 committee in the field of terminology and other language resources. StanSA TC 37 is the local standardisation committee regarding terminology and other language resources and is a mirror committee of the international ISO/TC 37.

Certain principles regarding standards are very important when the development of standards is considered. National standards are developed by consensus through technical committees, representing the needs of local stakeholders, bearing in mind discipline, modern technology, programme management, and cost effectiveness. Committees are comprised of a balanced representation of user groups, interest groups as well as the government. New standards have the advantage that they could focus on addressing the needs of the users of the standards. However, the process is very time-consuming and costly. Adopting regional and international standards has the advantage of having well-researched standards obtained relatively cheaply.

Standards can also be incorporated into legislation by referring to these regional and international standards. Compulsory specifications can be re-drafted as legislation. All standards are voluntary unless incorporated into legislation.

As standards are second to written law, they affect all aspects of human

life. ISO standardised terminology comprises of all terms and definitions occurring in ISO standards. As such, standards are an important element in the teaching of and training in subject-related topics: the initiation phase of any specialised field passes through learning of the main concepts (and the terms and definitions describing them).

Members of StanSA TC 37 represent South Africa at the annual international ISO TC 37 meetings. The StanSA TC 37 team has a very valuable role to play in deliberating at international level when standards dealing with terminology, lexicography and other language resources are under revision. They also gain valuable experience from the international scene, which they can then come to share with their South African colleagues.

Standardisation comprises a whole world of activities in various spheres of life which helps in creating benefits and sustainability for everyone in global society.

Dr Mariëtta Alberts
President: AFRILEX
PanSALB
Private Bag X08
ARCADIA 0007
Tel: +27 (0)12 341 9638   Fax: +27 (0)12 341 5938
Email: marietta@pansalb.org.za
AFRILEX home page:
<http://www.up.ac.za/academic/libarts/afrilang/homelex.html>

Redaksionele doelstellings

Lexikos is 'n tydskrif vir die leksikografiese vakspesialis en word in die AFRI-LEX-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 uitgee 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 leksiko-grafiese 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 be-werkstellig en te bevorder;

(4) om die interdissiplinêre aard van die leksikografie, wat ook terreine soos die taalkunde, algemene taalwetenskap, leksikologie, rekenaarweten-skap, 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* (AFRI-LEX) 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 han-del, 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 spora-dies 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 AFRI-LEX-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 Lexiko-graphen und Linguisten;

(3)   Kontakt herstellen und fördern zwischen südafrikanischen und aus-ländischen lexikographischen Projekten;

(4)   die Aufmerksamkeit lenken auf die interdisziplinäre wissenschaftliche Praxis der Lexikographie, die Beziehung aufweist zur Linguistik, allge-meinen 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 publiées sous forme de publications séparées dans la collection AFRILEX.

Standaardwoordeboek en kernwoordeboek: 'n Vergelykende evaluering

Herman L. Beyer, *Departement Germaanse en Romaanse Tale, Universiteit van Namibië, Windhoek, Namibië (hbeyer@unam.na)*

Opsomming: Hierdie artikel probeer om 'n evaluering te gee van die vierde uitgawe van die *Verklarende Handwoordeboek van die Afrikaanse Taal* wat in 2000 verskyn het, in vergelyking met *Kernwoordeboek*, waarvan die eerste uitgawe in 2000 gepubliseer is. Die ooreenkomste en verskille tussen die eksponente van hierdie twee woordeboektipes in die Afrikaanse leksikografie word uitgelig deur te fokus op die voorwerk, die makrostruktuur, die toegangstruktuur en aspekte van die mikrostruktuur in die sentrale teks. Gebruikersvriendelikheid en die leksikograaf se plig om verantwoording aan die gebruiker te doen, geld as belangrike kriteria. Die konsep *kernwoordeskat* vir leksikografiese doeleindes word ondersoek en metodes om die kernwoordeskat as makrostruktuur van 'n kernwoordeboek te bepaal, word krities beskou. Ander Afrikaanse woordeboeke word vergelykend betrek ten einde bepaalde aspekte van heersende leksikografiese praktyk uit te wys.

Sleutelwoorde: EKSTERNE TOEGANGSTRUKTUUR, FREKWENSIE, GEBRUIKERS-VRIENDELIK, INTERNE TOEGANGSTRUKTUUR, KERNWOORDEBOEK, KERNWOORDESKAT, LEKSIKALE ITEM, LEKSIKOGRAAF, LEKSIKOGRAFIE, LEKSIKON, LEMMA, LEMMATEKEN, LEMMATISERING, MAKROSTRUKTUUR, METATEKS, MIKROSTRUKTUUR, NESTING, NISTING, SELEKSIE, SENTRALE TEKS, STANDAARDTAAL, STANDAARDWOORDEBOEK, SUBLEMMATISERING, TEKSVERDIGTING, TIPOLOGIE, TOEGANGSTRUKTUUR, TOELIGTING, VERANTWOORDING, VOORWERK, WOORDESKAT

Abstract: Standard Dictionary and Core Vocabulary Dictionary: A Comparative Evaluation. This paper attempts to give an evaluation of the fourth edition of the *Verklarende Handwoordeboek van die Afrikaanse Taal*, which appeared in 2000, compared to *Kernwoordeboek*, of which the first edition was published in 2000. The similarities and differences between the exponents of these two dictionary types in Afrikaans lexicography is addressed by focusing on the front matter, the macrostructure, the access structure and aspects of the microstructure within the central text. Userfriendliness and the lexicographer's obligation to give account to the user serve as important criteria. The concept *core vocabulary* for lexicographical purposes is examined and methods to determine the core vocabulary as macrostructure of a core vocabulary dictionary is viewed critically. References to other Afrikaans dictionaries are made to point out certain aspects of current lexicographical practice.

Keywords: ACCESS STRUCTURE, ACCOUNT, CENTRAL TEXT, CORE VOCABULARY, CORE VOCABULARY DICTIONARY, EXTERNAL ACCESS STRUCTURE, FREQUENCY, FRONT MATTER, INTERNAL ACCESS STRUCTURE, LEMMA, LEMMA SIGN, LEMMATISA-

TION, LEXICAL ITEM, LEXICOGRAPHER, LEXICOGRAPHY, LEXICON, MACROSTRUC-
TURE, METATEXT, MICROSTRUCTURE, NESTING, NICHING, SELECTION, STANDARD DIC-
TIONARY, STANDARD LANGUAGE, SUBLEMMATISATION, TEXTUAL CONDENSATION,
TYPOLOGY, USER'S GUIDE, USERFRIENDLY, VOCABULARY

1.   Inleiding

Die *Verklarende Handwoordeboek van die Afrikaanse Taal* (voortaan HAT) het in
2000 sy vierde uitgawe (Odendal en Gouws 2000 — voortaan HAT[4]) beleef. Dit
is slegs ses jaar na die derde uitgawe (Odendal et al. 1994 — voortaan HAT[3]),
wat die kortste periode is wat tussen twee uitgawes van dié woordeboek ver-
loop het: Tussen die eerste uitgawe (Schoonees et al. 1965) en die tweede uit-
gawe (Odendal et al. 1979) het 14 jaar verloop, en tussen die tweede uitgawe en
HAT[3] het 15 jaar verloop. Sedert die eerste uitgawe het HAT onder leiding van
P.C. Schoonees en veral sedert die tweede uitgawe onder F.F. Odendal homself
as een van die mees gebruikte handwoordeboeke in Afrikaans gevestig en 'n
belangrike standaardiserings- en normeringsfunksie vervul (Carstens 1995: 139).
Volgens Carstens (1995: 139-142) is die HAT in sy verskillende uitgawes 'n
sosiohistoriese dokument wat veral sedert die demokratisering van Suid-Afrika
in 1994 'n fokusverskuiwing van elitisme na populisme ondergaan het. HAT
kan dus ook as sosiokulturele produk beskryf word.

    HAT[3] word in minstens twee resensieartikels, naamlik Carstens (1995) en
Van der Merwe (1994) bespreek. Volgens die "Voorwoord by die Vierde Uit-
gawe" in HAT[4] (2000: v) het dié uitgawe nie ingrypende verandering onder-
gaan nie. Die fokus by die hersiening van HAT[3] het naamlik geval op die daar-
stel van 'n meer verteenwoordigende makrostruktuur ten opsigte van kontem-
porêre Afrikaans, enkele wysigings in die interne toegangstruktuur ten einde
gebruikersvriendelikheid te verhoog, en die aangawe van meer sitate, ook uit
ligter literatuur, as voorbeeldmateriaal. Wanneer HAT[4] met HAT[3] vergelyk
word, sal daar dus nie van ingrypende mikrostrukturele veranderinge sprake
wees nie; daarom geld Carstens (1995) en Van der Merwe (1994) se kommen-
taar op HAT[3] ook grootliks vir HAT[4].

    In 2000 verskyn ook *Kernwoordeboek* (Odendal 2000 — voortaan KW), 'n
splinternuwe woordeboek in die Afrikaanse leksikografiese spektrum. Volgens
die "Voorwoord" (KW 2000: iii) vul KW 'n leemte "wat betref die omvang,
struktuur en hulpmiddels, wat tot dusver in Afrikaanse verklarende woorde-
boeke bestaan het. […] Eerstens is, met weinig uitsonderings, woorde en vorme
uit die standaardtaal opgeneem. […] Die strewe was om die kernwoordeskat
van Afrikaans te verklaar."

    Die doel van hierdie artikel is om HAT[4] met KW te vergelyk ten einde te
kan wys op die ooreenkomste en verskille tussen standaardwoordeboek en
kernwoordeboek ten opsigte van tipologiese plasing, makrostruktuur, toegang-
struktuur en mikrostruktuur, sowel as op moontlike tekortkominge in dié
struktuur. Dit gaan dus nie soseer oor 'n vergelyking tussen HAT[4] en HAT[3] nie,

hoewel enkele veranderinge in die jongste uitgawe wel uitgelig word. Ander beperkte en standaardwoordeboeke word ook vergelykenderwys betrek. Die voortekste van HAT[4] en KW sal ook krities beskou word.

2.   Tipologiese plasing

2.1   HAT[4]

HAT[4] staan in die sentrum van die tipologiese spektrum van Afrikaanse woordeboeke — as sinchroniese standaard verklarende woordeboek. Wat HAT[4] wil doen, "is om die algemeenste woorde van Standaardafrikaans op te neem en in hulle verskillende betekenisonderskeidinge in te deel, verder met soveel voorbeelde van hulle gebruikstoepassing as wat ruimte toelaat" (HAT[4] 2000: vi). HAT[4] vervul dus primêr 'n deskriptiewe funksie deurdat die leksikon van Standaardafrikaans beskryf word. 'n Normatiewe funksie is egter ook ter sprake, naamlik dat in die mikrostruktuur onder andere bepaalde oordele oor sommige lemmata as synde nie verteenwoordigend van Standaardafrikaans nie, in die vorm van hoofsaaklik leksikografiese etikette, maar by geleentheid ook minder eksplisiet, uitgespreek word. Dit hou onder meer in dat nie slégs standaardtaalitems in die makrostruktuur opgeneem word nie: "Nuwe woorde wat algemeen in die spreektaal voorkom, vind neerslag in hierdie uitgawe van die woordeboek. Dit sluit ook 'n aantal leenwoorde en verafrikaanste vorme uit Engels in wat veral in die geselstaal voorkom, bv. *cool, sorry, oukei, orraait*.[1] Hierdeur gee HAT ook erkenning aan die taalgebruik van 'n groot groep sprekers van die standaardvariëteit van Afrikaans" (HAT[4] 2000: v).

Die teikengebruiker van HAT[4] is waarskynlik die gemiddelde taalgebruiker, dit wil sê die gemiddelde volwasse moedertaalspreker van Afrikaans wat oor 'n graad 12-skoolkwalifikasie of die ekwivalent daarvan beskik.

2.2   KW

Die tipologiese plasing van KW is enigsins problematies. In die "Voorwoord" word geen melding van die spesifieke doel of teikengebruiker van dié woordeboek gemaak nie; daar word slegs genoem dat die woordeboek daarna strewe om die kernwoordeskat van Afrikaans te verklaar (KW 2000: iii). In die Afrikaanse leksikografiese spektrum bestaan 'n ander woordeboek wat ook daarop aanspraak maak dat dit die kernwoordeskat van Afrikaans verteenwoordig, naamlik *Basiswoordeboek van Afrikaans* (Gouws, Feinauer en Ponelis 1994 — voortaan BW). In die "Voorwoord" van BW word egter duidelik gemeld dat die woordeboek afgestem is op die aanleerder van Afrikaans — dit maak 'n sinvolle tipologiese plasing moontlik. (Oor die aanspraak op die verteenwoordiging van die kernwoordeskat van Afrikaans deur én BW én KW word in die bespreking van die makrostruktuur uitgebrei — vgl. 4.1).

Die feit dat KW geen spesifieke teikengebruiker of gebruiksdoel in sy voorwerk vermeld nie, laat die evalueerder en potensiële gebruiker oor aan die struktuur van die woordeboek om te bepaal vir watter doeleindes die woordeboek wel geskik kan wees. Dit is dus nie noodwendig onmiddellik seker watter "leemte" KW in die bestaande versameling Afrikaanse verklarende woordeboeke vul nie.

Die afwesigheid van 'n Afrikaanse verklarende skoolwoordeboek, die aanspraak dat KW die kernwoordeskat van Afrikaans verteenwoordig, en aspekte van die toegang- en mikrostruktuur van KW dui daarop dat dié woordeboek moontlik as 'n skoolwoordeboek op die moedertaal- of tweedetaalspreker van Afrikaans in die primêre en/of sekondêre skoolfases afgestem kan wees. Die woordeboek blyk nie geskik te wees vir vreemdetaalaanleerders van Afrikaans nie; daarvoor bestaan BW buitendien. KW is dus moontlik as skoolwoordeboek saamgestel. Die voorwerk van KW laat ongelukkig nie 'n sekerder begrensing van die teikengebruikergroep toe nie. In die res van hierdie artikel sal die perspektief van die skoolleerder as moontlike teikengebruiker gehandhaaf word.

3.   Die voorwerk

3.1   Leksikografiese verantwoording

Volgens Wiegand (1996: 135-136) se teksteorie kan die tekste in 'n woordeboek aan een van drie tipes behoort: die verpligte voorwerk, die verpligte sentrale teks en 'n opsionele agterwerk. Saam vorm hierdie tekste die makroteks. Die verband tussen die verskillende tekste in die makroteks word deur 'n min of meer gedifferensieerde woordeboekinterne mediostruktuur daargestel.[2] Dit is duidelik dat Wiegand se benadering 'n wisselwerking tussen die verskillende tekste in die woordeboek veronderstel; die onderskeie tekste staan apart, maar werk saam. Die verpligte voorwerk bevat 'n verpligte metateks, oftewel 'n gebruikersgids (Wiegand 1996: 136). Die gebruikersgids is 'n uiters belangrike teks in die woordeboek, want daarin verklaar die leksikograaf die werkswyse wat in die saamstel van die ander tekste (vernaamlik die sentrale teks) gevolg is. Soos wat vir die res van die woordeboek geld, is die gebruikersgids op die gebruiker van die woordeboek gerig; dus behoort die interne en eksterne struktuur van hierdie teks so aangebied te word dat dit vir die gemiddelde gebruiker toeganklik is.

3.2   Voortekste in HAT en KW

Wat metatekste betref, het HAT eers relatief onlangs op die toneel verskyn. Tot en met die tweede uitgawe (Odendal et al. 1979) is daar geen leksikografiese verantwoording nie. HAT[3] verteenwoordig 'n waterskeiding in gebruikersleiding in dié woordeboek. 'n Uitgebreide leksikografiese verantwoording is in die "Voorwoord by die Derde Uitgawe" opgeneem, wat sonder wysiging in

HAT[4] verskyn. Hierdie verantwoording sluit die woordeboek se doelstelling en werkswyse in.

Hoewel 'n volledige leksikografiese verantwoording soos wat in HAT[4] verskyn goeie leksikografiese praktyk verteenwoordig, doen die tekstuele struktuur waarin die metateks gegiet is egter afbreuk aan die meriete daarvan. Die probleem lê hier naamlik op twee vlakke. Eerstens behoort die gebruikersleiding nie deel van die "Voorwoord" te vorm nie, maar behoort onder 'n eie opskrif soos "Gids/Toeligting vir die gebruiker" as aparte voorteks te figureer. Tans val opmerkings van 'n meer persoonlike aard en die weergee van 'n tipiese artikelstruktuur binne een teks, wat die funksionaliteit daarvan kompromitteer. Volgens Wiegand (1996: 140) bestaan die voorwerk van 'n woordeboek nie slegs uit een teks nie; verskillende voortekste kan (en behoort) onderskei te word, waaronder die voorblad, die inhoudsopgawe, die voorwoord, die lys van afkortings, kort notas oor gebruik, leksikografiese verklarings en meer uitgebreide notas oor gebruik. Tweedens behoort die eksterne struktuur van die teks sodanig te wees dat die gebruiker maklik toegang tot 'n bepaalde onderafdeling kan kry. Dit kan bewerkstellig word deur die teks in duideliker onderafdelings te verdeel, elk met 'n eie opskrif wat tipografies van die res van die teks verskil. Voorbeelde van sulke opskrifte sou insluit "Aanduiding van woordsoort", "Afleidings en samestellings", e.s.m. Toeganklikheid tot die metateks sal sodoende verhoog word en die gebruiker kan gevolglik meer geneig wees om die metateks te raadpleeg om 'n spesifieke probleem op te los. Die relatiewe ontoeganklikheid van die teks soos dit tans daar uitsien, skrik die woordeboekgebruiker eerder af as wat dit hom/haar aanmoedig om die gebruikersleiding aanvanklik te lees en/of later te raadpleeg.

'n Welkome toevoeging tot die voorwerk in HAT[4] is 'n geannoteerde skematiese uiteensetting van die tipiese struktuur van 'n woordeboekartikel. Hierdie uiteensetting gee aan die gebruiker 'n duidelike oriëntering ten opsigte van die datakategorieë wat in woordeboekartikels voorkom, en ten opsigte van waar in die artikels elke kategorie gehuisves word. Die struktuuruiteensetting word ook woordeliks in die "Voorwoord" verduidelik.

Leksikografiese etikette word onder die opskrif "Etikette in hierdie woordeboek" in HAT[4] se voorwerk opgeneem. Van die 69 etikette wat hier aangebied word, bestaan slegs nege nié uit afkortings nie; dié etikette wat wel afkortings is, se volvorme word verstrek. 'n Kort verklaring van die omstandighede waaronder sodanige etikette in die sentrale teks aangewend word, word aangebied waar van toepassing. Voorbeelde van lemmata waar 'n bepaalde etiket tipieserwys sal voorkom, word ook ingesluit. Vergelyk die inskrywing van die volgende etiket:

(1)   **hist.** histories
      Aanduiding dat die woord om geskiedkundige redes opgeneem word alhoewel
         dit meestal nie meer algemeen gebruik word nie, bv. *Gaskonjer, gasreg, geheim-*
         seel, groenboek, independente fiskaal, aan die *kaak* stel, *kortelas, kroon* (muntstuk),
         Onderveld, wapenkneg.

Hierdie inskrywing gee aan die gebruiker volkome leiding oor die funksie en aanwending van etikette, en verteenwoordig goeie leksikografiese praktyk. 'n Enkele probleem met hierdie afdeling in HAT[4] se voorwerk is egter dat ook ander redaksionele afkortings onder dié opskrif tuisgebring word, naamlik **mntl.** (*moontlik*) en **wsk.** (*waarskynlik*). Volgens Gouws (1988: 4) verwys die term *etiket* na "daardie inskrywing(s) in 'n woordeboekartikel wat as merkers optree om die lemma of 'n ander inskrywing in die artikel se beperkinge ten opsigte van byvoorbeeld stylaard, gebruiksfeer en geografiese voorkoms aan te dui". Dit is duidelik dat hierdie term nie ander redaksionele afkortings in bestek het nie; trouens, in HAT[3] word dieselfde afdeling onder die opskrif "Gebruiklikste etikette en afkortings" tuisgebring.[3] Redaksionele afkortings het 'n ander funksie as etikette: hulle word naamlik ter wille van ruimtebesparing in veral die betekenisparafrases en koteksinskrywings aangewend, en verteenwoordig 'n teksverdigtingstrategie. Ander voorbeelde van redaksionele afkortings wat in HAT[4] voorkom, is **bv.** (*byvoorbeeld*), **e.d.** (*en dergelike*), **ens.** (*ensovoorts*) en **vgl.** (*vergelyk*). Die feit dat hierdie afkortings standaardafkortings in Afrikaans is en ook as lemmata opgeneem is, behoort volledigheidshalwe minstens genoem te word. Ter wille van akkurater leksikografiese verantwoording en groter gebruikersvriendelikheid sou twee aparte voortekste, een wat redaksionele afkortings verstrek en een wat etikette aangee, beter deug. (Etikette word klaarblyklik nie in KW aangewend nie; meer hieroor in die bespreking van die mikrostruktuur — vgl. 6.4.2.)

Die voorwerk in KW bestaan uit 'n "Voorwoord" waarin inligting oor die seleksie van lemmata, die metode van betekenisverklaring, die toepassing van voorbeeldsinne, die groepering van afkortings en enkele aspekte van die gebruik van nietipografiese struktuurmerkers aangebied word. 'n Sistematiese en volledige verklaring van die doelstelling en die werkswyse van die woordeboek, soos wat in HAT[4] verskyn, is afwesig (vgl. weer 2.2). Ook geen geannoteerde skematiese uiteensettings van tipiese woordeboekartikels verskyn ter toeligting in die voorwerk nie. Indien KW as skoolwoordeboek gebruik word, kan hierdie gebrekkige metateks leerders se opleiding in effektiewe woordeboekgebruik kniehalter. Een van die belangrikste punte wat by die opleiding van woordeboekgebruik onder die aandag van gebruikers gebring behoort te word, is (a) die ideaal dat die gebruiker die metateks moet deurlees voordat die sentrale teks geraadpleeg word, en (b) dat die gebruiker enige tyd na die metateks behoort terug te keer indien enige onduidelikheid in die raadpleeg van die sentrale teks voorkom. Indien die woordeboek wat as opleidingsinstrument gebruik word nie 'n volledige en gebruikersvriendelike metateks bevat nie, sal hierdie basiese gebruiksbeginsels deur die woordeboek self ondermyn word.

Ter opsomming: Die metateks in HAT[4] verteenwoordig goeie leksikografiese praktyk deurdat dit 'n volledige leksikografiese verantwoording aanbied. Die metateks in KW is onvoldoende en behoort verbeter te word, veral gesien in die lig van die potensiële teikengebruiker.

4.   Die makrostruktuur

Die makrostruktuur van 'n woordeboek het te doen met die leksikale items wat geselekteer is om as lemmata in die woordeboek opgeneem te word (Gouws 1989: 40).

In die geval van HAT[4] bestaan die makrostruktuur uit die algemeenste leksikale items wat in Standaardafrikaans voorkom, sowel as daardie leksikale items wat deel van die algemene geselstaal van sprekers van Standaardafrikaans uitmaak. Laasgenoemde groep lemmata behoort spesiale bewerking te ondergaan deur van leksikografiese etikette voorsien te word. Ander lemmata wat van etikette voorsien behoort te word, is onder meer leksikale items wat die leksikon van 'n bepaalde vaktaal of sosiolek verteenwoordig.

Beide KW en BW maak daarop aanspraak dat hulle die kernwoordeskat van Afrikaans verteenwoordig, hoewel dié twee woordeboeke verskillende gebruikersgroepe teiken. Hierdie verskynsel vra om die uitklaring van die begrip *kernwoordeskat*.

4.1   Die konsep *kernwoordeskat*

Die term *kernwoordeskat* het geen eenduidige denotasie nie. Ten minste drie breë toepassings kan ten opsigte van dié konsep ingebring word. Die eerste toepassing het te doen met die struktuur van daardie elemente waaruit die leksikon bestaan; die tweede is gebaseer op die produk van 'n bepaalde fase in die taalverwerwingsproses van die taalgebruiker; en die derde toepassing het betrekking op die gebruiksfrekwensie van die elemente waaruit die leksikon bestaan. Die potensiële toepassing en waarde van hierdie drie benaderings vir die leksikografie word hieronder uiteengesit, en word gevolg deur 'n voorstel vir 'n semanties bevredigende benadering tot die vasstel van die omvang van 'n kernwoordeskat.

4.1.1   Die struktuur van die leksikonelemente

Eksteen (1989: 359) gebruik 'n konsentriese model om die begrip *kernwoordeskat* te verduidelik, soos in Figuur 1 weergegee.
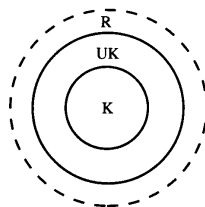
Figuur 1: Eksteen (1989) se konsentriese woordeskatstruktuur (K = kernwoordeskat; UK = uitgebreide kernwoordeskat; R = randwoordeskat)

Volgens Eksteen (1989: 358) bestaan die kernwoordeskat (K) uit "die sentrale groep ongelede leksikale items (simplekse items) waaruit die uitgebreide kernwoordeskat (UK) gevorm word, wanneer leksikaal gelede leksikale items (komplekse items) ontstaan, hoofsaaklik deur afleiding en samestelling. Die randwoordeskat (R) hieromheen bestaan uit versamelings van minder of meer gespesialiseerde leksikons op allerlei diverse terreine, vanaf ambagstale, groepstale en kunstale tot by wetenskaplike vaktale". Dit is 'n hoogs vereenvoudigde representasie van die leksikon wat nie vir leksikografiese doeleindes toepasbaar is nie. Beide simplekse en komplekse leksikale items verteenwoordig vanuit 'n leksikografiese oogpunt leksikonitems uit eie reg wat vir opname in 'n woordeboek kwalifiseer (vgl. o.m. Gouws 1989: 87).

4.1.2   Die bereik van 'n bepaalde fase in die taalverwerwingsproses

Die tweede benadering tot die beskrywing van die begrip *kernwoordeskat* gaan van die standpunt uit dat die kernwoordeskat bestaan uit die produk van 'n bepaalde fase in die taalverwerwingsproses wat deur 'n taalgebruiker in sy/ haar ontwikkeling bereik is. Hierdie produk vind neerslag in die versameling leksikale items wat die kernwoordeskat opmaak. Om die omvang van hierdie produk te bepaal, is egter geen eenvoudige taak nie.

Singleton (2000: 167-170) beskryf drie fases van woordeskatverwerwing. In die eerste fase maak die kind uit presies wat woorde is, hoe hulle gebruik kan word om na sake te verwys, e.s.m. Sowat 30 woorde word in hierdie fase verwerf. In die tweede fase vind 'n sogenaamde woordeskatontploffing plaas waartydens 'n groot aantal woorde teen 'n hoë tempo aangeleer word. Volgens Van der Merwe (1968: 111) leer 'n kind binne die eerste drie lewensjare die grootste deel van sy/haar alledaagse woordeskat aan. Die derde fase is een waarin leksikale kennis gekonsolideer, hersien en herorganiseer word. Volgens Singleton (2000: 170) begin die derde fase in die voorskoolse jare, en gaan voort in die primêreskooljare en selfs tot met volwassenheid. Trouens, Carroll (1971: 124, vlgs. Singleton 2000: 170) beweer dat 'n persoon se woordeskat tot op die ouderdom van 40 beduidend toeneem, en Diller (1971: 29, vlgs. Singleton 2000: 170) is van mening dat daar geen punt voor sterfte is dat 'n afname in leksikale verwerwing voorspel kan word nie. Hieruit volg dat dit dus baie moeilik sal wees om die omvang van 'n leksikon vir 'n bepaalde fase te probeer identifiseer. Boonop bevraagteken Dromi (1987: 12) die algemene gebruik om taalverwerwing in fases te verdeel sonder dat sodanige identifikasies wetenskaplik deeglik verantwoord is. Voeg hiertoe die byna oneindige aantal intrinsieke en ekstrinsieke faktore wat elke individu se woordeskatverwerwing beïnvloed (bv. die mentale vermoë om woorde in semantiese velde te groepeer, die mate van blootstelling aan die taal — vgl. Dromi 1987, Gleitman en Landau 1994 en Singleton 2000), en die gevolgtrekking moet wees dat dit "onmoontlik [is] om 'n naastenby eenvormige kernwoordeskat vir volwasse [of enige ander — HLB] Afrikaanssprekendes te bepaal" (Combrink 1990: 225). Selfs Combrink (1990: 225)

se intuïtiewe werksdefinisies is moeilik uitvoerbaar: "Die kernwoordeskat is dié woordeskat waaroor 'n kind beskik wanneer hy die eerste keer skool toe gaan. Die randwoordeskat is dié woordeskat wat 'n kind aanleer nadat hy begin skoolgaan het." Buitendien is daar volgens Van der Merwe (1968: 145) ook nie konsensus oor wat presies die omvang van 'n kernwoordeskat sou wees nie: "Hoe groot hierdie kernwoordeskat is, of behoort te wees, is nie met sekerheid te bepaal nie: sommige stel dit op minstens 10 000 woorde; andere meen dat 1 000 voldoende is."

Om 'n enigsins akkurate, eenvormige kernwoordeskat te identifiseer en te karteer na aanleiding van die taalontwikkelingsfases van taalgebruikers is tans dus nie 'n praktiese alternatief vir die leksikograaf nie. Hierdie teoretiese konstruk sal ter onderskeid vervolgens die *kognitiewe kernwoordeskat* genoem word.

4.1.3   Gebruiksfrekwensie as maatstaf

Waarskynlik die praktiesste en eenvoudigste manier — gegee hedendaagse tegnologie — om 'n (kern)woordeskat se omvang en aard vas te stel, is deur middel van frekwensietellings. Deur taalkorpora te gebruik, is dit moontlik om te bepaal watter items in die woordeskat die meeste in gesproke en geskrewe tekste voorkom. Deur die toepassing van 'n bepaalde drempelwaarde sou dus beweer kon word dat items met 'n frekwensie hoër as die drempelwaarde deel is van die kernwoordeskat, terwyl items met 'n laer frekwensie as die drempelwaarde buite die grense van die kernwoordeskat val. Die samestellers van BW meld dat dié woordeboek op die kernwoordeskat van Afrikaans fokus en dat "by die keuse van woorde vir behandeling [...] die skrywers hulle laat lei (het) deur 'n bestaande frekwensietelling wat 'n lysting bied van die gebruiklikste Afrikaanse woorde" (BW 1994: iii). Daar word egter nie gemeld watter spesifieke frekwensietelling ter sprake is nie. Die woordelys *40 000 woorde* (Bresler 1978) se makrostruktuur is ook grootliks op 'n frekwensietelling gebaseer (Bresler 1978: iv). Vanuit 'n kognitiewe perspektief op die leksikon sou die makrostruktuur van BW egter nie noodwendig 'n toereikende verteenwoordiging van die konsep *kernwoordeskat* wees nie. 'n Frekwensietelling sou in die eerste plek dui op die *aktiewe woordeskat*, dit wil sê daardie items in die woordeskat wat die frekwentste in taalgebruik voorkom. Die situasie kan wel wees dat sekere items wat as kernwoordeskatitems geld, nie noodwendig frekwent in taalgebruik voorkom nie. Een van die redes hiervoor is waarskynlik dat geen taalkorpus, hoe uitgebreid ook al, volledig verteenwoordigend van outentieke taalgebruik kan wees nie. Veral spreektaal word tans slegs in 'n beperkte mate in Afrikaanse korpora verteenwoordig. As voorbeeld kan die leksikale item *mamma* geneem word. Dié item val heel waarskynlik onder die eerste 30 woorde wat 'n kind aanleer en kan derhalwe tereg as 'n kernwoordeskatitem beskou word. Die vraag is egter of dié item ook onder die eerste 30 frekwentste woorde in 'n frekwensietelling sou opduik. Die punt hier is nie dat die hoë frekwensie van die item *mamma* in twyfel getrek word nie, maar eerder hoe akku-

raat 'n frekwensietelling die gebruik daarvan sal weergee. Op die oog af blyk dit nie 'n ernstige probleem te wees nie. Indien dit egter gaan om woorde wat nader aan die gekose drempelwaarde vir sogenaamde kernwoordeskatitems val, kan dit wees dat werklike kernwoordeskatitems uiteindelik nie in die makrostruktuur van 'n kernwoordeboek opgeneem word nie, óf omdat daardie items wél frekwent voorkom maar om die een of ander rede nie in die korpus opgeneem is nie, óf omdat daardie items wél deel uitmaak van die kernwoordeskat maar deel is van die passiewe kernwoordeskat. Teen hierdie agtergrond is 'n spreker se kernwoordeskat nie noodwendig gelyk aan sy/haar aktiewe woordeskat nie (vgl. Van der Merwe 1968: 146). 'n Woordeboek wat op frekwensietellings alleen steun vir die afbakening van 'n kernwoordeskat sal nie 'n akkurate weergawe van die kernwoordeskat kan bied nie. Waarskynlik as gevolg van 'n gebondenheid aan die genoemde frekwensietelling is onder meer die volgende 55 leksikale items nie in die makrostruktuur van BW in die artikeltrajek van S opgeneem nie, terwyl die items volgens intuïtiewe oordeel alledaagse Afrikaans verteenwoordig en waarskynlik deel van die kernwoordeskat van Afrikaans vorm:

(2)   *sebra, sekonde, seldery, September, serp, servet, sif, sjokolade, skaats, skaatsplank, ski, skilpad, skoenlapper, skolier, slaai, slaap* (sywand v.d. voorkop)*, smeer, smiddags, s'n, snags, sneeu, sonbrand, soontoe, sopnat, sosatie, sous, spaghetti, spanspek, spasie, spek, speld, spinasie, spoel* (water)*, spons, springbok, staander, staat, standbeeld, stapel* (~ boeke)*, steen, steenbok, stillees, stort, stortbad, stronk, strooimeisie, stroop* (s.nw.)*, suid, Suidpool, supermark, sus, sussie, suurlemoen, swik* (verstuit)*, sypaadjie.*

Terwyl die items in (2) nie in BW gelemmatiseer is nie, is die volgende items wél as lemmata opgeneem, hoewel sommige van hulle op die kernwoordeskatskaal waarskynlik laer sal lê (dus verder van die "kern" af) as die items in (2):

(3)   *saamsmelt, salueer, sambok, sambreel, sampioen, sarkasme, sauna, saxofoon, scenario, sedertdien, seekoei, seeman, seespieël, seggenskap, sekskat, selfvertroue, sentrum, senuweeagtig, sieldodend, sielkunde, sigbaar, simbool, sinoniem, skadeloos, skakelbeampte, skandaal, skeidsregter, skiereiland, skors* (straf)*, slagoffer, slagyster, sleepboot, soggens, solo, sondebok, soogdier, sowel, spandabel, spannend, speurder, sport* (trappie)*, spraakloos, spysenier, staatsdiens, staking, standaardtaal, steenkool, stikdonker, stokkiesdraai, strokiesverhaal, suiwel.*

Die items in (2) toon hoe gebondenheid aan 'n frekwensietelling sekere items wat in die makrostruktuur van 'n bepaalde woordeboek behoort, kan uitsluit. Otto (1995: 220) wys op 'n verdere nadeel van gebondenheid aan 'n frekwensietelling by die bepaling van die makrostruktuur, naamlik "[i]n 'n hele paar definisies verskyn woorde wat nie as lemmas opgeneem is nie".[4] Dit kan veroorsaak dat die gebruiker in 'n leksikografiese doodloopstraat beland indien hy/sy 'n woord wat in 'n definisie gebruik word, wil naslaan.

Dit is dus duidelik dat die gebruik van frekwensietellings nie daartoe bydra dat 'n kernwoordeboek die ware (kognitiewe) woordeskat sonder meer kan weergee nie. 'n Frekwensietelling voorsien die leksikograaf hoogstens van die *aktiewe woordeskat*, waarvan nie alle items noodwendig tot die kernwoordeskat behoort nie.

4.1.4 'n Gekombineerde benadering

Hierdie stand van sake vra om 'n benadering tot die vasstel van die omvang van 'n kernwoordeskat wat meer verteenwoordigend van die ware kernwoordeskat is én wat semanties bevredigend is. So 'n benadering behoort die voordele wat veral die frekwensietellingbenadering bied, suksesvol te kan inkorporeer.

Een van die sentrale probleme van die frekwensietellingbenadering tot die afbakening van 'n makrostruktuur is die feit dat 'n frekwensietelling 'n grootliks amorfe versameling leksikale items as resultaat lewer. So 'n resultaat is semanties onbevredigend, aangesien die leksikon nie uit 'n versameling lukrake leksikale items bestaan nie, maar eerder uit bepaalde groeperinge van leksikale items in verskillende semantiese velde wat in bepaalde verhoudings tot ander leksikale items (in ander semantiese velde) staan (vgl. Cruse 1986; De Stadler 1989; Gouws 1989, 1989a; Lyons 1977, 1981; Zgusta 1971). De Stadler (1989: 92) verduidelik die konsep *semantiese veld* soos volg:

> Semantiese velde moet in verband gebring word met konseptuele velde. Die buitetalige werklikheid is in **konseptuele areas** georganiseer, soos die konseptuele area van kleur, die konseptuele area van ameublement, die konseptuele area van temperatuur, ens. Wanneer hierdie konseptuele area struktureel georden word deur 'n taalsisteem, word dit 'n **konseptuele veld**. Die taalsisteem wat aan die konseptuele veld struktuur gee, is die **semantiese veld**. […] Die leksikale items van die taal kan dus in bepaalde woordfamilies, die sg. semantiese velde, geklassifiseer word op grond van die konseptuele velde waarmee hulle verband hou. Hulle verdeel die semantiese ruimte op bepaalde maniere.

Die ontoereikendheid van 'n frekwensietelling as enigste maatstaf in die afbakening van 'n makrostruktuur word aan die hand van twee voorbeeldgroepe hieronder geïllustreer.

Veronderstel 'n frekwensietelling met 'n bepaalde drempelwaarde lewer die resultate *Maandag* en *Junie*, elk met 'n frekwensiewaarde hoër as die drempelwaarde (dus 'n positiewe resultaat). Dit sal daartoe lei dat dié twee items in die makrostruktuur van 'n woordeboek wat op dié drempelwaarde gebaseer is, opgeneem word. Vanuit 'n semantiese oogpunt hou dit egter ook in dat die bepaalde semantiese velde waarin hierdie items val, geaktiveer behoort te word vir opname in die makrostruktuur. Dat hierdie semantiese beginsel nie (konsekwent) in die makrostruktuur van óf BW óf KW geld nie, blyk uit die data in Tabelle 1 en 2:

SEMANTIESE VELD: **Dae van die week**	BW	KW	HAT
Maandag	+	+	+
Dinsdag	+	+	+
Woensdag	−	+	+
Donderdag	+	+	+

SEMANTIESE VELD: **Dae van die week**	BW	KW	HAT
Vrydag	−	+	+
Saterdag	+	+	+
Sondag	+	+	+

Tabel 1:   Die lemmatisering van lede van die semantiese veld "dae van die week"

(Legende: "+" = betrokke item is gelemmatiseer; "−" = betrokke item is nie gelemmatiseer nie)

SEMANTIESE VELD: **Maande van die jaar**	BW	KW	HAT
Januarie	+	−	+
Februarie	−	+	+
Maart	+	+	+
April	−	−	+
Mei	−	−	+
Junie	+	+	+

SEMANTIESE VELD: **Maande van die jaar**	BW	KW	HAT
Julie	+	+	+
Augustus	−	−	+
September	−	+	+
Oktober	−	−	+
November	−	+	+ [5]
Desember	+	−	+

Tabel 2:   Die lemmatisering van lede van die semantiese veld "maande van die jaar"

(Legende: "+" = betrokke item is gelemmatiseer; "−" = betrokke item is nie gelemmatiseer nie)

Dit word dus duidelik dat 'n frekwensietelling sonder meer nie 'n semanties verantwoordbare makrostruktuur lewer nie. By uitbreiding kan ook op samestellings wat stamme van die semantiese veld "dae van die week" bevat, gewys word. Vergelyk Tabel 3 hieronder:

Leksikale items	BW	KW	HAT
Maandagaand	−	+	+
Maandagmiddag	−	+	+
Maandagmôre	−	−	+
Maandagnag	−	+	+
Maandagoggend	−	+	+
Dinsdagaand	+	−	−
Dinsdagmiddag	+	−	−
Dinsdagmôre	+	−	−
Dinsdagnag	−	−	−
Dinsdagoggend	+	−	−
Woensdagaand	−	+	+
Woensdagmiddag	−	+	+
Woensdagmôre	−	−	+
Woensdagnag	−	+	−
Woensdagoggend	−	+	−

Leksikale items	BW	KW	HAT
Donderdagaand	+	−	−
Donderdagmiddag	+	−	−
Donderdagmôre	+	−	−
Donderdagnag	−	−	−
Donderdagoggend	+	−	−
Vrydagaand	−	+	+
Vrydagmiddag	−	+	+
Vrydagmôre	−	+	+
Vrydagnag	−	−	+
Vrydagoggend	−	−	+
Saterdagaand	+	−	+
Saterdagmiddag	+	−	+
Saterdagmôre	−	−	−
Saterdagnag	−	−	+
Saterdagoggend	−	−	−

Leksikale items	BW	KW	HAT		Leksikale items	BW	KW	HAT
Sondagaand	–	+	+		*Sondagnag*	–	+	+
Sondagmiddag	–	+	+		*Sondagoggend*	–	–	–
Sondagmôre	–	+	+					

Tabel 3: Die lemmatisering van samestellings met lede van die semantiese veld "dae van die week"

(Legende: "+" = betrokke item is gelemmatiseer; "–" = betrokke item is nie gelemmatiseer nie)

By die saamstel van die makrostruktuur behoort dus ook op die oorvleueling van semantiese velde gelet te word, soos in die geval van die semantiese velde "dae van die week" en "tye van die dag" hierbo. So 'n oorvleueling word waarskynlik ook in die vorm van 'n verteenwoordigende frekwensietellingresultaat onder die aandag van die leksikograaf gebring; die leksikograaf moet dié resultaat dan sien as 'n sinjaal van semantiese oorvleueling en potensiële soortgelyke vorme wat vir opname in die makrostruktuur kwalifiseer, en moet hom/haar nie tot die resultaatitem beperk by die saamstel van die makrostruktuur nie. (Die data in Tabelle 1, 2 en 3 suggereer dat dit vir BW, KW en HAT[4] nie die geval was nie.) Uiteraard kan 'n frekwensietelling ook meer as een leksikale item uit 'n bepaalde semantiese veld as positiewe resultaat lewer. Hoe meer leksikale items uit 'n bepaalde semantiese veld positiewe resultate van 'n frekwensietelling verteenwoordig, hoe sterker is die saak vir die insluit van die semantiese veld se verteenwoordigende leksikale items by die makrostruktuur.

Sommige semantiese velde val in hul geheel buite die bestek van 'n bepaalde makrostruktuur, terwyl gedeeltes van ander semantiese velde binne die bestek kan val. So byvoorbeeld kan met 'n redelike mate van sekerheid gekonstateer word dat die semantiese veld van "die wrywingsprosesse tydens die afvuur van 'n ballistiese missiel" in sy geheel buite die kognitiewe kernwoordeskat val, terwyl sommige items van die semantiese veld "meganiek van 'n motor" binne die kognitiewe kernwoordeskat val (bv. *brandstoftenk, rathefboom, uitlaatpyp*) en ander daarbuite (bv. *aardleiding, MacPhersonstut, ontstekingspoel*). Waar dit duidelik word dat slegs 'n deel van 'n bepaalde semantiese veld vir opname in die makrostruktuur van 'n woordeboek geld, behoort die leksikograaf pogings aan te wend om die verdeling so akkuraat moontlik te begrens deur opnames, ondersteunende frekwensietellings, e.s.m.

Semantiese velde bestaan egter nie slegs uit verenigbare kontraste (vgl. De Stadler 1989: 86-87) soos die data in Tabelle 1, 2 en 3 kan suggereer nie: Die sistematiese betekenisverbande wat tussen die leksikale items in 'n semantiese veld bestaan, sluit in sinonimie, teenoorgesteldheid, hiponimie, onversoenbaarheid, skale, metonimie en sintagmatiese verhoudings (De Stadler 1989: 98-99; 1991). Hierdie verbande moet uiteraard ook in die makrostruktuur neerslag vind deurdat leksikale items wat op dié wyse met 'n positiewe frekwensietellingresultaat op dieselfde kognitief-leksikale vlak in verband staan, gelemmatiseer word. 'n Enkele voorbeeld uit BW wat die nievoldoening aan hierdie vereiste as gevolg van gebondenheid aan frekwensietellingresultate illustreer, is

die feit dat die leksikale item *landelik* as lemma opgeneem is, terwyl die leksikale item *stedelik* nié deel van die makrostruktuur van BW uitmaak nie.

Die toetsing van hierdie beginsel bring 'n volgende probleem met betrekking tot gebondenheid aan frekwensietellings aan die lig. 'n Bepaalde leksikale item wat as 'n positiewe frekwensietellingresultaat gelewer word, word gelemmatiseer. Dié enigsins oppervlakkige prosedure bied egter geen insig in die semantiese struktuur van die kwalifiserende leksikale item nie. 'n Bepaalde item wat byvoorbeeld 'n baie hoë frekwensietelling lewer, kan nie een nie, maar talle semantiese waardes verteenwoordig, wat daarvan 'n polisemiese leksikale item maak. So kan die positiewe frekwensietellingresultaat *klank* die volgende semantiese waardes verteenwoordig (HAT⁴ 2000: 553):

(4)   (a) geluid wat die gehoor tref
        (b) eienskap wat aan 'n geluid 'n eie karakter gee
        (c) musiekgeluid
        (d) spraaktaalgeluid
        (e) geartikuleerde spraakgeluid
        (f) standpunt soos onder woorde gebring
        (g) drankasem; slegte reuk a.g.v. bederf

Indien die korpuskoteks nie in berekening gebring word nie, is dit onwaarskynlik dat die toepaslike leksikale items uiteindelik gelemmatiseer sal word. Indien positiewe frekwensietellingresultate in hul onderskeie semantiese velde geplaas word, kan hierdie probleem grootliks uitgeskakel word. Vergelyk die semantiese veld "stukke in die skaakspel," wat uit die volgende leksikale items bestaan (betekenisbeskrywings aangepas uit HAT⁴)⁶:

(5)   (a) *biskop* stuk wat by die begin van die skaakspel langs die koning of koningin staan
        (b) *kasteel* skaakstuk in die vorm van 'n ronde vestingtoring
        (c) *koning* vernaamste stuk in die skaakspel
        (d) *koningin* op een na die vernaamste stuk in die skaakspel
        (e) *perd* skaakstuk in die afbeelding van 'n perd
        (f) *pion* skaakstuk van die minste waarde

Die leksikale items in (5) is, met die uitsondering van *pion*, almal (volgens HAT⁴) polisemies van aard. So verteenwoordig die leksikale item *biskop* byvoorbeeld ook die volgende semantiese waarde:

(6)   priester van hoë rang in verskeie kerkgenootskappe

Dit hou in dat die lemmatisering van die leksikale item *biskop* voorsiening moet maak vir minstens die twee polisemiese waardes in (5)(a) en (6). Indien alleen op 'n frekwensietelling staatgemaak word, kan dit gebeur dat die polisemiese waarde in (6) verteenwoordig word, maar nié die polisemiese waarde in (5)(a) nie, wat die semantiese veld van "stukke in die skaakspel" onvolledig in die woordeboek verteenwoordig sal laat.

Hierdie betoog onderstreep die potensiaal van die werk van teoretici soos Wierzbicka (1985, 1996) en Moe (2003) as 'n semanties aanvaarbaarder benadering tot die saamstel (en evaluering) van die makrostruktuur van 'n woordeboek. Waar uitgebreide (en verteenwoordigende) korpora reeds bestaan (soos in die geval van Afrikaans), behels die benadering naamlik dat 'n frekwensietelling gebruik word, nie in die eerste plek om lemmata vir die makrostruktuur te identifiseer nie, maar wel om bepaalde semantiese velde wat vir opname in die makrostruktuur kwalifiseer, leksikaal te vul. Die seleksie van leksikale items vir opname in die makrostruktuur behoort dus nie slegs uit 'n amorfe lys positiewe frekwensietellingresultate gedoen te word nie, maar uit die semantiese velde wat met behulp van positiewe frekwensietellingresultate uit verteenwoordigende korpora geïdentifiseer (en/or geverifieer) is.

Hoewel die teorie van semantiese velde nie sonder kritiek staan nie (vgl. De Stadler 1989: 94-96), is dit gewis 'n instrument wat meer deur leksikograwe benut kan word, nie net in die opbou van die mikro- en mediostrukture van 'n woordeboek nie, maar ook in die samestelling van 'n verteenwoordigende makrostruktuur. Dit vra om 'n konseptuele benadering, en nie 'n bloot lineêre benadering nie, tot die vasstel van die makrostruktuur. Trouens, binne dié benadering het die samestelling van die makrostruktuur 'n beduidende invloed op die effektiwiteit van die mikro- en mediostrukture.

5.   Toegangstruktuur

5.1   Die eksterne toegangstruktuur

5.1.1   Die snel eksterne toegangstruktuur

HAT[4] en KW vertoon identiese snel eksterne toegangstrukture. Lemmatekens is naamlik alfabeties gerangskik, en boaan elke bladsy van die sentrale teks verskyn die eerste en laaste lemmateken wat op daardie bladsy opgeneem is — lopende opskrifte. Nóg HAT[4] nóg KW bevat duimindekse. Duimindekse is nie 'n totaal vreemde konsep aan die Afrikaanse leksikografie nie: Die 1978-publikasie *40 000 woorde* (Bresler 1978) bevat byvoorbeeld 'n duimindeks. Onlangse publikasies van taalnaslaanwerke soos *Afrikaanse woordelys en spelreëls* (Taalkommissie 2002) en *Skryf Afrikaans van A tot Z* (Müller 2003) sluit ook duimindekse as snel eksterne toeganstruktuurelement in. Die byvoeging van duimindekse by toekomstige publikasies van HAT en KW behoort ook oorweeg te word, tensy die finansiële implikasies van so 'n stap enige gedeelte van die teikengebruikersgroep sou vervreem (Louw 1999: 110).

Louw (1999: 111) maak melding van die *Collins COBUILD English Language Dictionary* wat van groeperings van swartgedrukte en leë diamantjies by elke lemma gebruik maak om die relatiewe gebruiksfrekwensie van die lemma aan te dui: 'n lemma met vier swartgedrukte diamantjies in die regterkantlyn verteenwoordig 'n leksikale item met 'n hoë gebruiksfrekwensie ('n waarskyn-

like kernwoordeskatitem), terwyl 'n lemma met drie leë en een swartgekleurde diamantjie 'n leksikale item verteenwoordig waarvan die gebruiksfrekwensie relatief laag is. Die implementering van so 'n toegangstruktuurelement sal uiteraard waarde tot 'n woordeboek soos KW (én HAT) toevoeg.

5.1.2   Die standaard eksterne toegangstruktuur

Terwyl die snel eksterne toegangstruktuur se funksie is om die gebruiker na die onmiddellike omgewing (die betrokke woordeboekbladsy) van die gesoekte lemma te lei, is die standaard eksterne toegangstruktuur verantwoordelik daarvoor om die gebruiker tot die finale keuse van die gesoekte makrostruktuurelement (Louw 1999: 111) te lei. Beide HAT[4] en KW is monotoeganklik, dit wil sê elke monotoeganklike woordeboek se standaard eksterne toegangstruktuur is kongruent aan sy makrostruktuur — gebaseer op alfabetiese ordening van die lemmatekens.

In beide HAT[4] en KW word tussen homonimiese lemmata onderskei deur 'n verhewe syfer onmiddellik regs van elke lemmateken te plaas, byvoorbeeld **skool¹**, **skool²**. Hierdie gebruik is reeds goed gevestig in Afrikaanse (en anderstalige) woordeboeke.

Beide HAT[4] en KW maak van sublemmatisering gebruik. Vervolgens word op hierdie aspek van die standaard eksterne toegangstruktuur uitgebrei.

5.1.2.1   Sublemmatisering as teksverdigtingstrategie

5.1.2.1.1 Die stand van sake in HAT[4] en KW

Een van die leksikograaf se grootste uitdagings is beperkte ruimte. Teksverdigtingstrategieë is noodsaaklik om soveel moontlik inligting in so min moontlik spasie oor te dra. Sublemmatisering is 'n teksverdigtingstrategie wat op die makrostruktuur betrekking het. Die uitgangspunt is dat lemmata wat morfologies-semanties nou aan mekaar verwant is, nie telkens van 'n volledige betekenisparafrase voorsien hoef te word nie. Afleidings- en kompositavorme van 'n bepaalde lemma wat vir opname in die makrostruktuur kwalifiseer, ondergaan dus beperkte leksikografiese bewerking deurdat hulle as onverklaarde lemmata hanteer word, wat aan die einde van die artikel van die lemma waarmee hulle morfologies-semanties verband hou, saamgegroepeer word. Die hoofvoorwaardes vir 'n item om as onverklaarde lemma hanteer te word, is dat daardie item semanties deursigtig moet wees (Gouws 1989: 77-81). Die proses waardeur items wat as onverklaarde lemmata kwalifiseer, gelemmatiseer word, word *sublemmatisering* genoem.

Indien meer as een sublemma ten opsigte van 'n bepaalde hooflemma opgeneem word, moet die leksikograaf besluit op 'n ordeningsbeginsel vir die sublemmata wat dwarsdeur die makrostruktuur gehandhaaf sal word ten einde die gebruiker in staat te stel om sy soektog na 'n bepaalde lemma so

ongekompliseerd moontlik te maak (Gouws 2002: 140). Daar bestaan twee basiese sublemmatiseringstrategieë, naamlik *nesting* en *nisting* (vgl. Gouws 2002: 134-135). Aan die hand van verteenwoordigende uittreksels[7] uit HAT[4] en KW word die sublemmatiseringstrategieë in dié woordeboeke bespreek en vergelyk.

     Vergelyk die volgende uittreksels, dié in (7) uit HAT[4] (2000: 1119) en dié in (8) uit KW (2000: 642):

(7)   **sui'ker** ⊠ s.nw. Soet, plantaardige stof uit die sap van verskillende plante verkry, bv. uit suikerriet en beet: *Suiker in jou tee, koffie gooi. Daar is te veel suiker in die nagereg. Beet-, druiwesuiker.*

       ⊠ ww. (gesuiker) Soet maak, bestrooi met suiker: *Vrugte eers suiker, dan kook.* [F. *sucre*, oorspr. Arab.-Pers.] **suiker: ~agtig, ~boer, ~fabriek, ~fabrikant, ~fabrikasie, ~handel(aar), ~ig, ~industrie, ~kweker, ~onderneming, ~premie, ~produsent, ~sap, ~smaak, ~verbruik.**

    sui'ker·ba·ron s.nw. Eienaar van groot suikerplase.

(8)   **twak** *s.nw.* (twakke)

       1 Twak is 'n geselstaalvorm van **tabak**: *Sterk twak rook.*

       2 Iemand wat 'n twak is, is 'n niksbeduidende vent: *Wie neem nou notisie van so 'n twak!*

       3 Twak is onsin, kaf: *Jy praat twak! Ag, twak met jou.*

       UITDR. *Iemand se twak is nat*, hy of sy het geen kans op sukses nie.

       twak'kerig. twak'praatjies. twak'voorstel.

    twee

       ⊠ *tw.* Twee is een plus een: *Gee my tog twee rand. Twee is slimmer as een. Twee-twee loop*, in pare.

       ⊠ *s.nw.* (tweë of twees)

       Twee is die syfer 2: *Net 'n twee vir jou toets kry.*

       twee'de. twee'daags. twee'handig. twee'honderd. twee'jarig. twee'koppig. twee'manskool. twee'sang. twee'sydig. twee'tal. twee'talig. twee'voud(ig).

    tweed s.nw.

       Tweed is 'n growwe wol waarvan veral manspakke gemaak word; ook, klere word hiervan gemaak: *Hierdie baadjie is van tweed en sal jare der jare hou.*

Binne 'n bepaalde groepering van sublemmata, soos onder **suiker** in (7) en **twak** en **twee** in (8), geld 'n streng alfabetiese ordening van die sublemmata; daar is dus nie sprake van interne alfabetiese afwyking nie.[8] Die sublemmata wat saamgegroepeer is, pas egter nie alfabeties in die groter makrostrukturele omgewing wat deur die daaropvolgende hooflemmata (**suikerbaron** in (7) en **tweed** in (8)) verteenwoordig word nie; daar is dus sprake van eksterne afwyking. Hierdie sublemmatiseringstrategie staan as tweedevlaknisting bekend (vgl. Gouws 2002: 136), en dit word in beide HAT[4] en KW toegepas.

     'n Belangrike verskil tussen HAT[4] en KW wat sublemmatisering betref, is dat HAT[4] van nistoegangslemmata gebruik maak en dat die sublemmata in 'n nis uitfasering na links ondergaan (vgl. (7)), terwyl elke nislemma in KW volledig weergegee word (vgl. (8)). Die prosedure van uitfasering na links van nis-

lemmata in HAT[4] verteenwoordig 'n hoër graad van teksverdigting as in KW (vgl. Gouws 2002: 137), aangesien die teikengebruiker van HAT[4] oor meer gevorderde taal- en naslaanvaardighede moet beskik om die betrokke lemmatekens suksesvol te rekonstrueer.

Nóg in HAT[4] nóg in KW word in die groepe sublemmata 'n onderskeid getref tussen sublemmata wat produkte van afleiding is en dié wat produkte van samestelling is. Die sublemmata **suikeragtig** en **suikerig** in (7), en **twakkerig** en **tweede** in (8) verteenwoordig produkte van afleiding, terwyl die res produkte van samestelling is. Volgens Gouws (2002: 136) motiveer hierdie morfologiese verskil die gebruik van die verdelingsmerker ";" tussen die twee klasse sublemmata, wat tweedevlaknesting tot gevolg sou hê. Die sublemmata onder die hooflemma **suiker** sou dan soos volg in HAT[4] binne 'n tweedevlaknes aangebied word:

(9)   **suiker: ~agtig, ~ig; ~boer, ~fabriek, ~fabrikant, ~fabrikasie, ~handel(aar), ~industrie, ~kweker, ~onderneming, ~premie, ~produsent, ~sap, ~smaak, ~verbruik.**

Soos uit (7) en (8) blyk, word hierdie onderskeid nie getref nie: In HAT[4] word die verdelingsmerker "," en in KW die verdelingsmerker "." deurgaans gebruik om nislemmata van mekaar te onderskei. Tweedevlaknesting word in *Nasionale Woordeboek* (De Villiers et al. 1987) toegepas, en in *Verklarende Afrikaanse Woordeboek* (Labuschagne en Eksteen 1993) word nisting onreëlmatig aangewend. Dit wil dus voorkom asof daar in die Afrikaanse verklarende leksikografie geen reëlmatige tradisie ten opsigte van nesting en nisting bestaan nie.

Waar 'n hooflemma polisemies is, behoort sublemmata ten opsigte van die hooflemma se poliseme gegroepeer en van 'n inskrywing voorsien te word wat die onderskeie groepe aan die relevante poliseme verbind. Vergelyk die volgende uittreksel uit *Nasionale Woordeboek* (De Villiers et al. 1987: 525):

(10)   **'toevoer I** s.nw. 1. *vloei (van voorraad); voorsiening.* Die – van water, olie, petrol, goedere, e.s.m. reël, afsny, stopsit. **–** van hout in 'n saagmeule, gerwe in 'n dorsmasjien. Bron van –. 2. *voorraad wat voorsien word.* Die – neem af. 3. (rekenaar) *vloei van data na 'n rekenaar.* **toevoerbuis, -kanaal, -linie, -lyn, -pomp, -pyp, -vat** (by 1 en 2); **-apparaat, -formaat, -medium, -program, -toerusting** (by 3).

Ook hierdie toepassing van tweedevlaknesting kom nie in HAT[4] of KW voor nie.

5.1.2.1.2   Nadele van sublemmatisering in verklarende woordeboeke

Daar kan op minstens twee nadele van die praktyk van sublemmatisering gewys word. Hierdie nadele ontstaan as gevolg van die zerobewerking wat sublemmata ondergaan (uitgesonderd die aanduiding van hoofklem).

Die eerste nadeel het betrekking op die oordra van grammatiese inligting. Bepaalde datakategorieë word tans nie in die bewerking van onverklaarde lemmata verteenwoordig nie, byvoorbeeld woordklas, kollokasies en ander sintagmas. Vir die gemiddelde volwasse moedertaalgebruiker van 'n standaardwoordeboek behoort hierdie stand van sake nie 'n probleem te wees nie, aangesien die primêre kriterium van semantiese deursigtigheid sublemmatisering bepaal. Die afwesigheid van hierdie datakategorieë in 'n aanleerderwoordeboek kan die aanleerder egter tot nadeel strek; daar behoort dus na gelang van die teikengebruikersprofiel deeglik besin te word oor die toepassing al dan nie van sublemmatisering in spesiale woordeboeke. Vir 'n standaardwoordeboek soos HAT[4], wat buitendien 'n uitgebreide makrostruktuur in beperkte ruimte moet aanbied, is die praktyk van sublemmatisering 'n regverdigbare teksverdigtingstrategie, mits dit sistematies en konsekwent aangewend word en duidelik in die toeligting beskryf word.

Die tweede probleem is egter 'n enigsins groter bedreiging vir die taalkundige kwaliteit van die woordeboek. Volgens De Stadler (1989: 75) vind die leksikale betekenisse van leksikale items neerslag in die betekenisbetrekkinge wat tussen leksikale items bestaan. "Die strukturering van die leksikon hou dan verband met die mens se strukturering van sy kennis van die buitetalige werklikheid: die leksikale items vorm klasse wat verband hou met konseptuele velde binne hierdie kennisstruktuur, en die leksikale item verkry sy betekenis via hierdie verhouding met die buitetalige werklikheid" (De Stadler 1989: 75-76). Die aanduiding van leksikale betrekkinge wat tussen leksikale items geld, is dus 'n belangrike opdrag van die verklarende leksikograaf. Hierdie opdrag kom in die makro-, mikro- en mediostruktuur tot uitvoering deur die leksikografiese markering van leksikale betrekkinge soos homonimie, hiponimie, sinonimie, teenoorgesteldheid en sintagmatiese leksikale betrekkinge met behulp van 'n verskeidenheid inskrywingstipes. Dit volg dus dat by woordeboekartikels waar 'n mikro- en mediostruktuur ontbreek, soos by sublemmata in verklarende woordeboeke, elemente van hierdie belangrike semantiese inligting verlore kan gaan.

Hierdie probleem word geïllustreer aan die hand van die leksikale betrekking van sinonimie. Leksikografiese konvensie bepaal dat die lede van 'n sinoniemparadigma wat gelemmatiseer word, met mekaar in verband gebring moet word deur die lid met die hoogste gebruiksfrekwensie van 'n volledige betekenisparafrase te voorsien en die ander lid of lede van die paradigma van 'n sinoniem te voorsien wat die gebruiker per kruisverwysing na die verklaarde lid verwys (Gouws 1989: 145). Daarby moet 'n verwysing na elke lemma wat 'n lid van die betrokke sinoniemparadigma is, by die artikel van die frekwenste lemma ingesluit word (Gouws 1989: 146). In sommige gevalle verhoed die praktyk van sublemmatisering egter die implisiete en eksplisiete aanduiding van hierdie leksikale betrekking. Vergelyk die volgende uittreksels uit HAT[4] (2000: 31, 989):

(11)   **af'tas** ww. (afgetas) **1** Orals bevoel. **2** Met strale, lig e.d. oor 'n gebied beweeg om dit (teen gevaar, e.s.m.) te ondersoek. **aftas: ~ter, ~ting.**

(12)   **skan·deer'** ww. (geskandeer) **1** *(rek.)* Handeling waardeur 'n teks, grafika of fotomateriaal deur 'n apparaat, die skandeerder, in elektroniese vorm vasgelê word vir berging in 'n rekenaar: *As jy hierdie teks kan skandeer, hoef jy nie die hele ding in te tik nie.* **2** In versvoete verdeel: *'n Gedig skandeer.* [L. *scandere* klim] **skande'ring.**

   skan·deer'der s.nw. **1** *(rek.)* Rekenaarapparaat waarmee tekste, grafika, foto's, ens. geskandeer kan word; aftaster: *Is daar 'n skandeerder aan jou rekenaar gekoppel?* **2** *(med.)* Apparaat waarmee die intensiteit van bestraling, ultrasoniese uitstraling, ens. in die liggaam vir diagnostiese doeleindes waargeneem kan word; aftaster.

By elke poliseem van die lemma **skandeerder** is daar 'n verwysing na die lemma **aftaster**, wat volgens leksikografiese konvensie daarop dui dat die item *skandeerder* telkens die frekwentste lid van die sinoniemparadigma ₛₚ[*aftaster, skandeerder*]ₛₚ is.[9] Die feit dat die item *aftaster* slegs as sublemma verskyn en as sodanig nie leksikografies bewerk word nie, lei daartoe dat volledige kruisverwysing nie kan plaasvind nie. Die woordeboekgebruiker wat die nislemma **aftaster** onder die hooflemma **aftas** vind, word nie ingelig dat dié item 'n meer frekwente sinoniem, *skandeerder*, het nie. Hierin faal die verklarende woordeboek in sy opdrag as houer van taalkundige inligting.

   Die probleem kan op verskillende maniere opgelos word. Die eerste oplossing is om die rekenaarvaktaalbetekenis van die verbale leksikale item *aftas* in die artikel van die lemma **aftas** op te neem en kruisverwysing tussen dié leksikale item en die betrokke poliseem van die verbumlemma **skandeer** te bewerkstellig. Hierdie prosedure sal egter nog nie die betekenisverband tussen die leksikale items *aftaster* en *skandeerder* ekspliseer nie. 'n Verdere prosedure van tweedevlaknesting van sublemmata volgens poliseme waarop hulle betrekking het (vgl. (10)), sal die sublemma **aftaster** in indirekte verband bring met die leksikale item *skandeerder*, waarmee dit in 'n verhouding van sinonimie staan. 'n Optimale sinjalering van die leksikale betrekking wat tussen die twee items bestaan, sal 'n beperkte mikrostrukturele bewerking van sublemmata vereis. 'n Nis soos die volgende, waarin elke individuele sublemma 'n bewerkingseenheid verteenwoordig en wat ingelui word deur 'n verwyderde nistoegangslemma (vgl. Gouws 2002: 137), sou die produk van so 'n bewerking kon wees:

(13)   **aftas: ~ter** skandeerder**, ~ting** skandering.

Tweedevlaknesting kan 'n komplekser artikelstruktuur as resultaat toon:

(14)   **aftas: ~ting** skandering; **~ter** (by 3) skandeerder (bet. 1).

Die implikasie in (14) is dat die neslemma **aftasting** op al die poliseme van die hooflemma **aftas** slaan, terwyl die neslemma **aftaster** slegs op die derde (byge-

voegde) poliseem betrekking het,[10] en slegs met betrekking tot die eerste poli-
seem van die lemma **skandeerder** in 'n verhouding van sinonimie staan. Indien
hierdie prosedure leksikografiese praktyk word, word 'n bykomende voor-
waarde vir sublemmatisering toegevoeg: 'n sublemma mag nie die frekwentste
lid van 'n sinoniemparadigma verteenwoordig nie; sodanige item moet as hoof-
lemma volledig bewerk word. Die alternatief is dat sublemmata ook volledige
leksikografiese bewerking moet kan ondergaan. Die toepassing van sodanige
prosedure en die invloed op nes- en nisstrukture val buite die bestek van
hierdie artikel en dus word nie verder hierop ingegaan nie. Daar word volstaan
met die standpunt dat die praktyk van sublemmatisering in sy huidige vorm in
verklarende woordeboeke belangrike taalkundige inligting kan verberg.

5.1.2.1.3 Evaluering van sublemmatisering in HAT[4] en KW

Die teksverdigtingstrategie van sublemmatisering in HAT[4] is verdedigbaar op
grond van die feit dat die teikengebruiker van 'n standaardwoordeboek soos
HAT gewoonlik 'n (volwasse) moedertaalspreker is wat oor meer gevorderde
taalvaardighede as byvoorbeeld die tweedetaalspreker of laerskoolmoedertaal-
spreker beskik. Hierdie feit laat die leksikograaf toe om 'n hoër graad van teks-
verdigting in die standaardwoordeboek toe te pas, wat die gebruik van toe-
gangslemmata en die uitfasering van sublemmata na links insluit. Indien KW
as skoolwoordeboek geld, is die weergawe van sublemmata deur ongekonden-
seerde lemmatekens geregverdig. KW vertoon oor die algemeen 'n laer graad
van teksverdigting as HAT[4] omdat die waarskynlike gebruiker van KW (nog)
nie 'n volkome gesofistikeerde taalgebruiker is nie. Die risiko bestaan dat ver-
skynsels soos nistoegangslemmata en die kondensering van sublemmata na
links die teikengebruiker van 'n skool- of aanleerderwoordeboek se soekproses
onnodig kan kompliseer. In BW word dié prosedures byvoorbeeld ook nie
aangewend nie omdat die woordeboek op 'n taalaanleerder gerig is.

In 'n standaardwoordeboek, waarvan die gebruiker vollediger taalkun-
dige inligting kan verwag, behoort die prosedure van tweedevlaknesting eer-
der as tweedevlaknisting tydens sublemmatisering toegepas te word, waar die
groeperings binne 'n nes gebaseer is op die morfosemantiese aard van die sub-
lemmata, dit wil sê op die tweedeling tussen afleidings en samestellings (vgl.
(9)). Daarenteen is die waarskynlike gebruiker van 'n spesiale woordeboek soos
KW nie so gesteld op fyner taalkundige detail van dié aard nie, en kan tweede-
of eerstevlaknisting na gelang van die tipiese teikengebruiker se vermoëns en
behoeftes toegepas word. Die prosedure van tweedevlaknisting geld dan ook
tereg in KW, soos in (8) aangetoon.

In 'n standaardwoordeboek behoort egter ten minste 'n beperkte leksiko-
grafiese bewerking op sublemmata te volg wat nie op al die polisemiese waar-
des van die betrokke hooflemma slaan nie (vgl. (10)). Die afwesigheid van hier-
die bewerkings in HAT[4] behoort as 'n tekortkoming beskou te word. Afgesien
hiervan behoort in woordeboeke soos KW ook oorweging geskenk te word aan

die beperkte leksikografiese bewerking van sublemmata waar hulle relasies met ander leksikale items dit kan vereis, veral indien die vereiste sou geld as dié lemmata as hooflemmata in die woordeboek opgeneem sou word. Hoewel 'n beperkte leksikografiese bewerking soos in (13) voorgestel belangrike taal-kundige hulp aan die (taalaan)leerder kan bied, verteenwoordig sodanige be-werking ook die aanbod van meer gesofistikeerde semantiese inligting binne 'n teksverdigtingstrategie; dus behoort die insluit van hierdie prosedure in stan-daardwoordeboeke ook oorweeg te kan word.

Hoewel in HAT[4] se metateks na die toepassing van sublemmatisering verwys word, word geen bevredigende leksikografiese verantwoording aange-bied ten opsigte van die stelsel van sublemmatisering wat gevolg word nie. In KW se "Voorwoord" word geen melding van onverklaarde lemmata of sublem-matisering gemaak nie. Indien die standaard eksterne toegangstruktuur nie in die toeligting verklaar word nie, "this aspect of the dictionary's outer access structure loses almost all its value" (Louw 1999: 112).

Ter opsomming: Die teksverdigtingstrategie van tweedevlaknisting wat in HAT[4] geld, is morfosemanties onvoldoende en behoort deur die prosedure van tweedevlaknesting, gegrond op morfologiese struktuur, vervang te word; die prosedure van tweedevlaknisting in KW is bevredigend gegee die waarskyn-like teikengebruikersprofiel. In beide woordeboeke behoort sublemmata van beperkte leksikografiese bewerkings voorsien te word wat hulle betrekking op polisemiese waardes van die hooflemmata ekspliseer en wat hulle semantiese relasies met ander leksikale items uitlig. Die praktyk van sublemmatisering behoort volledig in die metatekste van dié woordeboeke verantwoord te word.

5.2   Die interne toegangstruktuur

Die doel van die interne toegangstruktuur is om die gebruiker te lei na die spe-sifieke datakategorie(ë) waarna hy/sy in 'n woordeboekartikel soek. Waar die eksterne toegangstruktuur die gebruiker dus bring tot by die woordeboekarti-kel wat die antwoord op sy/haar "vraag" bevat, navigeer die interne toegang-struktuur die gebruiker bínne die artikel tot by die spesifieke "antwoord" — dit gaan dus by laasgenoemde oor "die manier waarop inligting in die mikrostruk-tuur geplaas en aangebied word" (Stark 2001: 31).

5.2.1   Die snel interne toegangstruktuur

Die snel interne toegangstruktuur het hoofsaaklik struktuurmerkers in bestek, dit wil sê tipografiese elemente binne die woordeboekartikel wat die grense tussen verskillende datakategorieë sowel as die plasing en bestek van verskil-lende soeksones aandui (Gouws 1996: 23). Sodanige merkers is óf vreemde merkers óf merkers wat aan die gebruiker bekend is, maar wat buite hul nor-male konteks gebruik word.

In HAT[4] word onder meer die volgende struktuurmerkers vir die begren-sing van die genoemde datakategorieë aangewend (vgl. Wiegand 1990: 34):

(15)   (a) Die artikelgleuf vir uitspraakinligting word deur die saamgroeperingsmerker "( )" begrens en deur die nietipografiese struktuurmerker "uitspr." ingelui; vergelyk die lemma **arpeggio** se bewerking: "**ar·peg'gio** (uitspr. *ar-ped'zjoo*) […]" (HAT[4] 2000: 53).

(b) Leksikografiese etikette word deur die saamgroeperingsmerker "( )" in kursiefdruk (wat 'n tipografiese struktuurmerker verteenwoordig) gemarkeer; vergelyk die lemma **arteriosklerose** se bewerking: "**ar·te·ri·o·skle·ro'se** s.nw. *(med.)* […]" (HAT[4] 2000: 53).

(c) Lemmatiese meerfunksionaliteit word deur die nietipografiese struktuurmerker "⊠" voor elke woordklasaanduiding van die betrokke lemma aangedui.

(d) Waar 'n leksikale betrekking naas 'n betekenisparafrase in die semantiese kommentaar op 'n lemma aangebied word, word die aangawe van die betrekking van die betekenisparafrase geskei deur die verdelingsmerker ";" — vergelyk die lemma **baannommer**, wat soos volg bewerk is: "Wedloop op 'n baan; teenoor *veldnommer*" (HAT[4] 2000: 59).

(e) Waar 'n lemma slegs deur sinonieme verklaar word, word die sinonieme van mekaar geskei deur die verdelingsmerker "," — vergelyk die bewerking van die lemma **begelei**: "Saamgaan, vergesel: *Iemand na haar huis begelei. 'n Begeleidende brief*" (HAT[4] 2000: 71).

(f) Waar meer as een betekenisparafrase vir 'n bepaalde poliseem van 'n lemma aangebied word, word die parafrases van mekaar geskei deur die verdelingsmerker ";" — vergelyk die lemma **konkreet**[1], wat soos volg verklaar word: "Werklik bestaande; waarneembaar; wat aangeraak, gesien, gehoor, ens. kan word; nie abstrak nie" (HAT[4] 2000: 593).

(g) Voorbeeldmateriaal word van die betekenisparafrase van 'n lemma geskei deur die verdelingsmerker ":" — vergelyk die lemma **begeerte**, wat soos volg bewerk is: "Sterk wens, verlange: *'n Sterk begeerte koester om sy land nog eens te besoek*" (HAT[4] 2000: 71).

(h) Sitate in die voorbeeldmateriaal van HAT[4] word deur die nietipografiese struktuurmerker "▽" voorafgegaan; vergelyk die bewerking van die lemma **klam**: "**1** Vogtig, natterig: *Die plek is klam, selfs in die somer. ▽Die grond is klam en die graaf gly daardeur as ek dit wegtrap* (C.J.M. Nienaber [...]" (HAT[4] 2000: 552).

(i) Die groepering van vaste uitdrukkings onder 'n bepaalde lemma word deur die nietipografiese struktuurmerker "UITDR.:" ingelui, soos by die lemma **hot**[1]: "Links, na die linkerkant toe; teenoor *haar*: *Hierdie perd trek hot. Draai die moer hot om*. UITDR.: *Hot en **haar**, (hot en bles),* rond en bont, orals. *Hot is hot en **haar** is haar*, dit is soos dit is; daar is geen twyfel nie […]" (HAT[4] 2000: 413).

Van die nege struktuurmerkers wat hierbo ten opsigte van HAT[4] lukraak geëkstraheer is, word slegs een ((15)(c)) leksikografies verantwoord in die toeligting. In die geannoteerde skematiese voorstelling van die tipiese woordeboekartikel onder die opskrif "Artikeluitleg" (HAT[4] 2000: xx) word slegs drie ((15)(c), (h) en (i)) verteenwoordig, hoewel alle datakategorieë in hulle teksvorm gemerk en geïdentifiseer word. Die soeksones in die woordeboekartikel word telkens uitgelig, maar nie die merkers wat die soeksones van mekaar (onder)skei nie.

KW vaar beter in sy "Voorwoord" deur die aangawe van die meeste struktuurmerkers wat in dié woordeboek gebruik word. 'n Skematiese voorstelling van 'n verteenwoordigende artikel (soortgelyk aan die een wat onder die opskrif "Artikeluitleg" in HAT⁴ aangebied word) waarin die verskillende elemente waaruit die woordeboekartikels bestaan, geïdentifiseer en verduidelik word, ontbreek egter. So 'n voorstelling sou die samestelling van die woordeboekartikels (en die denotasies van die onderskeie struktuurmerkers) vir die teikengebruiker toegankliker gemaak het.

Indien KW as skoolwoordeboek geld, sou aanbeveel kon word dat elke soeksone so ver prakties moontlik deur 'n unieke struktuurmerker afgebaken word. Sodanige struktuurmerkers sou 'n tweeledige voordeel inhou: eerstens is hulle eenduidig en kan die gebruiker nie verwar nie; tweedens neem die jong gebruiker deeglik van die onderskeie inligtingtipes wat hy/sy in die woordeboek kan terugvind, kennis — kennis wat later met vrug by die gebruik van standaardwoordeboeke soos HAT⁴ aangewend kan word.

5.2.2   Die standaard interne toegangstruktuur

Die standaard interne toegangstruktuur verwys na organisering van die boustene van 'n woordeboekartikel, dit wil sê die plasing van die onderskeie artikelgleuwe. Ten opsigte van die interne ordening van die woordeboekartikel oorvleuel die standaard interne toegangstruktuur met die artikelstruktuur (Louw 1999: 117). Vir die doel van hierdie evaluerende vergelyking sal nie op artikelstruktuur (vgl. o.m. Wiegand 1990) ingegaan word nie. Wat die plasing van ooreenstemmende datakategorieë in HAT⁴ en KW betref, is die standaard interne toegangstruktuur in hierdie twee woordeboeke byna identies. KW se standaard interne toegangstruktuur verskil van dié van HAT⁴ slegs ten opsigte van die aangawe van etimologiese inligting. Terwyl HAT⁴ wel by sommige lemmata etimologiese inligting voorsien, bevat KW tereg geen etimologiese inligting nie, gegee die waarskynlikste doel van dié woordeboek (vgl. 6.6).

Die grootliks identiese ordening van mikrostruktuurelemente wat deur HAT⁴ en KW gedeel word, verseker 'n gladde oorskakeling van die een na die ander wanneer die teikengebruiker van KW (die leerder) eventueel oorbeweeg na die teikengebruikersgroep van HAT⁴ (die volwasse taalgebuiker).

5.2.3   Mikroargitektuur

Beide HAT⁴ en KW vertoon 'n mikroargitektuur. In woordeboeke sonder 'n mikroargitektuur word die artikels as teksblokke (vergelykbaar met 'n paragraaf gewone teks) aangebied. Volgens Stark (2001: 32) dra die "konvensionele, lineêre aanbieding van die inligtingstipes" in woordeboeke sonder 'n mikroargitektuur daartoe by dat dit moeiliker is "vir die gebruiker om inligting te herwin en die herwinningsproses neem ook langer". HAT⁴ vertoon 'n mikroargitektuur deurdat die subsemantiese kommentaar op 'n bepaalde woordklas van

'n meerfunksionele lemmateken telkens op 'n nuwe reël begin en dus 'n aparte teksblok binne die artikel vorm. Hierdie mikroargitektuur kan as 'n *beperkte mikroargitektuur* beskryf word, in vergelyking met die mikroargitektuur wat KW vertoon. Vergelyk die artikels van die lemma **hap** uit HAT[4] ((16)(a)) en KW ((16)(b)):

(16)   (a) **hap** ⊠ ww. (gehap) Die geopende mond (bek) skielik sluit om daarmee iets tussen die tande te vat; met die tande gryp; skielik, vinnig na iets byt: *Hap 'n stuk uit my appel. Die hond het na haar gehap.*
⊠ s.nw. (-pe; -pie) **1** Handeling van hap: *Die leeu het die bokkie se been met een hap afgebyt*. **2** Soveel as wat op 'n slag afgebyt kan word: *Vat 'n groot hap*. **3** Leemte waar 'n deel uit iets verwyder is: *Die onvoorsiene uitgawe het 'n hap uit my geld gemaak. As veroweraar 'n hap uit 'n ander land vir jou annekseer.* UITDR.: *In (met) **een** hap*, in een slag. **hap'beweging**.

      (b) **hap**
⊠ *ww.* (gehap)
As 'n mens of dier iets hap, byt hulle dit: *Ek hap 'n stuk uit die appel. Na iets hap*, dit probeer byt, bv. *die hond hap na die kind*.
⊠ *s.nw.* (happe; happie)
1 'n Hap is die daad van hap: *Met een hap alles opeet*.
2 'n Hap is ook die hoeveelheid wat 'n mens op een keer hap: *'n Groot hap uit die appel*.
hap'perig.

KW vertoon 'n *uitgebreide mikroargitektuur*: Ten opsigte van kommentaar op die meerfunksionaliteit van die lemmateken, elke polisemiese waarde se volledige bewerking, kruisverwysings en die aanbied van sublemmata vorm elke kategorie 'n aparte teksblok binne die artikel. Hierteenoor vertoon HAT[3] geen mikroargitektuur nie:

(17)   **hap** ☐ ww. (gehap) Die geopende mond (bek) skielik sluit om daarmee iets tussen die tande te vat; met die tande gryp; skielik, vinnig na iets byt: *Hap 'n stuk uit my appel. Die hond het na hom gehap*. ☐ s.nw. (-pe; -pie) **1** Handeling van hap: *Die leeu het sy arm met een hap afgebyt*. **2** Soveel as wat op 'n slag afgebyt kan word: *Vat 'n groot hap*. **3** Leemte waar 'n deel uit iets verwyder is: *Die onvoorsiene uitgawe het 'n hap in my geld gemaak. As veroweraar 'n hap uit 'n land vir jou annekseer*. UITDR.: *In (met) **een** hap*, in een slag. **hap'beweging.**

Die volledige artikel van die lemma **hap** word in HAT[3] in een teksblok aangebied. In HAT[3] is daar dus van 'n hoër graad van teksverdigting sprake as in HAT[4]. Die beperkte mikroargitektuur in HAT[4] maak dié woordeboek se inligting 'n graad toegankliker vir die gemiddelde gebruiker. Die uitgebreide mikroargitektuur wat in KW geld, verteenwoordig 'n laer graad van teksverdigting as wat in HAT[4] die geval is. Indien KW as skoolwoordeboek aangewend word, is die uitgebreide mikroargitektuur goeie leksikografie, aangesien dit die teikengebruiker (leerder) help om die inligting in die woordeboek maklik te ontsluit. Woordeboekonderrig word ook hierdeur gefasiliteer, wat tot gevolg sal

hê dat skoolverlaters oor goed ontwikkelde woordeboekvaardighede kan be-skik teen die tyd dat HAT[4] hul standaardnaslaanwerk vir taalkundige inligting word.

Die ordening van elemente binne die artikels van HAT[4] en KW verteen-woordig 'n geïntegreerde mikrostruktuur deurdat alle inligting wat op 'n be-paalde polisemiese waarde betrekking het, aangebied word alvorens die vol-gende polisemiese waarde aan bod kom (vgl. weer (16) en (17)). Vir 'n woorde-boek soos KW is 'n geïntegreerde mikrostruktuur die aangewese struktuur, aan-gesien die naslaan- en semantiese koördinasievaardighede van die teikenge-bruiker moontlik nie ontwikkeld genoeg kan wees om aan die meer komplekse verbindinge wat 'n ongeïntegreerde mikrostruktuur vereis, te beantwoord nie.

Ter opsomming: Die toegangstrukture van HAT[4] en KW kom in 'n groot mate ooreen. KW vertoon oor die algemeen tereg 'n laer graad van teksverdig-ting as HAT[4], wat veral in die interne toegangstruktuur gemanifesteer word en as sodanig tot groter gebruikersvriendelikheid lei. Indien KW as skoolwoorde-boek geld, behoort die oorgang van KW na HAT[4] vir die gebruiker glad te ver-loop.

6.   Die mikrostruktuur

Carstens (1995) en Van der Merwe (1994) lewer in hulle resensieartikels van HAT[3] ook op die mikrostruktuur kommentaar. In die "Voorwoord by die Vier-de Uitgawe" in HAT[4] (2000: v) word gemeld dat "die vierde uitgawe van HAT nie 'n ingrypende hersiening is nie". Hierdie opmerking hou in dat daar ook nie (ingrypend) aan die mikrostruktuur verander is nie. Die volgende kommentaar op die mikrostruktuur sal dus beperk word tot die ooreenkomste en verskille tussen HAT[4] en KW, en tot enkele aspekte van die mikrostruktuur wat as sodanig aandag verdien. Alle aspekte van die mikrostruktuur sal dus nie (in detail) bespreek word nie. Vergelyk Carstens (1995) en Van der Merwe (1994) vir bykomende besprekings.

6.1   Uitspraakinligting ten opsigte van die lemma

Beyer (2000: 285-287) gee 'n oorsig van die uiteenlopende opinies van (meta)-leksikograwe wat die aanbied van uitspraakinligting betref, en kom tot die gevolgtrekking dat Gouws (1989: 265) die felste kritiek op die aanbod van hier-die datakategorie lewer wanneer hy beweer dat dit nie met genoeg sorg en uit-voerigheid in Afrikaanse woordeboeke hanteer word nie. HAT[4] en KW is geen uitsonderings nie.

6.1.1   Aanduiding van hoofklem

Volgens Gouws (1989: 266) is hoofklemaanduiding in Afrikaanse woordeboeke meestal op die lemmateken gerig: "Uitspraakinligting word […] ten opsigte van

die geskrewe vorm van die lemma gegee. Alhoewel dit in die leksikografiese praktyk van waarde is, is dié werkswyse taalkundig nie houdbaar nie, aangesien dit 'n vermenging van twee tekenstelsels verteenwoordig", naamlik die ortografiese tekenstelsel en die fonologiese tekenstelsel. Hierdie vermenging kom ook in HAT[4] en KW voor, wat identiese stelsels van hoofklemaanduiding vertoon, naamlik die plasing van 'n aksentteken ná die sillabe waarop die hoofklem val:

(18)   **hang'kas** […]
         rond'te […]

Hierdie gebruik word leksikografies in HAT[4] en KW se metatekste verant-woord.[11] In die "Voorwoord by die Derde Uitgawe" van HAT[4] (p. x) word ook die keuse van die plasing van die aksentteken ná die betrokke sillabe verant-woord: "Navraag het […] bewys dat die gebruiker die algemener gebruik van die aanduiding ná die beklemtoonde sillabe verkies." Hierdie stelling gee blyke van 'n poging om die gebruiker se behoeftes in ag te neem by die aanbied van taalkundige inligting, en dit toon goeie leksikografiese praktyk, mits die "na-vraag" onder 'n verteenwoordigende monster van die teikengebruikersgroep gedoen is. Dit beteken egter nie dat die huidige formaat van die aanbod van uitspraakinligting noodwendig in alle gevalle gebruikersvriendelik is nie, veral ten opsigte van die moontlike teikengebruiker van KW.

6.1.2 Transkripsie

In gevalle waar die samestellers van HAT[4] oordeel dat 'n lemma uitspraak-probleme kan oplewer, word sodanige lemma deur middel van ortografiese herskrywing getranskribeer. Vergelyk die uitspraakleiding in die volgende HAT[4]-artikels:

(19)   (a) **chif·fon'** (uitspr. *sjif-fon'*) […]
         (b) **Chris'ten** (uitspr. *ch* = *k* of *g*) […]
         (c) **com'me il faut'** (uitspr. *kom'-miel fo'*) […]

In 19(a) en (c) word volledige ortografiese herskrywing aangebied, terwyl (19)b 'n voorbeeld van gedeeltelike ortografiese herskrywing is. Gouws (1989: 268) lewer soos volg kommentaar op ortografiese herskrywing as transkripsietipe:

> Ondanks die voordele van eenvoud en toeganklikheid wat 'n ortografiese her-skrywing as transkripsietipe inhou, bly 'n fonetiese transkripsie uit 'n taalkun-dige oogpunt die suksesvolste manier van uitspraakleiding. Dit skep ook die ideale terrein vir die toepassing van ander uitspraakinligtingstipes soos klem-aanduiding en sillabeverdeling.

Die argument dat fonetiese transkripsies te moeilik verstaanbaar vir die gemid-delde gebruiker is, is ongeldig. Volgens Louw (1997: 85) sou die aanpassing wat die aanleer of gebruik van die fonetiese alfabet van die gemiddelde gebrui-

ker vereis, nie veel groter wees as wat die aanleer en gebruik van byvoorbeeld *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Dictionary (Grobbelaar 1987)* se sleutels sal verg nie. Indien die gemiddelde gebruiker buitendien woordeboekopleiding ontvang het aan die hand van 'n skoolwoordeboek wat hierdie stelsel gebruik, sal hy/sy min moeilikheid met HAT ondervind.

6.2   Woordklasaanduiding van die lemma

Anders as sy voorgangers, dui HAT[4] die woordklas van elke leksikale lemma byna konsekwent aan. Die enkele lemmata wat nie van hierdie inligting voorsien is nie, is waarskynlik per abuis oorgesien. Hierdie praktyk vind sy eggo in KW, waarin die aanbod van hierdie datakategorie essensieel is. Hierdie mate van konsekwentheid is 'n prestasie in die Afrikaanse leksikografie; daar word derhalwe by hierdie opmerking volstaan sonder om op die bewerking van subleksikale en multileksikale items ten opsigte van leksikale kategorie (vgl. hieroor Gouws 1989: 213-226) in te gaan.

6.3   Morfologiese inligting oor die lemma

HAT[4] is steeds inkonsekwent met die aanbied van morfologiese inligting oor lemmata. Vergelyk die volgende uittreksels uit HAT[4]:

(20)   (a) **jan·blom'** s.nw. (-me, -s) Donderpadda.
        (b) **mid'del·voor** Middelste voorspeler in bv. hokkie of sokker […]

(21)   (a) **glad** b.nw. en bw. (-der, -ste) **1** Met 'n oppervlak […]
        (b) **hard'han·dig** b.nw. en bw. Hard, ru van hand […]

Volgens die "Voorwoord by die Derde Uitgawe" in HAT[4] (2000: ix) word samestellings in HAT[4] nie van meervoudsvorme voorsien nie. Die veronderstelling is dat samestellings se meervoude teruggevind kan word deur die laaste stam van die samestelling na te slaan, waar dié se meervoudsvorm aangedui word; dié vorm is dan van toepassing op die samestelling. Daarom word die meervoudsvorm van 'n lemma soos **janblom** ((20)(a)) aangebied, maar die kompositumlemma **middelvoor** ((20)(b)) word nie van hierdie inligting voorsien nie. Laasgenoemde voorbeeld illustreer egter die probleem met hierdie beginsel: Die stam *-voor* in *middelvoor* slaan naamlik nie op die substantief *voor* waarvan die meervoudsvorm *vore* is nie, maar wel op die setsel *voor* wat geen meervoud kan neem nie. Dieselfde geld die eerste stam *middel-*. 'n Meervoudsvorm behoort dus by die substantieflemma **middelvoor** aangebied te word. Om hierdie probleme te vermy, sal dit raadsaam wees om ten opsigte van alle substantieflemmata meervoudsvorme te verstrek. Die konsekwente toepassing van so 'n beginsel sal ook die verdere nut hê dat dit implisiet onderskeid tussen soortname (waar meervoudsvorme aangebied word) enersyds en massaname en abstrakta (waar meervoudsvorme nie aangebied word nie) andersyds tref, mits dié gebruik in die metateks verantwoord word.

Wat die aanbied van dimunitiefvorme ten opsigte van lemmata betref, word die volgende werkswyse in die metateks (HAT[4] 2000: ix) verklaar: "Soms word die verkleiningsvorm ook gegee, veral as dit wat die spelling betref, probleme kan skep." Om hierdie (enigsins vae) beginsel te toets, is 21 lukraak geselekteerde leksikale items in die dimunitiefvorm se opname in HAT[4] nagegaan. Die resultaat van hierdie toets word in Tabel 4 opgesom:

Leksikale element	Formaat in HAT[4]		Leksikale element	Formaat in HAT[4]	
	Morfol.	Lemma		Morfol.	Lemma
apie	+	+	*lampie*	+	−
bordjie	+	+	*lettertjie*	−	−
broekie	+	×	*liggie*	+	−
dogtertjie	−	−	*mannetjie*	+	+
drukkie	+	−	*pennetjie*	+	−
glasie	+	+	*prentjie*	−	+
hondjie	+	+	*rekenaartjie*	−	−
karretjie	+	+	*sebratjie*	−	−
katjie	+	+	*seuntjie*	−	−
kindjie	+	+	*vroutjie*	−	+
knoppie	−	−			

Tabel 4:   Die lemmatisering van 'n aantal dimunitiefvorme in HAT[4]

(Legende: "Morfol." = opname van die betrokke afleiding as grammatiese inligting ten opsigte van die lemma wat die stam verteenwordig; "Lemma" = opname van die dimunitiefvorm as selfstandige lemma; "+" = die betrokke leksikale item is opgeneem in die betrokke formaat; "−" = die betrokke leksikale item is nie in die betrokke formaat opgeneem nie; "×" = die dimunitiefvorm lei nie tot alfabetiese uiteenplasing in die standaard eksterne toegangstruktuur nie en verskyn derhalwe nie as aparte lemma nie)

Uit die selfverklarende resultate in Tabel 4 blyk dat die beginsel wat in die metateks verklaar word, nie konsekwent deurgevoer is nie.

In HAT[4] se metateks word geen melding gemaak van die hantering van adjektiewe nie. Uit die bestudering van die mikrostruktuur moet die gebruiker self aflei dat 'n adjektiwiese lemma in die predikatiewe vorm verskyn met geen inligting oor die attributiewe vorm in die artikelgleuf vir morfologiese inligting nie. Die attributiewe vorm van 'n adjektieflemma kan klaarblyklik uit die voorbeeldsinne afgelei word. In (21)(a) word die fleksieverbuigings vir trappe van vergelyking van die adjektieflemma **glad** aangebied, hoewel in die metateks (HAT[4] 2000: x) verklaar word dat dié inligting in die reël nie aangebied word nie, soos in (21)(b). Vir ongewone vorme word die gebruiker na die betrokke lemma se voorbeeldsinne verwys. Hierdie onsistematiese voorkoms behoort opgeklaar te word deur óf die beginsels wat in die metateks verklaar word sonder uitsondering toe te pas óf om die inligting ten opsigte van elke lemma in die artikelgleuf vir morfologiese inligting oor die lemmateken te plaas. Laasge-

noemde metodiek sou aanbeveel word ter wille van gebruikersvriendelikheid, hoewel ten koste van ekonomie.

Hierteenoor vertoon KW konsekwente toepassing van leksikografiese konvensies wat in die metateks verantwoord word: by substantieflemmata word die meervoudsvorm en die verkleiningsvorm gegee; by verbumlemmata word die verlededeelwoordvorm gegee; by adjektieflemmata word sowel die attributiewe vorm as die trappe van vergelyking van die betrokke lemma voorsien. Hierdie eksplisiete en konsekwente aangawe van grammatiese inligting oor die lemmata is gesonde leksikografiese praktyk, veral indien in ag geneem word dat die moontlike gebruiker van KW 'n leerder is.

6.4   Semantiese kommentaar

In die bespreking oor semantiese kommentaar word primêr op die ooreenkomste en verskille tussen HAT[4] en KW gewys; HAT[4] word dus nie hier met HAT[3] vergelyk nie.

6.4.1   Betekenisparafrases

In HAT[4] word betekenisparafrases geformuleer sodat hulle min of meer die lemmata waarop hulle betrekking het, binne konteks kan vervang. So kan die betekenisparafrase van die eerste polisemiese waarde van die adjektieflemma **groot** in (22)(a) (HAT[4] 2000: 332) die leksikale item *groot* in (22)(b) vervang om (22)(c) te lewer sonder verlies aan betekenis:

(22)   (a) **groot** […] **1** Van meer as normale afmeting […]
          (b) Dit is 'n *groot* padkaart.
          (c) Dit is 'n padkaart *van meer as normale afmeting*.

Hierdie metode van parafrasering verteenwoordig 'n teksverdigtingstrategie, aangesien betekenisparafrases in volsinvorm groot hoeveelhede kosbare ruimte in die woordeboek in beslag sou neem. Landau (2001: 164) wys op die probleem dat die beginsel van vervangbaarheid nie deurgaans toegepas kan word nie, soos onder meer op die leksikale items *aan, die, dit, in, is, ja* en *wat*. Hierop word nie verder ingegaan nie. Gouws (1989: 114-121) en Svensén (1993: 122-139) bespreek ook betekenisparafrases ("leksikografiese definisies") in besonderhede.

Volgens Landau (2001: 164) is *Collins COBUILD English Language Dictionary* (Sinclair et al. 1987) die eerste woordeboek wat die innoverende stelsel van betekenisparafrases in volsinvorm ingevoer het: "The editors of *Cobuild* decided that sentence definitions were easier for their readers to understand and gave them more assistance in learning how to use the word than a substitutable definition." Hierdie stelsel is volgens Landau prysenswaardig, veral die toepassing daarvan in woordeboeke gerig op aanleerders van die taal. KW het hierdie voorbeeld gevolg deur lemmata deurgaans van volsinparafrases te voorsien. In

die "Voorwoord" van KW (2000: iii) word die besluit om hierdie metode van parafrasering te volg, verantwoord: "Daardeur lees die definisie nie net makliker nie, en is dus nie net makliker verstaanbaar nie, maar die hoofwoord word onmiddellik gegee in 'n verband waarin dit in gewone Afrikaans gebruik word." Vergelyk die volgende uittreksels uit KW:

(23)   (a) **ge·sant'** […] **1** 'n Gesant is iemand wat 'n staat in 'n ander staat verteenwoordig, net laer in rang as 'n ambassadeur […]
       (b) **ne'tjies** […] Iemand wat netjies is, is presies, skoon, ordelik […]
       (c) **stuk** […] **1** 'n Stuk is 'n brok, gedeelte, deel van iets anders […] **4** In die geselstaal beteken **stuk** "nooi of kêrel" […]

Uit (23)(c) blyk dat wanneer 'n lemma in sy selfnoemfunksie in 'n betekenisparafrase in KW optree, daardie leksikale item in vetdruk verskyn (dit wil sê dieselfde tipografiese struktuurmerker as dié van die lemmateken tree op), hoewel geen melding van sodanige stylbeginsel in die metateks gemaak word nie.

Volgens die "Voorwoord" in KW (2000: iii) word pragmatiese inligting oor die lemma ook in die betekenisparafrase gegee, wat die gebruik van leksikografiese etikette uitskakel. Vergelyk die vierde polisemiese waarde van die lemma **stuk** in (23)(c), wat in 'n standaardwoordeboek met 'n etiket soos *(geselstaal)* gemerk sou wees. Hieroor kortliks meer in die volgende afdeling (vgl. 6.4.2).

Die volsinparafrases in KW verteenwoordig 'n strategie wat in die eerste plek op die gebruiker (waarskynlik die skoolleerder) gerig is. Dit is goeie leksikografie. Op die spesifieke aard van volsinparafrases soos dit in KW voorkom, sal nie hier ingegaan word nie.

6.4.2   Etikette

In HAT[4] word leksikografiese etikette gebruik om pragmatiese kommentaar op (bepaalde polisemiese waardes van) lemmata te lewer. Hierdie gebruik word effektief in die sentrale teks toegepas en deeglik in die metateks verantwoord (vgl. 3.2).

In die "Voorwoord" van KW (2000: iii) word gemeld dat die spesifieke verklaringstegniek wat in dié woordeboek gebruik word, die aangee van leksikografiese etikette in 'n hoë mate uitsluit omdat die betekenisparafrase self die verband aandui. Dit veronderstel dat leksikografiese etikette as spesifieke inskrywingstipe nie in KW figureer nie. Aangesien die gebruik van leksikografiese etikette nie net in Afrikaanse woordeboeke voorkom nie, maar reeds as 'n internasionale leksikografiese tradisie beskryf kan word, is dit jammer dat hierdie inskrywingstipe nie ook in KW ingesluit is nie. Indien die teikengebruiker van KW 'n leerder is, sou selfs die aangawe van etikette tesame met gelykluidende pragmatiese kommentaar in die betekenisparafrase voorgestel kon word. Om die leerder egter van toereikende woordeboekopleiding te voorsien, sou vereis dat minstens leksikografiese etikette as draers van pragmatiese inligting oor lemmata by die mikrostruktuur ingesluit word.

6.4.3   Voorbeeldmateriaal

Studies soos Gouws (1989), Lombard (1992), Rademeyer (1992, 1993) en Van Niekerk (1992) bespreek die aanwending en funksies van voorbeeldmateriaal (koteks) in Afrikaanse woordeboeke. Verder lewer Carstens (1995: 158-159) en Van der Merwe (1994: 236) spesifiek kommentaar op dié datakategorie in HAT[3], wat oorgedra kan word op HAT[4]. Noemenswaardig is wel dat "heelwat meer sitate, ook van jonger skrywers en skrywers van ligter literatuur" opgeneem is "om juis 'n beeld van die werklike gebruik van Afrikaans te gee" (HAT[4] 2000: v). Hier word derhalwe bloot kortliks gewys op die belangrikste verskille tussen standaardwoordeboek en kernwoordeboek wat hierdie datakategorie betref.

Waarskynlik die belangrikste onderskeid tussen HAT[4] en KW wat voorbeeldmateriaal betref, is dat KW ten opsigte van elke (polisemiese waarde van 'n) lemma ten minste een koteksinskrywing aanbied, terwyl daar by talle lemmata in HAT[4] geen koteksinskrywings gegee word nie. Verder poog KW kennelik in soveel gevalle as wat ruimtelik toelaatbaar is om voorbeeldmateriaal in volsinvorm weer te gee, terwyl HAT[4] uitgebreid frases waarin die lemma as kernelement optree, gebruik. Gegee die teikengebruikers van hierdie twee woordeboeke, is dié enigsins uiteenlopende tipes weergawes van voorbeeldmateriaal funksioneel. KW verhoog sy toeganklikheid en daarmee sy gebruikersvriendelikheid deur hierdie praktyk wat die waarskynlikste gebruiker (die leerder) tot voordeel strek. Die gebruik van frases in KW kan ook pedagogies aangewend word in die opleiding van die teikengebruiker vir latere raadpleging van standaardwoordeboeke.

6.4.4   Illustrasies

Die gebruik van illustrasies in die sentrale tekste van Afrikaanse standaard- en beperkte woordeboeke het tot op hede nog geen verskyning gemaak nie. Veral ten opsigte van eentalige aanleerderwoordeboeke en skoolwoordeboeke behoort hierdie inligtingstipe aandag te geniet.[12] Vergelyk Gouws (1994), Landau (2001: 143-147) en Svensén (1993: 167-180) vir besprekings oor ostensiewe adressering.

6.5   Vaste uitdrukkings

Botha (1991) en Gouws (1990) bespreek die leksikografiese hantering van vaste uitdrukkings in Afrikaanse woordeboeke. Hier sal volstaan word met 'n kort verwysing na die ooreenkomste tussen HAT[4] en KW wat die bewerking van dié tipe multileksikale item betref.

Hoewel vaste uitdrukkings as multileksikale items leksikologies gesproke volledige lemmastatus in woordeboeke behoort te geniet, is dit in die Afrikaanse leksikografiese tradisie tot nog toe nie die geval nie (vgl. Gouws 1989: 97-101). In beide HAT[4] en KW is die gebruik om vaste uitdrukkings onder 'n lem-

ma wat deel van die betrokke uitdrukking vorm, in 'n rubriek gemerk "UITDR."
te verklaar. Dié gebruik word soos volg in die metateks van HAT⁴ (2000: xi-xii)
verantwoord:

> Indien daar 'n selfstandige naamwoord in die uitdrukking voorkom, sal die uit-
> drukking by die (eerste) selfstandige naamwoord voorkom. […] Indien daar
> meer as een selfstandige naamwoord voorkom, word die uitdrukking onder die
> eerste selfstandige naamwoord verklaar […] Sou daar geen selfstandige naam-
> woord voorkom nie, soek onder die eerste byvoeglike naamwoord, eerste by-
> woord, eerste werkwoord en eerste voorsetsel, in dié volgorde.

Volgens dié beginsel word die vaste uitdrukking *met huid en haar* onder die
lemma **huid** in HAT⁴ en KW opgeneem. In (24) word die betrokke artikel in
KW weergegee:

(24)   **huid** *s.nw.* (huide)
    1 Huid is 'n deftige woord vir die vel van 'n mens: *'n Bruingebrande huid*.
    2 Huid is die afgeslagte vel van 'n groot dier (teenoor vel van 'n kleiner dier):
    Die handel in huide en velle.
    UITDR. *Met huid en haar*, totaal, bv. *'n prooi met huid en haar opvreet*.
    huid'aandoening.
    huid'arts.
    huid'plooi. huid'siekte.
    huid'vlek. huid'wond.

Hoewel KW 'n identiese stelsel as HAT⁴ in hierdie verband volg, word die ern-
stige leksikografiese fout in KW begaan dat dié stelsel glad nie in die metateks
verantwoord word nie. Gebruikers (leerders) moet dus self probeer om die stel-
sel te ontrafel, wat hul woordeboekgebruikspogings frustreer as gevolg van
ontoeganklikheid. Indien KW as skoolwoordeboek geld, is hierdie oorsig 'n
selfs ernstiger tekortkoming.

6.6   Etimologiese inligting oor die lemma

In die "Voorwoord by die Derde Uitgawe" in HAT⁴ (2000: viii) word te kenne
gegee dat etimologiese inligting ten opsigte van lemmata tereg perifere inlig-
ting verteenwoordig, en sedert die tweede uitgawe is nie daarop uitgebrei nie.
    Etimologiese inligting oor lemmata is nie vir die teikengebruiker van KW
van belang nie, en word nie in dié woordeboek se mikrostruktuur opgeneem
nie.

7.   Gevolgtrekking

In 'n studie van hierdie omvang kan die standaardwoordeboek en kernwoor-
deboek nie volledig vergelyk word nie. Daarom is op enkele belangrike aspekte
gekonsentreer. HAT⁴ toon bepaalde ooreenkomste met KW wat (tans) funksio-

neel is: Indien KW beskou word as skoolwoordeboek, kan dié woordeboek as 'n soort voorloper vir HAT[4] beskou word deurdat die skoolverlater uiteindelik as volwasse taalgebruiker sal oorbeweeg van die gebruik van KW na die gebruik van HAT. In hierdie opsig sluit die twee woordeboeke tipologies mooi by mekaar aan. Problematiese aspekte van die makrostruktuur, toegangstrukture en mikrostrukture in beide woordeboeke behoort egter uitgeklaar te word sodat die woordeboeke uiteindelik hul volle potensiaal as houers van taalkundige inligting kan verwesenlik. Hierin is die fokus op die teikengebruiker essensieel, wat by verstek die deeglike en toeganklike verantwoording van die leksikografiese praktyk in dié woordeboeke se onderskeie metatekste veronderstel.

Aantekeninge

1.   Die lemma **cool** is slegs as leenitem uit Engels geëtiketteer. Kennelik kort die etiket *(geselst.)*.
2.   Die mediostruktuur sal nie pertinent in hierdie artikel aan bod kom nie; vergelyk Gouws en Prinsloo (1998) en Svensén (1993: 194-199) vir besprekings van die mediostruktuur in woordeboeke. Carstens (1995: 159-163) bespreek spesifiek elemente van die mediostruktuur in HAT[3].
3.   'n Probleem met hierdie opskrif is dat die woord *gebruiklikste* nie volledigheid veronderstel nie. In hierdie opsig is die opskrif van dieselfde afdeling in HAT[4] leksikografies meer bevredigend.
4.   Juis vir die doel van BW, naamlik om vreemdetaalsprekers van diens te wees, is dit essensieel dat alle leksikale items wat in betekenisparafrases en koteksinskrywings voorkom, ook as lemmata opgeneem behoort te wees.
5.   Die lemmateken van die leksikale item *November* is foutiewelik met 'n kleinletter gedruk.
6.   Hier word nie op die meriete van die betekenisparafrases as sodanig kommentaar gelewer nie.
7.   In HAT[4] is 100 groeperings van sublemmata lukraak as toets gekies. Dié seleksie het 1 236 sublemmata in bestek gehad. In KW is alle groeperings van sublemmata onder die letter S as toets gekies, wat 1 674 lemmata in bestek gehad het. (Die letter S is in KW gekies omdat dié deel van die sentrale teks 115 bladsye van die totale sentrale teks van 757 bladsye beslaan, wat 'n tekssteekproef van 15,19% verteenwoordig.)
8.   In HAT[4] toon 6 (0,5%) van die 1 236 getoetste sublemmata alfabetiese afwyking, terwyl 24 (1,44%) van die 1 674 sublemmata in KW alfabetiese afwyking toon. Teen die agtergrond van hierdie lae frekwensies en die afwesigheid van semantiese struktuur in dié gevalle, is daar eerder sprake van ordeningsfoute as van tweedevlaknesting.
9.   Legende: "sp" = sinoniemparadigma.
10. Die rekenaarvaktaalpoliseem sal by inkorporering onder die lemma **aftas** op grond van relatiewe gebruiksfrekwensie moontlik die eerste polisemiese waarde word. Hier word duidelikheidshalwe egter aangeneem dat dié poliseem as derde poliseem naas die twee bestaandes bygevoeg sal word.
11. Hier word nie nou uitgebrei op die formaat van die verantwoordings nie.
12. Intussen het Rufus Gouws, Mariza Stark en Liezl Gouws se *Nuwe Woordeboek sonder Grense* (2004) by Maskew Miller Longman, Kaapstad, verskyn, waarin illustrasies wel in die sentrale teks voorkom.

Verwysings

Woordeboeke en ander taalnaslaanwerke

Bresler, G.C. 1978. *40 000 woorde*. Johannesburg: McGraw-Hill.

De Villiers, M., J. Smuts, L.C. Eksteen en R.H. Gouws. 1987[7]. *Nasionale Woordeboek*. Kaapstad: Nasou.

Gouws, Rufus, Ilse Feinauer en Fritz Ponelis. 1994. *Basiswoordeboek van Afrikaans*. Pretoria: J.L. van Schaik.

Grobbelaar, P. (Red.). 1987. *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Diction-ary*. Kaapstad: The Reader's Digest Association.

Labuschagne, F.J. en L.C. Eksteen. 1993[8]. *Verklarende Afrikaanse Woordeboek*. Pretoria: J.L. van Schaik.

Müller, Dalene. 2003. *Skryf Afrikaans van A tot Z*. Kaapstad: Pharos.

Odendal, F.F. 2000. *Kernwoordeboek*. Midrand: Perskor.

Odendal, F.F. en R.H. Gouws. 2000[4]. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Midrand: Perskor.

Odendal, F.F., P.C. Schoonees, C.J. Swanepoel, S.J. du Toit en C.M. Booysen. 1979[2]. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Johannesburg: Perskor.

Odendal, F.F., P.C. Schoonees, C.J. Swanepoel, S.J. du Toit en C.M. Booysen. 1994[3]. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Midrand: Perskor.

Schoonees, P.C., C.J. Swanepoel, S.J. du Toit en C.M. Booysen. 1965. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Johannesburg: Voortrekkerpers.

Sinclair, J. et al. (Reds.). 1987. *Collins COBUILD English Language Dictionary*. Londen: Harper-Collins.

Sinclair, J. et al. (Reds.). 1995[2]. *Collins COBUILD English Language Dictionary*. Londen: Harper-Collins.

Taalkommissie van die S.A. Akademie vir Wetenskap en Kuns. 2002[9]. *Afrikaanse woordelys en spelreëls*. Kaapstad: Pharos.

Ander bronne

Beyer, Herman L. 2000. Nuwe dinge in *Nuwe woorde* — 'n Oorsig oor buitemikrostrukturele leksi-kografiese vernuwings. *Lexikos* 10: 275-290.

Botha, T.J.R., F.A. Ponelis, J.G.H. Combrink en F.F. Odendal (Reds.). 1989[2]. *Inleiding tot die Afri-kaanse taalkunde*. Pretoria/Kaapstad: Academica.

Botha, W.F. 1991. Die lemmatisering van uitdrukkings in verklarende Afrikaanse woordeboeke. *Lexikos* 1: 20-36.

Carroll, J.B. 1971. Development of Native Language Skills beyond the Early Years. Reed. C. (Red.). 1971: 97-156.

Carstens, Adelia. 1995. 'n Kritiese beskouing van HAT[3]. *Lexikos* 5: 138-165.

Combrink, J.G.H. 1990. *Afrikaanse morfologie: Capita exemplaria*. Pretoria: Academica.

Cruse, D.A. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

De Stadler, L.G. 1989. *Afrikaanse semantiek*. Johannesburg: Southern Boekuitgewers.

De Stadler, L.G. 1991. Sintagmatiese leksikale betrekkinge in Afrikaans. *Lexikos* 1: 61-82.

Diller, K. 1971. *Generative Grammar, Structural Linguistics and Language Teaching*. Rowley, Mass.: Newbury House.

Dromi, Esther. 1987. *Early Lexical Development*. Cambridge: Cambridge University Press.

Eksteen, L.C. 1989. Die Afrikaanse woordeskat. Botha, T.J.R. et al. (Reds.). 1989: 358-381.

Emejulu, James Duplessis (Red.). 2001. *Éléments de Lexicographie Gabonaise. Tome 1.* New York: Jimacs-Hillman.

Gleitman, Lila en Barbara Landau. 1994. *The Acquisition of the Lexicon*. Cambridge, Mass./Londen: Massachusetts Institute of Technology.

Gouws, R.H. 1988. Die gebruik van etikette as leksikografiese hulpmiddel. *Suid-Afrikaanse Tydskrif vir Taalkunde,* Geleentheidsuitgawe 6, April 1988.

Gouws, R.H. 1989. *Leksikografie*. Pretoria/Kaapstad: Academica.

Gouws, R.H. 1989a. Leksikale semantiek. Botha, T.J.R. et al. (Reds.). 1989: 382-411.

Gouws, R.H. 1990. Vaste uitdrukkings as multileksikale lemmas in Afrikaanse woordeboeke. *Tydskrif vir Geesteswetenskappe* (30)4: 265-283.

Gouws, R.H. 1994. Ostensiewe adressering in vertalende woordeboeke. *Lexikos* 4: 61-85.

Gouws, R.H. 1996. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 14-31.

Gouws, Rufus H. 2001. Lexicographic Training: Approaches and Topics. Emejulu, James Duplessis (Red.). 2001: 58-94.

Gouws, Rufus H. 2002. Niching as a Macrostructural Procedure. *Lexikos* 12: 133-158.

Gouws, R.H. en D.J. Prinsloo. 1998. Cross-Referencing as a Lexicographic Device. *Lexikos* 8: 17-36.

Landau, Sidney I. 2001². *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

Lombard, F.J. 1992. Voorbeeldmateriaal in woordeboeke. *Lexikos* 2: 148-164.

Louw, Phillip Adriaan. 1999. Access Structures in a Standard Translation Dictionary. *Lexikos* 9: 108-118.

Louw, Phillipus Adriaan. 1997. *Kriteria vir 'n standaard vertalende woordeboek*. Ongepubliseerde M.A.-skripsie. Stellenbosch: Universiteit van Stellenbosch.

Lyons, John. 1977. *Semantics. Volume 2*. Cambridge: Cambridge University Press.

Lyons, John. 1981. *Language and Linguistics: An Introduction*. Cambridge: Cambridge University Press.

Moe, Ronald. 2003. Compiling Dictionaries Using Semantic Domains. *Lexikos* 13: 215-223.

Otto, A.N. 1995. Resensie van *Basiswoordeboek*. *Lexikos* 5: 219-231.

Rademeyer, L. 1992. *Die funksies van verbale voorbeeldmateriaal in eentalige woordeboeke*. Ongepubliseerde M.A.-skripsie. Stellenbosch: Universiteit van Stellenbosch.

Rademeyer, L. 1993. Funksies van voorbeeldmateriaal in eentalige woordeboeke. *Lexikos* 3: 205-214.

Reed, C. (Red.). 1971. *The Learning of Language*. New York: Appleton–Century–Crofts.

Singleton, David. 2000. *Language and the Lexicon: An Introduction*. Londen: Arnold.

Stark, Mariza. 2001. *Die toegangstruktuur in verklarende aanleerderwoordeboeke*. Ongepubliseerde M.A.-skripsie. Stellenbosch: Universiteit van Stellenbosch.

Svensén, Bo. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.

Van der Merwe, H.J.J.M. (Red.). 1968. *Afrikaans: Sy aard en ontwikkeling*. Pretoria: J.L. van Schaik.

Van der Merwe, H.J.J.M. 1968. Die Afrikaanse woordeskat. Van der Merwe, H.J.J.M. (Red.). 1968: 111-147.

Van der Merwe, Michele. 1994. 'n Evaluering van die *Verklarende Handwoordeboek van die Afrikaanse Taal* as standaard verklarende woordeboek. *Tydskrif vir Geesteswetenskappe* (34)4: 231-236.

Van Niekerk, A.E. 1992. Kollokasies: 'n Leksikografiese perspektief. *Lexikos* 2: 254-264.

Wiegand, Herbert Ernst. 1990. Printed Dictionaries and their Parts as Texts: An Overview of More Recent Research as an Introduction. *Lexicographica* 6: 1-126.

Wiegand, Herbert Ernst. 1996. A Theory of Lexicographic Texts: An Overview. *Suid-Afrikaanse Tydskrif vir Taalkunde* 14(4): 134-149.

Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.

Wierzbicka, Anna. 1996. *Semantics, Primes and Universals*. Oxford: Oxford University Press.

Zgusta, Ladislav. 1971. *Manual of Lexicography*. The Hague: Mouton.

The Use of Pictorial Illustrations in African Language Dictionaries[*]

Lilian Gangla-Birir, *Unit for Language Skills Development, University of Pretoria, Pretoria, Republic of South Africa (kenyan@mweb.co.za)*

Abstract: This article examines the use of pictorial illustrations in African language dictionaries, with particular reference to items generally included in dictionaries based on Stein's (1991) categories. It argues for the need to include more illustrations of cultural items specific to a particular community and language group in dictionaries. A corpus analytic survey is done to determine if African language dictionaries employ pictorial illustrations and, if so, which items are illustrated. The results of this survey indicate that only a few African language dictionaries employ pictorial illustrations, even those that do correspond to Stein's categories. Of the dictionaries examined, only one uses illustrations in the form of full-colour photographs, namely of different kinds of cattle, a very important cultural item within that specific community.

Keywords: PICTORIAL ILLUSTRATIONS, CULTURE, LEXICOGRAPHY, BILINGUAL DICTIONARY, MONOLINGUAL DICTIONARY, AFRICAN LANGUAGE

Opsomming: Die gebruik van beeldillustrasies in Afrikataalwoordeboeke.
Hierdie artikel ondersoek die gebruik van beeldillustrasies in Afrikataalwoordeboeke met spesiale verwysing na items wat gewoonlik in woordeboeke ingesluit word, gebaseer op Stein (1991) se kategorieë. Dit betoog vir die noodsaaklikheid om meer illustrasies van kulturele items eie aan 'n besondere gemeenskap of taalgroep in woordeboeke in te sluit. 'n Analitiese oorsig van 'n korpus is gedoen om te bepaal of Afrikataalwoordeboeke beeldillustrasies gebruik, en, indien wel, watter items geïllustreer word. Die resultate van hierdie oorsig toon aan dat slegs 'n paar Afrikataalwoordeboeke beeldillustrasies gebruik, selfs dié wat met Stein se kategorieë ooreenstem. Van die woordeboeke wat ondersoek is, gebruik net een illustrasies in die vorm van volkleurfoto's, naamlik van verskillende soorte beeste, 'n baie belangrike kulturele item binne daardie spesifieke gemeenskap.

Sleutelwoorde: BEELDILLUSTRASIES, KULTUUR, LEKSIKOGRAFIE, TWEETALIGE WOORDEBOEK, EENTALIGE WOORDEBOEK, AFRIKATAAL

Introduction

The role and function of pictorial illustrations for educational purposes cannot be overemphasised. Despite various studies carried out to determine the exact contribution of illustrations to learning and knowledge acquisition, the find-

---

ings have varied significantly. However, education systems are to a large extent based on reading and writing and most source materials are produced in written form. Among the source materials used within education systems are dictionaries which are written to serve different purposes. These are divided into two main categories: dictionaries that focus on language (linguistic dictionaries) and those that focus on the world (encyclopaedic dictionaries) (cf. Zgusta 1971, Carstens 2001).

Zgusta (1988) states that, since the dictionary is a text couched in natural language, it possesses pragmatic aspects itself. In this article, the role of illustrations in conveying cultural information in monolingual and bi- or multi-lingual dictionaries is dealt with. Though an examination of the included pictorial illustrations was done, at this stage the purpose of the dictionary survey was to investigate which objects, if any, are illustrated in these dictionaries. It was attempted to establish if there is a kind of pattern in what is illustrated. Fleming (1967: 247) defines pictorial illustrations as 'those configurations of line, dot or area and any combination of these three resembling events or objects (persons, places and/or things) either as perceived or as generally conceived'. Al-Kasimi (1977: 77) comments that the definition should be expanded to include such borderline cases as number lines, geometric figures, structural chemical formulas, curves, graphs and time lines. Putter (1999: 51) regards the use of illustrations in dictionaries as a type of definition, which is known in lexicography as 'ostensive definition' or 'ostensive addressing' (in Gouws' terminology). It is normally a type of definition that is used to augment and to elucidate the verbal definition (Gouws 1989: 168; Swanepoel 1990: 186).

The inclusion of encyclopaedic information in dictionaries

For a long time, in the history of lexicography, a strict distinction was made between linguistic and encyclopaedic (pragmatic, real-world) information. During the nineties, lexicographers came to realize that part of what is known about the meaning of a word belongs to what is known about the surrounding world. This realization led to a less strict division between linguistic and encyclopaedic information, and between dictionaries and encyclopaedias.

Cultural information is a type of information that merits the inclusion of illustrations in dictionaries. In itself it may be regarded as a type of encyclopaedic information. Illustrations may serve as vehicles to clarify and/or disambiguate the meanings of culture-specific words. Lexicographers also came to realize that the criteria for inclusion of illustrations in different types of dictionaries might differ on the basis of the cultural differences between dictionary users and the culture in which the target language is embedded. This gap would, for instance, determine how much encyclopaedic information should be presented to bridge the gap. Pedagogical dictionaries would particularly benefit from the inclusion of encyclopaedic information, such as illustrations. Due to their comprehensiveness, certain overall descriptive dictionaries, such as the *Woordeboek van die Afrikaanse Taal* (WAT), in its first volumes, included several illustrations.

Metalexicographical perspectives on the function and use of illustrations in dictionaries

Gouws (1994: 61) argues that pictorial illustrations can be employed as a microstructural type to enhance semantic disambiguation. He further adds that pictorial illustrations have a semantic importance, a lexicographical function and a range of applications, and although the primary utilization of ostensive addressing is found in the differentiation of senses in equivalent relations of semantic divergence, they are also used extensively where a lemma has a low degree of translatability. The explicitness that can be achieved through the use of pictorial illustrations enhances the retrieval of information. This applies to both monolingual descriptive and translation dictionaries. Al-Kasimi (1977) argues that pictorial illustrations should be systematically and consistently employed in bilingual dictionaries, not for the purpose of advertisement but as an essential lexicographical device.

The role of culture in the compilation of dictionaries

Culture is defined as that complex whole which includes knowledge, beliefs, arts, morals, law, custom and any other capabilities and habits acquired by man as a member of society (Tylor 1871 in Ocholla-Ayayo 1980: 2). Casson (1981: 17), on the other hand, defines culture as 'idealised cognitive social systems of knowledge, beliefs, and value that exist in the collective minds of individual members of society. It is the mental equipment that society members use in orienting, transacting, discussing, defining, categorising, and interpreting social behaviour in their society.' It is the means by which society members generate appropriate social behaviour and appropriately interpret the behaviour of others.

The lexicographical history of African languages has been dominated by bilingual lexicography. There are, however, a number of monolingual dictionaries in some of the African languages. The users of these dictionaries vary from mother-tongue speakers, to second language speakers to foreign learners. An examination of the cultural setting and the equivalence in bilingual dictionaries reveals that language does not exist in a vacuum but occurs in context. Hence, relationships between the various languages as well as within a language become clear. One of the important aspects is the cultural context in which a word is used. The cultural context determines its meaning and to a large extent also the ability of a lexicographer either to translate the word into another language or to describe it in another language.

The relationship between language and culture is a major issue in cross-cultural communication. Ndoleriire (2000: 272) points out that language can be considered a cultural practice, and that it is both an instrument to and a product of culture. He further adds that 'the different languages spoken by different groups of people who have the same culture are used to express cultural norms

and practices which are the same in those groups of people'. Therefore, it is apparent that culture is learnt through human interaction and linguistic communication and that both language and communication are functions of culture. According to the Sapir–Whorf hypothesis, language and culture are intimately related and language provides a key to understanding culture. This statement further indicates that language is an important instrument in the totality of culture. The degree of misunderstanding between cultures depends on the degree of difference between the cultures concerned. Language is a part of culture and culture is a part of language; the two are so intricately interwoven that they cannot be separated without losing the significance of either language or culture (Brown 1987: 123).

Ndoleriire (2000: 274) further exemplifies that 'individual people learn the values, norms, beliefs, views, and behavioural patterns of their group through linguistic interaction and groups give expression to their cultural identities, and practise their cultural life, not only through music, art, dancing, and dress, but also through language'. Accordingly, the link between peoples' way of life and their language can be clearly seen. Being a written record of this language, the dictionary must find a way to carry on the peoples' culture as part of their written heritage. Language is the key component of the dictionary and therefore the lexicographer needs to look at the best way to describe it. In this regard, where there are semantic gaps due to culture-specific words or ideas, it is important that another way is used to refer to the concept. Pictorial illustrations provide an accurate way of representing the concept, given that the illustrations — whether photographs or line drawings, full colour or black and white illustrations — communicate the idea effectively.

Users of bilingual dictionaries, in particular, may benefit from illustrations. According to Zgusta (1988: 247), the most pragmatically effective equivalent of a word in a bilingual dictionary is a translational one; that is, an equivalent which is the counterpart of the target language lexical unit in all respects, namely with the same denotations, stylistic levels, and collocations. Particularly due to cultural differences, however, incomplete equivalence or the absence of any equivalent is experienced. This is not only limited to cultural-bound words but also differences in the pragmatics and ethnology of speaking. This aspect must be taken into account when compiling a dictionary, particularly a bilingual dictionary, and in such instances pictorial illustrations will go a long way in bridging the semantic gap between the languages.

In spite of this, illustrations should not necessarily have a uni-directional function in dictionaries. According to Zgusta (1988: 246), lexicography strongly takes into consideration the culture of the language described. He illustrates this by using as example the *Bilingual Dictionary of Zoque* (Engel and Allhiser de Engel 1987). This dictionary, which has many illustrations, does not only use them to explain peculiarities of Zoque culture to speakers of Spanish, but they also illustrate Spanish cultural items to Zoque speakers. For instance, there is an illustration of a boy buying something in a shop because objects and situa-

tions like these belong not so much to Zoque culture as to the culture surrounding them. Thus, the dictionary not only explains 'the Zoque culture to the Spanish but also the Spanish culture to the Zoque, certainly a rather novel attitude in a dictionary of a language of a small, rural group' (Zgusta 1988: 246). This is in contrast to, for instance, the *South African Multi-language Dictionary and Phrase Book* where the illustrations are more Eurocentric than African. For example, the rooms in a house include a pantry while the sports played are cricket and rugby. This is despite the fact that this is a multi-language dictionary that covers seven languages and apart from English and Afrikaans the rest are African languages (Northern Sotho, Sesotho, Tswana, Xhosa and Zulu).

Material and methods

In pursuance of this task, a deductive research method was employed, based on existing literature within the field of pictorial illustrations. A survey of dictionaries was done to determine which dictionaries employ pictorial illustrations. The dictionaries were not representative of all the African languages and for some of the languages there was more than one dictionary. This limitation partly due to the fact that the study was limited to dictionaries available in the Merensky Library at the University of Pretoria and additional dictionaries obtained from the Department of African Languages at the University of Pretoria. As a result of this, most of the African language dictionaries are of African languages spoken within South Africa.

   The inclusion of English dictionaries was to compare the objects they illustrate to those illustrated in the African language dictionaries and to see if there are any similarities or differences.

Dictionary survey

Stein (1991: 106) distinguishes four main types of illustrations in dictionaries:

1.   Illustrations showing common animals, objects and plants.

2.   Illustrations showing things that are not easily explained in words, such as shapes, complex actions, or small differences between words which are similar but not the same.

3.   Illustrations depicting groups of related objects with a view to explaining the differences between similar objects, showing the range of shapes and forms covered by a particular word and serving as an important aid to vocabulary expansion.

4.   Illustrations showing the basic or physical meaning of words commonly used in an abstract or figurative way.

Culture-specific terms can be added to the above main types, particularly in African language dictionaries.

Corpus-analytic research

For this survey, the dictionaries were picked at random and scanned for any form of pictorial illustrations. The purpose of the exercise was to determine which dictionaries make use of pictorial illustrations and which types of lemmas are typically illustrated, making use of Stein's (1991) typology outlined above. The survey further considered the placement of the illustrations: in the front matter, the central text or the back matter of the dictionary.

Results

Of the dictionaries scanned, only four employed pictorial illustrations, namely *Thanodi ya Setswana*, a monolingual dictionary, the *Bilingual Dholuo–English Dictionary Kenya*, the *Learner's English–Venda Dictionary Illustrated in Colour* and the *South African Multi-language Dictionary and Phrase Book*, which covers English, Afrikaans, Northern Sotho, Sesotho, Tswana, Xhosa and Zulu.

The only African language monolingual dictionary with pictorial illustrations is *Thanodi ya Setswana*. It gives full colour pictures of cattle and trees. However, these pictures do not appear in the central text of the dictionary but are given at the end of the dictionary, forming the back matter. The full colour pictures are accompanied by captions that provide the name of the particular cattle or plant in Setswana. The difficulty the lexicographers would have been faced with if they opted to describe the different kinds of cattle or plants can clearly be realised. It would probably not have been as effective as the pictures used. However, there is no cross-reference from the central text to these illustrations in the back matter. The illustrations therefore serve an encyclopaedic, rather than a lexicographical function.

The *Bilingual Dholuo–English Dictionary Kenya* has line drawings of birds, insects, plants, pots, animals and fish. The dictionary is bi-directional, but the bulk of the illustrations occur in the Dholuo–English section. Of particular relevance are the illustrations of different kinds of pots. In English there is only one word for *pot* but in Dholuo, a Nilotic language spoken in East Africa, there are several words for the concept **pot.** In the Dholuo–English section of this dictionary, the lexicographer chooses to lemmatise the different words for *pot* alphabetically. The illustrations are in the central text of the dictionary.

In the *English–Venda Dictionary Illustrated in Colour*, two full-colour photographs are incorporated along with a word list in which the headword is lexicalised in English, the word class indicated using appropriate abbreviations and the translations provided in Tshivenda. The photographs are of various activities, items, objects, places, animals, birds, insects and flowers. However,

the captions of these photographs only contain the name in English. To find the Tshivenda translation, the user will first have to locate the English word in the A–Z section where the Tshivenda translation is given. The illustrations therefore are not very effective.

The *South African Multi-language Dictionary and Phrase Book* has a visual dictionary in the front matter. The contents of the visual dictionary include pictures of the human body, a house plan (lounge, dining-room, kitchen, pantry, bedroom, bathroom, garden), an office, a car (instruments, mechanics) and sports (soccer, rugby, cricket). Though these items are illustrated, these illustrations will not be further analysed as they fall outside the scope of this study which deals particularly with cultural items.

Discussion

On considering the pictorial illustrations found in the dictionaries surveyed, the ideal criteria as specified by Al-Kasimi (1977: 100-102) have been met. The illustrations are compact as visual noise has been avoided and have fidelity as they are realistically executed. The interpretability of the message conveyed is easily understood. They are relevant as they are related to the user's past environmental and realistic experiences, which are determining factors in the user's success in interpreting the picture. The pictures are simple and precise. The use of captions and titles aid in making the illustrations complete whereas the appropriate size and adequate artistic performance make them clear.

The English dictionaries used for comparison show the inclusion of numerous pictorial illustrations. This may be attributed to a number of factors including a well-developed and highly competitive lexicographical department within the English language industry with a varied target market ranging from first language speakers of English, second language speakers of English and foreign language learners of English with varied needs. The comparison between the English and African language dictionaries was therefore not effective due to a number of factors, including the dissimilar availability of sources.

Based on the findings of the dictionary survey, there is a close correlation between the categories Stein (1991) gives and the items actually illustrated in these dictionaries. The items illustrated across the board include birds, animals, plants and houses. These items fall under Stein's first category of illustrations of common animals, objects, and plants. However, there are differences in the way in which the items are illustrated. In some dictionaries, for example the *Bilingual Dholuo–English Dictionary Kenya*, the lemma **bird** in the Dholuo–English section of the dictionary is illustrated according to the Dholuo word for the particular bird, whereas in the other dictionaries different kinds of birds are illustrated under the single lemma **bird**.

In showing how a dictionary may facilitate the learning of the lexicon of a foreign language, the concept **basket** (as known in English) shall be considered and compared to the same broad concept in Dholuo. In English, there is only

one generic word, but, as indicated below, Dholuo has different names for different baskets depending on their shape and function.

In addition to this, the concept **cattle** (termed *cattle* in the English-speaking world) will also be considered. It will be compared with the way in which the Batswana lexicalized this concept, and the way a Setswana dictionary has, by making use of illustrations, dealt with the range of lexical items for *cattle*.

Basket types and their functions among the Luo people of East Africa

According to Ocholla-Ayayo (1980: 138), it is thought that, in an attempt to establish the Luo cultural pattern, one important method is to look at the socio-cultural or socio-economic function of the manufactured objects. The Luo do not use a basket for more than one purpose. Below is a list of a few of the many different types of baskets of the Luo people and their various functions.

okapu	Basket made by weaving from the same material, having a close interwoven pattern and used for transportation, especially of vegetables.
aduda-biye	Market basket made by weaving and intertwining.
atonga	Basket made of wicker in a twinned weave, commonly used for transporting harvested grain from the fields.
odheru	Dish-shaped basket similar to a large tray, made of wicker and used for transporting and winnowing grain.
osech-winy	Non-return trap, used for catching birds.
osech-rech	Non-return cage, used for fishing.
lweru	Cradle-shaped basket, made of sisal ropes and used for fishing.
rimba or *gogo*	Large, long fish basket or net, used for fish drives.

Without using pictorial illustrations, it would obviously be impossible to explain the differences between these various baskets. The illustrations below have been reduced to their fundamentals and would therefore be compact enough to be included in a dictionary. To a large extent these illustrations are realistic and even if a particular basket has never been seen, this lack of prior knowledge would not hinder an understanding of what is being defined. The illustrations are also simple and precise. In addition to this, they have clarity and are easily distinguished. They therefore meet the criteria for pictorial illustrations as proposed by Al-Kasimi (1977). Furthermore, they possess distinguishing properties and thus establish the concept **basket** more efficiently than a verbal equivalent would be able to do. Moreover, the verbal equivalents require an uneconomical number of words and may not show the spatial or sequential relationships effectively.
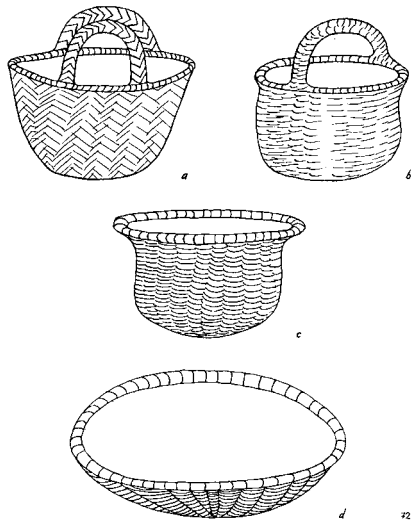
Figure 1: Baskets for general use
   a. A vegetable basket (*okapu*)
   b. A market basket (*aduda-biye*)
   c. A basket for transporting grain from the fields (*atonga*)
   d. A winnowing tray (*odheru*)
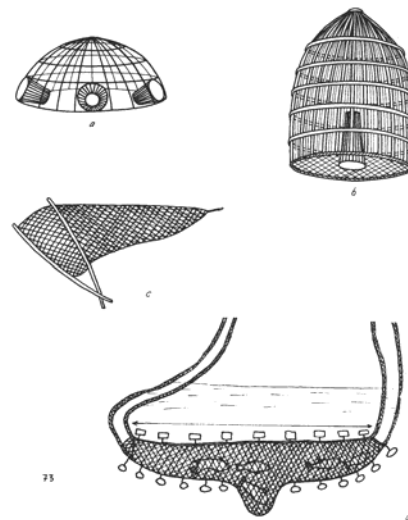
Figure 2: Baskets with special uses
   a. A bird trap (*osech-winy*)
   b. A fish basket (*osech-rech*)
   c . A cradle-shaped basket (*lweru*)
   d. A large long fish basket (in Lango) or fish net (in Kadem and other southern Luo tribes) (*rimba* or *gogo*)

Cattle types of the Batswana

According to Breutz (1991: 15), the Batswana people exclusively breed the long-horn cattle for the size of their horns and their colour which arouses the enthusiasm of the men, chiefly serve as sacrificial animals and supply the essential portion of bride wealth. This possession constitutes the economic wealth of their owner. The cattle have different names depending on the colour of and patterns on their skin. In *Thanodi ya Setswana* the different kinds of cattle are included in the back matter of the dictionary. The lexicographer has used full colour photographs of the various kinds of cattle so that the dictionary user will clearly see the differences between the various kinds of cattle.

Due to the fact that the names of the different kinds of cattle are derived from the colour of and patterns on their skins, the use of photographs provides realistic illustrative material which has fidelity. The illustrations are simple, and not open to double interpretation. They are precise as the user's attention is directed only to the colour of and patterns on the skins of the cattle, which is what is relevant to the specific concept. The illustrations are also completed by captions providing the names of the particular kinds of cattle. Furthermore, the

dictionary user easily distinguishes the features of the illustrations thus indicating that they have clarity.

As they cue and reinforce the verbal equivalents, especially when the dictionary user can identify and respond to the photographs, these illustrations serve the functions proposed by Gropper (1963). They also serve as examples as several different but relevant photographs have been given in order to establish the concept of the different names for different kinds of cattle depending on the colours of and patterns on their skins. Landau (1989: 112) states that photographs are necessary in the case of unidealized individual things. From these photographs, it can clearly be seen how the use of pictures depicting different concepts, particularly in African language monolingual and bilingual dictionaries, can go a long way in actually showing the dictionary user the differences that would be very difficult to describe verbally.

According to Bergenholtz and Tarp (1995: 156), encyclopaedic sections may be explicitly or implicitly integrated in the word list, depending on whether direct cross-references are made from the individual dictionary articles or on whether such references are understood. However, in the *Thanodi ya Setswana*, there are no cross-references to or from the central word list. The illustrations may therefore be included to serve a supplementary function and are comparable to the colour plates in earlier encyclopaedias.

The use of these two examples does, however, not mean that objects such as baskets and cattle are the only items that should be illustrated in dictionaries. There are also other cultural artefacts as well as initiation and religious rituals and marriage and cleansing ceremonies.

Conclusion

On the basis of the research described above, it can be concluded that not only pedagogical dictionaries do and should include illustrations. Other dictionary types may also include illustrations, for instance, bilingual and multilingual dictionaries with significant differences between the cultures of the speakers, for instance, the *South African Multi-language Dictionary and Phrase Book,* and, in addition to this, monolingual dictionaries for languages embedded in cultures that have not given way to popular culture but have retained their customs, traditional norms and values like the Luo people of East Africa and the Setswana people of Southern Africa. This is important due to the fact that dictionaries for many African languages are being compiled for the first time and that because of limited resources, these languages do not have different types of dictionaries, for instance, overall descriptive, standard or pedagogical dictionaries. For many of the African languages, a standard monolingual dictionary is still non-existent. As a result, African language dictionaries have to serve many purposes simultaneously. Their target audience ranges from mother-tongue speakers and learners, to children and scholars. For these reasons, it would

make these dictionaries accessible to a wider audience if lexical items for cultural concepts could be illustrated.

According to Al-Kasimi (1977: 99), the past experiences of the perceiver play an important role in the process of perception when he/she is familiar with the object pictured. But when the object pictured is new to the perceiver, it is the stimulus properties of the picture which will be the determining factor in the formation of the concept. Accordingly, pictorial illustrations can be employed in the bilingual dictionary to illustrate objects familiar to the user as well as those cultural items which are peculiar to the foreign learner.

On the basis of this study, it can be said that, in general, pictorial illustrations add to the communicative value of dictionaries, help in bridging the semantic gaps that may occur between languages and save space which would otherwise be taken up by long descriptions of a lemma.

The relative limited research done on pictorial illustrations in the lexicographical, particularly African language lexicographical field, made it difficult to limit theories and principles to lexicography per se and a great deal of borrowing had to be done from other fields. The scope of the study had also to be limited as this is a broad field that needs much more research. This study touches on the basis of inclusion of pictorial illustrations in dictionaries, particularly African language dictionaries which are still being developed for some of the languages.

However, further analysis of pictorial features should be done in order to develop and cultivate the culture of the inclusion of pictorial illustrations in dictionaries so that their explanatory value will be realised by lexicographers.

Bibliography

Dictionaries

Abraham, R.C. 1958. *Dictionary of Modern Yoruba.* London: University of London Press.

Benson, T.G. (Ed.). 1964. *Kikuyu–English Dictionary.* Oxford: Oxford University Press.

Botne, R. and A.T. Kulemeka. 1995. *A Learner's Chichewa and English Dictionary.* Cologne: Rüdiger Köppe.

Brown J.T. 1925. *Setswana Dictionary Setswana–English/English–Setswana.* Gaborones: Botswana Book Centre.

Bryant, A.T. 1956. *An Abridged English–Zulu Wordbook.* Eighth Edition. Mariannhill: Mariannhill Mission Press.

Bullen, H.B. 1951. *Zande English Dictionary.* London: The Sheldon Press.

Capen, C.A. 1998. *Bilingual Dholuo–English Dictionary Kenya.* Tucson: Carole A. Capen.

Cuenod, R. 1991. *Tsonga–English Dictionary.* Braamfontein: Sasavona Publishers and Booksellers.

Corbeil, J.C. and A. Archambault. 1994. *The Macmillan Visual Dictionary: Multilingual.* Montreal: Éditions Quèbuec/Amérique Inc.

Cowie, A.P. (Ed.). 1989. *Oxford Advanced Learner's Dictionary of Current English.* Fourth Edition. Oxford: Oxford University Press.

Dekker, A.M and J.H. Ries. 1958. *Woordeboek Afrikaans–Zoeloe Zoeloe–Afrikaans/Isichazimazwi Isibhunu–IsiZulu IsiZulu– Isibhunu.* Johannesburg: Afrikaanse Pers.

De Schryver, G.-M. and N.S. Kabuta. 1997. *Lexicon Cilubà–Nederlands.* Ghent: Research Centre for African Languages and Literatures (Recall), University of Ghent.

De Schryver G.-M. and N.S. Kabuta. 1998. *Lexicon Cilubà–Nederlands.* Ghent: Research Centre for African Languages and Literatures (Recall), University of Ghent.

Doke, C.M., J.M Malcolm, J.M. Sikakana and B.W. Vilakazi. 1990. *English–Zulu/Zulu–English Dictionary.* Johannesburg: Witwatersrand University Press.

Education and Training ISCOR. 1966. *Woordeboek Afrikaans–Fanakalo/Dictionary English–Fanakalo/ Lo Buk Ka Lo Izwi-Fanakalo–Sibulu–Singis.* Pretoria: ISCOR.

Eise, B. and W.H. Haacke. 1999. *Khoekhoegowab Language of Damara, Haillom and Nama.* Windhoek: Gamsberg Macmillan Publishers.

Fisher, A., E. Mdala, S. Tshabe and E. Weiss. 1985. *English–Xhosa Dictionary.* Cape Town: Oxford University Press.

Grobler, G.M, J.C. le Roux, L.E. Jennings, R.H. Moeketsi and P.C. Taljaard. 1995. *The Concise Multilingual Dictionary.* Johannesburg: Ad Donker Publishers.

Kgasa, M.L.A. and J. Tsonope. 1995. *Thanodi ya Setswana.* Gaborones: Longman Botswana.

Kriel, T.J. 1965. *Pukuntšu Noord-Sotho–Afrikaans/Afrikaans–Noord-Sotho Woordeboek.* Pretoria: Dibukeng.

Kriel, T.J. 1976. *Popular Northern Sotho Dictionary. Northern Sotho–English/English–Northern Sotho.* Pretoria: J.L. van Schaik.

Kriel, T.J. 1977. *Pukuntšu Noord-Sotho–Afrikaans/Afrikaans–Noord-Sotho Woordeboek.* Pretoria: J.L van Schaik.

Kriel, T.J. 1988. *Popular Northern Sotho Dictionary.* Third Edition. Pretoria: J.L. van Schaik.

Librarius Felicitias. 1986. *Learner's English–Venda Dictionary Illustrated in Colour.* Johannesburg: Librarius Felicitias.

Louwrens, L.J. 1994. *Dictionary of Northern Sotho Grammatical Terms.* Pretoria: Via Afrika.

Mantzel, D. and B. Schulz (Eds.). 1997. *Illustrated School Dictionary for Southern Africa.* Cape Town: Francolin Publishers.

Matumo, Z.I. 1993. *Setswana–English–Setswana Dictionary.* Gaborone: Botswana Book Centre and Macmillan Botswana Publishers.

McLaren, J. 1963. *A New Concise Xhosa–English Dictionary.* Cape Town: Maskew Miller Longman.

Nesbitt, R. 1996. *Chambers-Macmillan South African Dictionary, Junior Primary, Illustrated.* Manzini: Macmillan Boleswa Publishers.

Paroz, R.A. 1961. *Southern Sotho–English Dictionary.* Revised and Enlarged Edition. Morija: Morija Sesuto Book Depot.

Prinsloo, D.J. and B.P. Sathekge. 1996. *New Sepedi Dictionary English–Sepedi (Northern Sotho)/Sepedi (Northern Sotho)–English.* Pietermaritzburg: Shuter and Shooter.

Prinsloo, D.J. and B.P. Sathekge. 1997. *Popular Northern Sotho Dictionary.* Cape Town: Pharos Dictionaries.

Prinsloo, D.J, B.P. Sathekge and E. Kapp. 1997. *Nuwe Sepedi Woordeboek Afrikaans–Sepedi (Noord-Sotho/Sepedi (Noord-Sotho)–Afrikaans.* Pietermaritzburg: Shuter and Shooter.

Reynierse, C. (Ed.). 1996. *South African Multi-language Dictionary and Phrase Book.* Cape Town: Reader's Digest Association of South Africa.

Rycroft D.K. 1981. *Concise SiSwati Dictionary SiSwati–English / English–SiSwati.* Pretoria: J.L van Schaik.

Sanderson, G.M. 1954. *A Dictionary of the Yao Language.* Zomba: The Government Printer.

Snyman, J.W., J.S. Shole and J.C. le Roux. 1990. *Dikisinare ya Sestwana–English–Afrikaans Dictionary / Woordeboek.* Pretoria: Via Afrika.

Summers, D. (Ed.). 1987. *Longman Dictionary of Contemporary English.* New Edition. Harlow: Longman.

Taasisi ya Uchunguzi wa Kiswahili Chuo Kikuu Dar es Salaam. 1981. *Kamusi ya Kiswahili Sanifu.* Nairobi: Oxford University Press.

Taasisi ya Uchunguzi wa Kiswahili Chuo Kikuu Dar es Salaam. 1990. *Kamusi Sanifu ya Isimu na Lugha.* Dar es Salaam: Educational Publishers and Distributors.

Taasisi ya Uchunguzi wa Kiswahili Chuo Kikuu Dar es Salaam. 1996. *TUKI English–Swahili Dictionary. Kamusi ya Kiingereza–Kiswahili.* Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili (TUKI).

Turner, Wm. Y. 1996. *Tumbuka / Tonga–English / English–Tumbuka / Tonga Dictionary.* Blantyre: Central Africana Ltd.

Vallardi, A. 1992. *Italiano–Swahili / Swahili–Italiano.* Italy: Garzanti Editore.

Van Warmelo, N.J. 1989. *Venda Dictionary English–Tshivenda.* Pretoria: J.L van Schaik.

Ziervogel, D. and P.C. Mokgokong. 1979. *Klein Noord-Sotho Woordeboek.* Third Edition. Pretoria: J.L van Schaik.

Other sources

Al-Kasimi, A.M. 1977. *Linguistics and Bilingual Dictionaries.* Leyden: E.J. Brill.

Bergenholtz, H. and S. Tarp (Eds.). 1995. *Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries.* Amsterdam/Philadelphia: John Benjamins.

Breutz, P.-L. 1991. *The Social and Political System of the Sotho–Tswana. Four Generations of Traditional Culture Change.* Ramsgate: P.-L. Breutz.

Brown, H.D. 1987. *Principles of Language Learning and Teaching.* Englewood Cliffs: Prentice-Hall Inc.

Carstens, A. 2001. *Principles of Document Design.* Class Notes. Pretoria: University of Pretoria.

Casson, R.W. 1981. *Language, Culture, and Cognition: Anthropological Perspectives.* New York/London: Macmillan.

Fleming, M. 1967. Classification and Analysis of Instructional Illustrations. *AV Communication Review* 15: 246-258.

Gouws, R.H. 1989. *Leksikografie.* Cape Town: Academica.

Gouws, R.H. 1994. Ostensiewe adressering in vertalende woordeboeke. *Lexikos* 4: 61-85.

Gropper, G.L. 1963. Why a Picture is Worth a Thousand Words. *A V Communication Review* 11: 76-95.

Landau, S.I. 1989. *Dictionaries: The Art and Craft of Lexicography.* Cambridge: Cambridge University Press

Ndoleriire, O.K. 2000. Cross-cultural Communication in Africa. Webb, V. and E. Kembo-Sure (Eds.). 2000. *African Voices. An Introduction to the Languages and Linguistics of Africa*: 268-285. Cape Town: Oxford University Press.

Ocholla-Ayayo, A.B.C. 1980. *The Luo Culture: A Reconstruction of the Material Culture Patterns of a Traditional African Society.* Wiesbaden: Steiner.

Putter, A.M. 1999. *Verklaringsmeganismes in kinderwoordeboeke*. Unpublished M.A. Thesis. Potchefstroom: University of Potchefstroom.

Stein, G. 1991. Illustrations in Dictionaries. *International Journal of Lexicography* 4(2): 99-127.

Swanepoel, P.H. 1990. *Definisies in vakkommunikasie*. Pretoria: HSRC.

Tylor, E.B. 1871. *Primitive Culture: Researches into the Development of Mythology, Philosophy, Religion, Art and Custom.* 2 volumes. London: John Murray.

Zgusta, L. 1971. *Manual of Lexicography*. The Hague: Mouton.

Zgusta, L. 1988. Pragmatics, Lexicography and Dictionaries of English. *World Englishes* 7(3): 243-253.

Oor die verhouding tussen woordeboekstrukture, woordeboek-inhoud en leksikografiese funksies

Rufus H. Gouws, *Departement Afrikaans en Nederlands, Universiteit van Stellenbosch, Stellenbosch, Republiek van Suid-Afrika (rhg@sun.ac.za)*

Opsomming: Navorsing op die gebied van die metaleksikografie het oor die laaste paar dekades 'n betekenisvolle fokusverbreding ondergaan. Die aanvanklike fokus was veral op die woordeboekinhoud. Mettertyd het die fokus verskillende woordeboekstrukture ingesluit en tans val leksikografiese funksies binne die bestek van metaleksikografiese navorsing. Al drie hierdie komponente van die leksikografie is ter sake vir die beplanning en samestelling van gebruiker-gedrewe woordeboeke. In hierdie artikel word aangetoon dat die vertrekpunt in die beplanning van 'n nuwe woordeboek nie in die eerste plek 'n bestaande woordeboek moet wees nie. Die tei-kengebruiker moet die vertrekpunt van die hele leksikografiese proses uitmaak. Dit moet lei tot die vasstelling van leksikografiese funksies, die keuse van leksikografiese data en die bepaling van die relevante woordeboekstrukture. Hierdie strukture moet die data op so 'n manier aanbied dat die gebruiker in staat is om 'n optimale onttrekking van inligting te bereik. Terselfdertyd moet die leksikografiese funksies van die woordeboek probleemloos vervul kan word.

Sleutelwoorde: DATAVERSPREIDINGSTRUKTUUR, GEBRUIKERGEDREWE BENADE-RING, GEBRUIKERSITUASIE, GEBRUIKERSPERSPEKTIEF, GEBRUIKSITUASIE, KENNISGE-RIGTE FUNKSIES, KOMMUNIKASIEGERIGTE FUNKSIES, LEKSIKOGRAAF, LEKSIKOGRA-FIESE DATA, LEKSIKOGRAFIESE FUNKSIES, LEKSIKOGRAFIESE PROSES, TEKSPRODUKSIE, TEKSRESEPSIE, TOEGANGSTRUKTUUR, WOORDEBOEKBEPLANNING, WOORDEBOEKGE-BRUIKER, WOORDEBOEKINHOUD, WOORDEBOEKSTRUKTURE

Abstract: About the Relationship between Dictionary Structures, Dictionary Contents and Lexicographic Functions. Over the last few decades, research in the field of metalexicography has undergone a significant broadening of scope. The original focus had chiefly been on the contents of dictionaries. In the course of time, the focus also included different dictionary structures and the scope of current metalexicographic research accommodates lexico-graphic functions. These three components of lexicography are all relevant in the planning and compilation of user-driven dictionaries. In this article, it is shown that the point of departure in the planning of a new dictionary should in the first instance not be an existing dictionary. The target user should constitute the point of departure of the whole lexicographic process. This should lead to the identification of lexicographic functions, the choice of lexicographic data and the selection of relevant dictionary structures. These structures should present the data in such a way that the user is able to achieve an optimal retrieval of information. At the same time the lexicographic functions of the dictionary should be accomplished in an unproblematic way.

1.   Inleidend

'n Oorsig van die ontwikkeling van die teoretiese leksikografie bied 'n duidelike beeld van die verskuiwing in die relatiewe posisionering van die leksikografie as vakgebied. Tydens die vroeë ontwikkelingsjare van die metaleksikografie, veral in die tydperk kort na die verskyning van Ladislav Zgusta se baanbrekende werk *Manual of Lexicography* (1971), was dit een van die groot deurbrake dat die teoretiese leksikografie die vroeëre skeptisisme van talle linguiste kon afskud om 'n plek te verdien as 'n wesenlike komponent van die taalkunde. Die teoretiese leksikografie, onder meer ook die benadering in Zgusta (1971), was veral gerig op die taalkundige inhoud van woordeboeke en met dié navorsings- en studiefokus het die posisie van die teoretiese leksikografie algaande verbeter tot 'n volwaardige subdissipline van die taalkunde (vgl. in hierdie verband Gouws 2004).

Die taalkunde het taal en die wetenskaplike beskouing van taal as studieterrein. Die teoretiese leksikografie het woordeboeke en 'n wetenskaplike beskouing van woordeboeke as studieveld en in die laaste dekades het hierdie beskouing van woordeboeke in so 'n mate ontwikkel dat dit tans in die leksikografie om veel meer as net die taalkundige aanbod in woordeboeke gaan. In hierdie verband is die navorsingsaanslag van teoretiese leksikograwe aangepas om naas die taalkundige inhoud ook die struktuur en in die onlangse jare ook die funksies van woordeboeke in te sluit. As gevolg van hierdie fokusverbreding wat tot meer as net die taalkundige beskouing van woordeboeke gelei het, het die siening onder teoretiese leksikograwe posgevat dat die leksikografie nie meer 'n subdissipline van die taalkunde is nie, maar dat dit tot 'n volwaardige en selfstandige dissipline ontwikkel het (vgl. in hierdie verband Hartmann en James 1998, Wiegand 1998 en Tarp 2000). Hierdie verandering het gelei tot 'n wesenlike behoefte aan navorsing oor die rol en waarde van woordeboekstrukture en leksikografiese funksies asook hulle onderlinge verhouding en wisselwerking met die inhoud van woordeboeke, woordeboektipologie en die belang van die gebruikersperspektief. Dit is van dié aspekte wat in hierdie artikel ter sprake kom. In hierdie verband is dit egter ook ter sake om aandag te gee aan die leksikograaf en aan die gebruiker- en gebruiksituasie van 'n bepaalde woordeboek.

2.   Die rol van die leksikograaf

'n Opvallende kenmerk van die moderne leksikografie is die aktiewe wisselwerking tussen teorie en praktyk. Gelukkig is baie praktiserende leksikograwe daarvan bewus dat hulle die gehalte van hulle woordeboeke kan verbeter indien die woordeboeke 'n stewige teoretiese basis het. Gelukkig is baie teoretiese leksikograwe daarvan bewus dat die werklike sukses van hulle navorsingsopdrag 'n werklikheid word wanneer dit neerslag vind in die leksikografiepraktyk en dat hulle teorieë op so 'n manier geformuleer moet word dat dit interpreteerbaar is in terme van die eise van die praktyk. 'n Leksikografiese proses is gerig op die uiteindelike produksie van woordeboeke en vir hierdie proses om suksesvol te verloop, mag die rol van die leksikograaf wat die betrokke woordeboek saamstel, nooit onderskat word nie.

Leksikograwe kan in ten minste vier groepe verdeel word: dié wat slegs by die leksikografiepraktyk betrokke is, sonder enige kennis van of belangstelling in die teoretiese leksikografie; dié wat veral by die leksikografiepraktyk betrokke is, maar met 'n kennis van en belangstelling in die teoretiese leksikografie; dié leksikograwe wat veral betrokke is by die teoretiese leksikografie, maar met 'n belangstelling in en ervaring van die praktiese leksikografie, en, laastens, dié leksikograwe wat slegs betrokke is by die teoretiese leksikografie, sonder enige belangstelling in of ervaring van die leksikografiepraktyk. Die sukses van leksikografiese werk op sowel die praktiese as die teoretiese vlak is dikwels afhanklik van die relatiewe posisie van die leksikograaf binne die genoemde vierdeling. Die twee middelste groepe wat die wisselwerking tussen teorie en praktyk verteenwoordig, vorm dikwels die basis vir suksesvolle leksikografiese prosesse. Die twee groepe op die eindpunte van die vierdeling het ook 'n rol om te speel, maar hulle werk het dikwels bykomende ondersteuning nodig om sukses te verseker. 'n Leksikograaf wat geen begrip van die leksikografiese praktyk het nie en wat slegs op 'n abstrakte vlak werk se navorsingsbevindinge het dikwels weinig toepassingswaarde. In so 'n geval gaan dit te maklik slegs om teorie ter wille van die teorie. Die ander uiterste is dat praktiserende leksikograwe te min teoretiese kennis het om suksesvol deel te hê aan die beplanning van innoverende woordeboekprojekte. Hulle kan 'n woordeboek volgens 'n voorgeskrewe woordeboekplan saamstel en kan selfs help met die ontwerp van 'n nuwe leksikografiese proses, maar dan moet daar meestal staatgemaak word op die voorbeeld van reeds bestaande projekte en woordeboeke. Hulle kan vanaf 'n bestaande model na die gebruiker toe werk, maar kan nie op grond van die gebruiker se eiesoortige behoeftes 'n nuwe en gepaste woordeboekkonseptualiseringsplan ontwerp nie. Die uiteindelike sukses van 'n woordeboek word medebepaal deur die kennis en ervaring van die leksikograaf, die leksikograaf se vertroudheid met sowel die leksikografieteorie as die leksikografiepraktyk en 'n erkenning van die noodsaaklikheid van 'n gebruikersgedrewe benadering in die beplanning en samestelling van die woordeboek. Die gebruikersgedrewe benadering vereis ook 'n insig in die tipiese gebruiker- en gebruiksituasie.

3.   Gebruiker- en gebruiksituasies

Wanneer daar gekyk word na 'n woordeboek-spesifieke leksikografiese proses beperk 'n mens te maklik die omvang van dié proses deur veral of selfs uitsluit-lik te fokus op die woordeboeksamestellingsfase van die proses. Die leksiko-grafiese proses sluit egter veel meer as net die woordeboeksamestelling in, want die skryf van die woordeboek word voorafgegaan en gevolg deur 'n reeks fases, onder meer 'n fase waartydens indringend aandag gegee word aan die tipiese situasie waarbinne die betrokke woordeboek gebruik gaan word. Ty-dens die beplanningsfase van 'n woordeboek is dit belangrik om te weet wie die potensiële gebruikers van die woordeboek sal wees, maar ook wat die tipiese gebruiker- en gebruiksituasie sal wees, wat die probleme van die teiken-gebruikers in die bepaalde gebruiksituasies is, watter data aangebied word en ook hoe dit aangebied moet word om die gebruiker optimaal ter wille te kan wees. 'n Vertroudheid met die gebruiker, die gebruiker- en gebruiksituasie is voorvereistes vir besluite oor die relevante leksikografiese funksies van 'n gege-we woordeboek. Hierdie kennis van die gebruiker en die gebruiker- en gebruik-situasie vorm die basis vir die grondlegging van 'n teoretiese benadering waar leksikografiese funksies 'n wesenlike rol speel.

Die hoofmikpunt van die leksikografie as 'n wetenskaplike praktyk is die totstandbrenging van 'n leksikografiese proses wat gerig is op die produksie van woordeboeke wat die teikengebruikers se behoeftes bevredig. Woorde-boeke word saamgestel om as praktiese gebruiksinstrumente te dien en hulle sukses of mislukking word eers bepaal wanneer die werklike gebruikers hier-die woordeboeke in die werklike gebruiksituasie benut. Die beskikbaarstelling van 'n woordeboek aktiveer 'n volgende fase in die leksikografiese proses, te wete die kulturele praktyk van woordeboekgebruik (vgl. Wiegand 1989: 251). Om die suksesvolle gebruik van 'n woordeboek te verseker, moet die woorde-boek op so 'n manier beplan word dat daar voldoende voorsiening gemaak word vir die eise van die tipiese situasie waarin die woordeboek gebruik gaan word.

In sowel die teoretiese as die praktiese leksikografie is daar tot dusver veels te min aandag gegee aan 'n ondubbelsinnige vasstelling van die tipiese gebruiker- en gebruiksituasies van woordeboeke. 'n Leksikografiese benade-ring waar leksikografiese funksies 'n wesenlike rol speel, word daardeur ge-kenmerk dat die woordeboek nie net op 'n spesifieke gebruiker gerig is nie maar ook op 'n spesifieke gebruiksituasie. Die gebruikersituasie het in die meta-leksikografiese literatuur meer aandag gekry as die gebruiksituasie. Tarp (2004: 232) dui byvoorbeeld aan dat die volgende gebruikersituasies ter sake is in aanleerderleksikografie: teksresepsie in die vreemde taal, teksproduksie in die vreemde taal, doelgerigte studie van die grammatikale stelsel van die vreemde taal en 'n doelgerigte studie van die leksikale stelsel van die vreemde taal.

Die gebruiksituasie waaraan nog min aandag gegee is, het onder meer te make met die fisiese omgewing waarin en praktiese omstandighede waaronder

'n woordeboek gebruik moet word. Hierdie faktore behoort 'n invloed te hê op verskillende aspekte van die beplanning van woordeboeke. Die fisiese omgewing waarbinne 'n woordeboek gebruik word, kan byvoorbeeld 'n regstreekse invloed hê op die suksesvolle onttrekking van inligting. Die gebruikers van byvoorbeeld 'n standaard verklarende woordeboek vir moedertaalsprekers van Britse Engels in Engeland sal waarskynlik die woordeboek onder ander omstandighede gebruik as byvoorbeeld 'n laerskoolleerder van Engels as addisionele taal met Tswana as moedertaal wat in 'n plattelandse Suid-Afrikaanse omgewing woon. Die gebruiker van die Britse woordeboek raadpleeg waarskynlik die woordeboek in 'n biblioteek, tuis of in 'n kantoor binne 'n omgewing wat ondersteun word deur 'n goed gevestigde woordeboekkultuur. Daarteenoor kan dit gebeur dat die Tswanaleerder nie op dieselfde mate van woordeboekkultuur kan staatmaak nie en in 'n omgewing verkeer wat veel minder geskik is vir suksesvolle woordeboekgebruik as dié van die Britse gebruiker. By die beplanning van hierdie woordeboek vir die Tswanaleerder moet leksikograwe voorsiening maak vir 'n verskeidenheid gunstige, maar ook minder gunstige gebruiksituasies. Die leksikograaf moet besef dat die woordeboek dalk gebruik word in 'n klaskamer met of sonder die hulp van 'n onderwyser, of tuis met of sonder die ondersteuning van ouers. "Tuis" verwys ook nie noodwendig na 'n goed toegeruste studeerkamer nie. Hierdie gebruiksituasie mag onder 'n boom of in 'n plakkershuis sonder elektrisiteit wees. Waar 'n gebruiksituasie voorsien word met min of geen ondersteuning van onderwysers of ouers nie en veral waar die fisiese omgewing ongunstig is vir maklik lees, bv. as gevolg van swak lig weens gebrek aan elektrisiteit, moet die data wat in die woordeboek verstrek word en die manier waarop dit aangebied word, aangepas word om die gebruiker so goed as moontlik te help. Om seker te maak dat die werklike doel van 'n woordeboek bereik word, is dit noodsaaklik dat daar aan 'n wye spektrum aangeleenthede aandag gegee word in die beplanningsfase van die woordeboek. 'n Woordeboek wat saamgestel word vir gebruikers wat nog nie vlot lesers is nie, moet byvoorbeeld liefs 'n groter lettertipe gebruik as 'n woordeboek wat op vlot lesers gerig is — op 'n soortgelyke manier as in handboeke vir die vroeë grade op skool. Ook oor die aanbod van data en die mate van teksverdigting moet met inagneming van die gebruiker en gebruiksituasie besluit word. Woordeboekuitleg is nie 'n kosmetiese deel van die publikasie van woordeboeke of 'n blote kwessie van die druk van 'n boek nie, maar 'n wesenlike komponent van die leksikografiese proses wat veel daartoe kan bydra dat die werklike doel van 'n woordeboek bereik word, deurdat dit die sukseskoers van die woordeboeknaslaanpogings kan verhoog. Die uitleg van 'n woordeboek mag gevolglik nie slegs aan die uitgewer oorgelaat word nie, want dit is 'n leksikografiese verantwoordelikheid wat verrig moet word as deel van 'n gebruikersgerigte benadering en as 'n reaksie op die spesifieke behoeftes en naslaanvaardighede van die teikengebruikers (vgl. Swanepoel 2001). Die spesifieke gebruiker- en gebruiksituasie van 'n woordeboek kan bepalend wees vir die uitleg daarvan.

In die metaleksikografie benadruk 'n fokus op leksikografiese funksies die gebruikergedrewe benadering, insluitend die gebruiksituasie, baie sterk. Die gebruiksituasie moet op 'n ondubbelsinnige manier omskryf word en sluit veel meer besonderhede in as wat dikwels deur leksikograwe besef of erken word.

4.   'n Tradisionele vertrekpunt in die beplanning van woordeboeke

Die sukses van 'n leksikografiese proses is onder meer afhanklik van die regte vertrekpunt vir die beplanning van die woordeboek. Die eiesoortige aard van verskillende teikengebruikersgroepe is sodanig dat hierdie vertrekpunt, ook met betrekking tot die ontwerp en formulering van 'n woordeboekkonseptua-liseringsplan, liefs nie te sterk op bestaande woordeboeke moet steun nie.

In 'n gebruikergedrewe leksikografiese proses waar die beplanning van 'n woordeboek bepaal word deur die leksikografiese funksies sal die vertrekpunt vir die formulering van die woordeboekkonseptualiseringsplan nie bestaande woordeboeke wees nie — alhoewel bestaande woordeboeke en hulle benutting in werklike gebruiksituasies wel 'n ondergeskikte rol mag speel. Dolezal (1986: 47) verwys na die vroeë Engelse leksikografie as 'n reeks geredigeerde en her-siene tekste. Vir baie jare was daar in die Engelse leksikografie min tekens van werklike vernuwing waar 'n nuwe woordeboek met 'n aanbod of benadering gekom het wat grootliks afwyk van dié van bestaande woordeboeke, en 'n by-drae lewer wat gerig is op die veranderde behoeftes van die teikengebruikers. Leksikograwe het te dikwels in die beplanning van hulle woordeboeke onkri-ties staat gemaak op reeds verskene woordeboeke in stede daarvan om innove-rend te werk te gaan en 'n woordeboek te beplan wat nie op ander woorde-boeke gebaseer is nie, maar regstreeks volg uit die behoeftes van die teikenge-bruikers. Die beplanningsfase van 'n leksikografiese proses moet nie primêr 'n herhaling van dié van 'n bestaande woordeboek of 'n geringe aanpassing daar-van wees nie. So 'n proses wat bloot gebaseer is op 'n ander woordeboek vir die lewering van 'n model vir 'n nuwe woordeboek, word in figuur 1 voorgestel:

Figuur 1

Bestaande woordeboek(e)   ►   Beplanningsfase   ►   Nuwe woordeboek

Hierdie tipe benadering mag tot die produksie van 'n goeie woordeboek lei indien 'n goeie woordeboek as vertrekpunt gebruik is, maar dit beteken nie noodwendig dat die nuwe woordeboek op 'n gepaste manier vir die spesifieke teikengebruikersgroep saamgestel is nie; veral nie wanneer die leksikografiese funksies en teikengebruikers van die twee woordeboeke nie ooreenstem nie. Vir die bes moontlike nuwe woordeboekprojekte is dit wel nuttig om vertroud te wees met bestaande woordeboeke, maar sulke woordeboeke behoort nie 'n beslissende rol in die beplanning te speel nie. Die geïdentifiseerde teikenge-bruikers, hulle behoeftes en naslaanvaardighede en die tipiese gebruiksituasies van die beplande woordeboek behoort die beplanning te rig.

5.   Die gebruiker as vertrekpunt

Wanneer die gebruiker die vertrekpunt van 'n nuwe woordeboek se beplanning is, skep dit die ideale situasie vir die totstandbrenging van 'n bevredigende verhouding tussen leksikografiese funksies, woordeboekstrukture en woordeboekinhoud. Om dit moontlik te maak, is dit belangrik dat die leksikograaf in 'n heel vroeë fase van die woordeboekbeplanning die teikengebruiker van die beplande woordeboek ondubbelsinnig moet identifiseer en uitsluitsel moet kry oor die behoeftes en naslaanvaardighede van hierdie teikengebruiker. Vervolgens moet vasgestel word hoe die beplande woordeboek in die behoeftes van die teikengebruiker kan voorsien en wat die woordeboekgebruiker met die woordeboek moet kan verrig. 'n Antwoord hierop lei tot die vasstelling van die leksikografiese funksies. Die vasstelling van die leksikografiese funksies moet gesien word as grondliggend aan die vroeë beplanning van die woordeboek en verteenwoordigend van dit waarmee die woordeboek as praktiese instrument die teikengebruiker help. Die gebruiker staan sentraal in die vasstelling van die betrokke funksies.

6.   Leksikografiese funksies, woordeboekstrukture en leksikografiese data

Om die gebruiker en die gebruiker se behoeftes as vertrekpunt te hê in die beplanning van 'n nuwe woordeboek noop die leksikograaf om te besin oor die noodsaaklikheid van leksikografiese funksies en die tipe woordeboekstrukture, en om die woordeboekplan dienooreenkomstig op te stel. 'n Leksikografiese funksie (vgl. Tarp 2000: 196) verteenwoordig die hulp wat 'n woordeboek verskaf aan 'n spesifieke gebruiker ter bevrediging van die behoeftes van daardie gebruiker in 'n spesifieke gebruikersituasie. Omdat elke woordeboek op spesifieke behoeftes van spesifieke teikengebruikers gerig is, is dit moeilik om 'n bestaande woordeboek te vind met 'n identiese teikengebruikergroep en identiese behoeftes wat presies ooreenstem met die omstandighede van die nuwe woordeboek. Die moderne metaleksikografie gaan van die standpunt uit dat geen enkele woordeboek alles vir almal kan wees nie. Spesifieke woordeboeke is op spesifieke gebruikers, behoeftes, gebruiker- en gebruiksituasies en leksikografiese funksies gerig. Daarom behoort elke woordeboek uniek beplan te word. In hierdie verband speel leksikografiese funksies 'n beslissende rol.

Binne 'n funksiegedrewe benadering word 'n onderskeid gemaak tussen twee hooffunksietipes, te wete kennisgerigte en kommunikasiegerigte funksies (vgl. Tarp 2000, Bergenholtz en Tarp 2002 en Tarp en Gouws 2004). Kennisgerigte funksies help gebruikers met algemene kulturele en ensiklopediese data, spesiale data oor die onderwerpsveld en data oor die taal. Kommunikasiegerigte funksies help met die oplossing van probleme met betrekking tot teksproduksie in die moedertaal of vreemde taal, teksresepsie in die moedertaal of vreemde taal en die vertaling van tekste vanaf die moeder- na die vreemde taal en omgekeerd.

Die leksikografiepraktyk het in 'n preteoretiese era ontwikkel en in die verlede, selfs die heel onlangse verlede, is baie woordeboeke beplan sonder enige aanduiding van hulle leksikografiese funksies. Desondanks het die data-aanbod in woordeboeke dikwels daartoe gelei dat sekere funksies wel vervul is, maar nie op 'n beplande, gestruktureerde of voorspelbare manier nie. Vandag word 'n vroeë vasstelling van die leksikografiese funksies as 'n verpligte komponent van die beplanningsfase van enige woordeboek-spesifieke leksikografiese proses gesien en woordeboekopstellers moet bewus wees van hierdie belangrike opdrag. In 'n ideaal-tipiese leksikografiese wêreld sou elke woordeboek slegs 'n enkele leksikografiese funksie vervul. Om al die onderskeie funksies in verskillende monofunksionele woordeboeke gerealiseer te kry, sou tot die samestelling van vier of selfs agt woordeboeke van dieselfde tipologiese kategorie gelei het (vgl. in hierdie verband Ščerba 1940, Hausmann 1977 en Kromann et al. 1984). Om praktiese en ekonomiese redes is dit nie moontlik nie en gevolglik word daar vandag dikwels meerfunksionele woordeboeke saamgestel (vgl. bv. Wiegand 1996 se pleidooi vir 'n enkele polifunksionele woordeboek). So 'n woordeboek hoef nie gelyke aandag te gee aan al die genoemde of ook ander leksikografiese funksies nie, maar die leksikograaf moet al daardie funksies wat in die woordeboek aan bod kom, op 'n ondubbelsinnige manier vasstel, omskryf en konsekwent en sistematies toepas. Nadat die funksies vasgestel is, is dit belangrik om te besluit hoe die funksies ten beste in die woordeboek aan die orde kom. In hierdie verband is dit noodsaaklik om noukeurig te let op die verskillende woordeboekstrukture, die datakeuse en die data-aanbod om die uitvoering van die betrokke leksikografiese funksies te verseker. Dit voorvereis weer eens 'n kennis van die gebruikerprofiel en die woordeboekgebruiksituasie.

In die metaleksikografie moet daar steeds daarteen gewaak word om nie een navorsingskomponent so te isoleer dat dit nie meer as deel van die groter navorsingsgebied van die teoretiese leksikografie gesien word nie. Of dit oor byvoorbeeld woordeboekstruktuur, leksikografiese funksies of die inhoud van woordeboeke gaan, bly dit noodsaaklik om steeds die onderlinge verhouding tussen die verskillende komponente in gedagte te hou en nie een komponent te verabsoluteer as sou dit die kernaspek van die leksikografieteorie verteenwoordig nie. 'n Benadering wat net aandag gee aan een van die komponente, bv. woordeboekstrukture of leksikografiese funksies, kan maklik die balans wat binne 'n algemene leksikografieteorie bestaan, versteur en daardeur die ontwikkeling van 'n model vir 'n gebruikersgerigte woordeboek benadeel. Dit is belangrik dat navorsing gedoen word oor die verskillende individuele komponente van die leksikografieteorie, maar die relatiewe verband met die algemene teorie en die posisionering met betrekking tot die ander komponente moet steeds voor oë gehou word.

Die belang van woordeboekstrukture in die studie van die leksikografie mag nooit ontken word nie. Dit is juis een van die fokuspunte in die verskuiwing van die leksikografie vanaf 'n subdissipline van die taalkunde tot 'n volwaardige en selfstandige dissipline in eie reg (vgl. Gouws 2004). Een van die

opwindende benaderings in die metaleksikografie van die afgelope twintig jaar is die vasstelling en uitvoerige beskrywing van 'n groot verskeidenheid woordeboekstrukture (vgl. in hierdie verband veral die werk van Wiegand, bv. Wiegand 1989a, 1989b, 1989c, Hausmann en Wiegand 1989 en Gouws en Prinsloo 2005).

Vir die produksie van suksesvolle woordeboeke, die ontwikkeling van 'n verbeterde leksikografiese praktyk en die aktivering van suksesvolle woordeboekgebruik, is die regte verhouding tussen woordeboekstrukture en leksikografiese funksies van die uiterste belang. Waar navorsing oor veral woordeboekstrukure in die laat tagtiger- en negentigerjare van die vorige eeu aan die orde was, is hierdie navorsing in die laaste dekade aangevul deur navorsing gerig op die rol van leksikografiese funksies. Sommige metaleksikograwe (vgl. Bergenholtz en Tarp 2002), gaan so ver om hulle navorsing te rig op die ontwikkeling van 'n teorie van leksikografiese funksies wat vir hulle die basis van die teoretiese leksikografie vorm. Die gevaar verbonde aan so 'n benadering is die oorbeklemtoning van leksikografiese funksies ten koste van woordeboekstrukture.

Eerder as die verheffing van een komponent tot 'n dominante faktor behoort daar gelet te word op die aktiewe wisselwerking tussen funksies en strukture. 'n Leksikografieteorie behoort nie verskraal te word tot 'n teorie van leksikografiese funksies of 'n teorie van woordeboekstrukture nie. Binne 'n algemene leksikografieteorie moet daar voldoende ruimte geskep word vir sowel leksikografiese funksies as woordeboekstrukture en vir 'n bepaalde verhouding tussen hierdie twee komponente van die teorie.

Tydens die ontwerp van 'n leksikografiese proses moet leksikograwe daarvan bewus bly dat die woordeboekkonseptualiseringsplan sentraal staan in die vroeë stadiums van die beplanningsfase. Die wesenlike kenmerke van enige woordeboek sluit die leksikografiese funksies, die woordeboekstrukture asook die leksikografiese data in (vgl. in hierdie verband Nielsen 2003). 'n Woordeboek bevat leksikografiese data, vertoon bepaalde strukture en moet sekere spesifieke funksies verrig. Die woordeboekkonseptualiseringsplan vereis 'n kombinasie van hierdie drie kenmerke om die sukses van die betrokke woordeboek te verseker. Tradisionele benaderings in die beplanning van woordeboeke het veral sterk gefokus op die inhoud, dit is die data wat in die woordeboeke aangebied moet word. Die aanbod van data in bestaande woordeboeke het meestal as model gedien. Aanvanklik in 'n mindere, maar later in 'n toenemende mate, is daar aandag gegee aan bepaalde strukture van woordeboeke, maar nogmaals is daar veral gewerk met die voorbeeld van bestaande woordeboeke. Ondanks die werk van vroeë metaleksikograwe soos Ščerba (vgl. Ščerba 1940), is daar oor die algemeen baie lank min aandag aan die rol en waarde van leksikografiese funksies gegee. Waar leksikografiese funksies wel ter sprake gekom het, was dit dikwels toevallig en selde die gevolg van 'n beplanningsproses.

Een van die eerste kwessies ter sake vir die leksikograaf in die formulering van 'n woordeboekkonseptualiseringsplan raak die leksikografiese funksies van

die voorgenome woordeboek. Dié funksies word nie arbitrêr gekies nie, maar word deur die gebruikersprofiel bepaal en help die gebruiker om daardie probleme op te los wat aanleiding gegee het tot 'n bepaalde woordeboeknaslaanhandeling. Om te verseker dat bepaalde funksies wel in 'n woordeboek geld, is dit belangrik om hulle vooraf te identifiseer en die beplanning en keuse van woordeboekstrukture en leksikografiese data dienooreenkomstig te doen. Sodra die funksies van 'n woordeboek vasgestel is, kan die data en die strukture gekies word wat nodig is om die funksies te laat realiseer. So 'n benadering werk met die funksies as vertrekpunt in die beplanning van 'n woordeboek en sien die verwesenliking van die funksies as 'n produk van, onder meer, die leksikografiese data en die woordeboekstruktuur. Die leksikografiese strukture speel 'n belangrike rol, want die data-items mag nie op 'n arbitrêre manier georden word nie, maar wel volgens die kriteria wat ontwikkel is om die verwerkliking van die funksies en die werklike doel van die woordeboek te verseker. Waar woordeboekbeplanning op hierdie manier gedoen word, dien 'n bestaande woordeboek nie as vertrekpunt nie, maar die gebruikersgedrewe benadering neem leksikografiese funksies en dus die gebruiker as vertrekpunt. Nogmaals is ondersteuning vanuit die teoretiese leksikografie nodig, soos blyk uit figuur 2:

Figuur 2
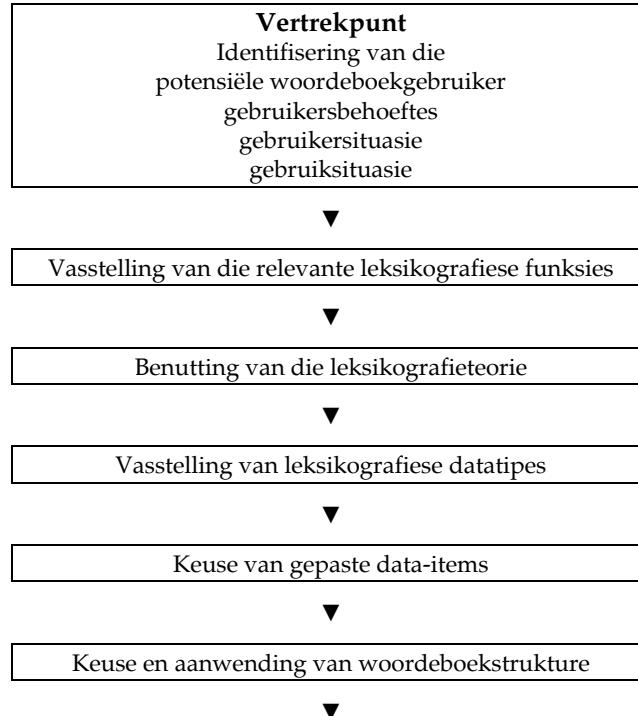
Toepassing van die dataverspreidingstruktuur en die aanbieding van data-items

▼

Totstandbrenging van 'n nuwe woordeboek

'n Sodanige beplanning van 'n woordeboek vereis 'n doeltreffende en behoorlike kennis en benutting van woordeboekstrukture. Een van die uitdagings in hierdie verband aan die leksikograwe is die keuse van strukture wat relevant is vir die betrokke woordeboek. 'n Funksiegedrewe keuse van woordeboekstrukture het implikasies vir die tipe struktuur, maar ook vir die manier waarop die struktuur benut word. Nogmaals moet die gebruikersperspektief deurslaggewend wees in die besluit.

7.   Die keuse van woordeboekstrukture

Wanneer die teikengebruiker, eerder as 'n bestaande woordeboek, die motivering vir die beplanning en samestelling van 'n nuwe woordeboek is, lê een van die uitdagings van die leksikograaf daarin om van die geïdentifiseerde leksikografiese funksies te beweeg na 'n vasstelling van die relevante datatipes, die keuse van data-items en van tersaaklike woordeboekstrukture om die data te kan akkommodeer en leksikografies te kan vergestalt (vgl. weer eens figuur 2). Hier is die wisselwerking tussen leksikografieteorie en -praktyk akuut ter sake. Praktiserende leksikograwe moet vertroud wees met 'n verskeidenheid aspekte van die teoretiese leksikografie, maar dit is ewe belangrik dat teoretiese leksikograwe bewus moet wees van die eiesoortige aard van die uitdagings van die leksikografiepraktyk om hulle navorsing te kan rig op die komplekse probleme van die praktiserende leksikograaf. Die bevindinge van die navorsing van die teoretiese leksikografie moet in terme van die eise van die praktyk interpreteerbaar en toepasbaar wees. Navorsing oor woordeboekstrukture kan te maklik gesien word as 'n uiters abstrakte poging wat geïsoleerd staan van die leksikografiepraktyk. Woordeboekstrukture mag nooit beskou word as die uiteindelike doelwit van leksikografiese navorsing nie. Indien woordeboekstrukture en die navorsing daaroor van die praktyk geskei word, laat dit maklik daardie indruk ontstaan. Woordeboekstrukture moet steeds binne die breër leksikografiese konteks gesitueer word waar dit gesien word as 'n middel tot 'n doel. Hierdie doel is om met die data te kombineer en 'n data-aanbod moontlik te maak wat die vervulling van die relevante leksikografiese funksies kan bewerkstellig. Om dit te verseker, is baie navorsing oor woordeboekstrukture nog nodig sodat werklike leksikograwe wat werklike woordeboeke vir werklike gebruikers saamstel, die werklike probleme die hoof kan bied.

In die beplanningsfase van die leksikografiese proses van enige nuwe woordeboek moet die leksikograwe daarvan bewus wees dat dit nie nodig is om 'n plek te vind vir elk van die talle strukture wat reeds in die metaleksikografiese literatuur geïdentifiseer is nie. Hier speel woordeboektipologie en lek-

sikografiese funksies, en by implikasie dus ook die gebruikersperspektief, 'n bepalende rol. Slegs daardie strukture wat ter sake is vir die vervulling van die vasgestelde leksikografiese funksies van die betrokke woordeboeksoort binne die raamwerk van sy tipologiese kriteria, moet vir aanwending in die woordeboek oorweeg word. In hierdie verband het teoretiese leksikograwe die verantwoordelikheid om die praktiserende leksikograwe te help met die vasstelling van daardie strukture wat oorweeg moet word sodat 'n spesifieke funksie of funksies in 'n woordeboek vervul kan word. Sulke strukture mag nie 'n kosmetiese toevoeging tot die woordeboek wees nie, maar moet meehelp om die vervulling van die relevante leksikografiese funksies te bevredig.

In die studie van woordeboekstrukture kan daar onder meer 'n onderskeid gemaak word tussen kernstrukture (vgl. Mihindou in voorbereiding) en gidsstrukture (vgl. Louw 1997). Die kernstrukture is onder meer die dataverspreidingstruktuur, die makrostruktuur en die mikrostruktuur, en die gidsstrukture is die toegangstruktuur en die mediostruktuur. In die hieropvolgende afdeling word daar aan die hand van voorbeelde uit die leksikografiepraktyk gewys op maniere waarop van hierdie woordeboekstrukture aangewend word. Daar word ook verwys na implikasies daarvan vir die bereiking van bepaalde leksikografiese funksies.

8.   Van teorie na praktyk

Met die gebruiker as vertrekpunt, bepaal die tipologiese kategorie van 'n woordeboek en die leksikografiese funksies watter soort strukture benut word om die woordeboekinhoud weer te gee sodat die vasgestelde leksikografiese funksies bereik kan word. In enige woordeboek is kernstrukture soos die dataverspreiding-, makro- en mikrostruktuur, asook die toegangstruktuur as een van die gidsstrukture, deurgaans ter sake. Die manier waarop hierdie strukture op 'n woordeboek-spesifieke manier aangewend word, is ter sake vir die sukses van die betrokke woordeboek. Vervolgens word daar gekyk na aspekte van woordeboekstrukture in *Nuwe Woordeboek sonder Grense* (voortaan NWSG) (vgl. ook Steyn en Gouws 2005).

NWSG is beplan as 'n multifunksionele woordeboek met veral kommunikasiegerigte, maar ook kennisgerigte funksies. Die kommunikasiegerigte funksies sluit sowel teksproduksie as teksresepsie in. Hierdie woordeboek is saamgestel vir addisioneletaalleerders van Afrikaans in die senior primêre en junior sekondêre skoolfases. Dit is 'n eentalige verklarende woordeboek met 'n tweetalige dimensie, soos blyk uit die insluiting van 'n Engelse ekwivalent van die lemma in elke voorbeeldsin. Vergelyk die volgende artikel:

ons   voornaamwoord ∎
❶ Ons word gebruik wanneer die spreker
na homself en ander mense verwys.
⇨ *Ons [we] gaan môre see toe.* ⇨ *Ons [we]*

> *ry elke dag met die bus skool toe.*
> ❷ Ons word gebruik om te wys dat iets
> aan die spreker en aan een of meer
> ander persone behoort. ⇨ *Ek hoop nie*
> *hulle het al ons [our] lekkers opgeëet terwyl*
> *ons weg was nie.*

NWSG is daarop gerig om sy teikengebruikergroep te help met die lees en verstaan, maar ook die produksie van Afrikaanse tekste. Uit gesprekke met die uitgewer asook met skrywers van handboeke vir die betrokke leerdergroep kon die inskrywingstipes waaraan die betrokke teikengebruikers behoefte het ten einde die vasgestelde leksikografiese funksies te bevredig, geïdentifiseer word. Hierdie datatipes is betekenisparafrases, grammatiese leiding, veral met betrekking tot morfologie, gebruiksvoorbeelde en gebruiksaantekeninge. Gevolglik het die dataverspreidingstruktuur voorsiening gemaak vir die insluiting van inskrywings uit die tersaaklike datatipes. Die fokus was egter nie net op die sentrale teks van die woordeboek nie, maar voorsiening is ook gemaak vir die totstandbrenging van 'n raamstruktuur (vgl. Kammerer en Wiegand 1998, en Gouws 2001; 2004a), met die voortekste-afdeling wat onder meer 'n gebruiksleidingsteks insluit en die agtertekste-afdeling wat die kennisgerigte funksie van die woordeboek ondersteun deur die insluiting van lyste van die Suid-Afrikaanse provinsies en hulle hoofstede, die amptelike tale van Suid-Afrika en die name van die maande van die jaar en die dae van die week. Die woordeboek word 'n politoeganklike bron deurdat die agtertekste-afdeling as buiteteks 'n alfabetiese ekwivalentregister bevat met al die Engelse ekwivalente wat in die sentrale teks se voorbeelde verstrek is. Hierdie ekwivalente word telkens van 'n Afrikaanse inskrywing of inskrywings voorsien wat die gebruiker lei na die lemma van die artikel waarin die betrokke vertaalekwivalent optree. Vergelyk die volgende gedeeltelike artikeltrajek uit die ekwivalentregister:

> **junk;** gemors
>
> **just;** juis, sommer
>
> **justify;** regverdig

Uit die gebruikersprofiel wat vir hierdie woordeboek vasgestel is, het geblyk dat die beplande teikengebruikers nie vertroud is met 'n woordeboekkultuur nie en dat baie van die onderwysers wat hulle behulpsaam moet wees in die tipiese gebruiksituasie nie in 'n posisie is om hulle te help met 'n optimale ontsluiting van die inligting in die woordeboek nie. Hierdie realiteit strook met Carstens (1995) se bevindinge. Die woordeboekstrukture van NWSG moes sodanig wees dat die teikengebruiker daartoe in staat sou wees om 'n suksesvolle woordeboeknaslaanhandeling te kan uitvoer sonder deurlopende bystand; ook omdat die gebruiksituasie sodanig sou wees dat die woordeboek nie altyd in 'n klaskamer gebruik word nie, maar ook tuis waar die gebruiker nie noodwen-

dig op ander hulp aangewese kan wees nie. Gevolglik moes die struktuur een-
voudig en doeltreffend wees. Dit het gelei tot die keuse van 'n streng alfabe-
tiese makrostruktuur wat slegs uit hooflemmata bestaan en geen nes- of nis-
lemmatisering vertoon nie. Dit het 'n geïntegreerde mikrostruktuur waar voor-
beeldsinne in die onderskeie semantiese subkommentare telkens onmiddellik op
die betekenisparafrases volg en waar prosedures van regstreekse lemmatiese
en nielemmatiese adressering geld. Die behoefte aan gebruiksaantekeninge het
die leksikograwe genoodsaak om 'n heterogene artikelstruktuur in te voer wat
naas die verstekartikelstruktuur ook vir 'n meer komplekse uitgebreide ver-
pligte mikrostruktuur voorsiening gemaak het. Gebruiksaantekeninge is gevolg-
lik in tekskassies in spesiale artikelgleuwe geplaas. Twee tipes tekskassies word
onderskei, te wete een wat gemerk is met 'n tekening van 'n by en een wat
gemerk is met 'n tekening van 'n skilpad. Die by waarsku leerders teen tipiese
foute terwyl die skilpad bykomende inligting, veral oor die leksikale item wat
deur die lemmateken voorgestel word, verstrek. Vergelyk die volgende voor-
beeldartikels in hierdie verband:

grootte   naamwoord ■

Die grootte van iets is hoe groot dit is.

⇨ *Wat is die grootte [size] van jou skoene?*

⇨ *Sy weet nie wat die grootte [size] van die*
swembad is nie.

Pasop vir **grootte** en
grote.
⇨ *Daardie vis is 'n grote*
[big one]. ⇨ *Watter*
grootte [size] broek dra jy?

haar[2]   naamwoord ■

(hare, haartjie)

'n Haar is'n lang ding wat uit mense en
sommige diere se velle groei. ⇨ *Borsel*
asseblief jou hare [hair] voordat jy skool toe
gaan. ⇨ *Ons kat verloor baie hare [hair].*

Haar word meestal in
die meervoud gebruik,
dit wil sê: ⇨ *Daardie meisie het pragtige,*
lang hare [hair].

Wiegand (1989a: 427) onderskei twee belangrike kategorieë funksionele teks-
segmente op mikrostrukturele vlak, te wete aanduiders en struktuurmerkers.

Aanduiders is funksionele tekssegmente waaruit die gebruiker inligting oor die voorwerp waarop die woordeboek gerig is of oor die vorm van die woordeboek kan bekom. Daarteenoor is struktuurmerkers inskrywings wat die gebruiker help om aanduiders beter te identifiseer, te onderskei en sistematies en vinnig te vind. Vir die teikengebruiker van NWSG is die toegangstruktuur en die soekgebiedstruktuur van groot belang. Struktuurmerkers help met die soekroete na die verskillende aanduiders waaraan gebruikers die tersaaklike inligting kan onttrek om sowel teksresepsie as teksproduksie moontlik te maak. Duidelik onderskeibare soekvelde in artikels wat 'n mikro-argitektuur vertoon, is nodig om die sukses van die kennis- en die kommunikasiegerigte funksies te verseker. Die volgende artikel het onder meer 'n kennisgerigte funksie, soos blyk uit die betekenisparafrase asook die ingevoegde teks in die tekskassie:

be<u>vel</u>sin   naamwoord ■
(bevelsinne, bevelsinnetjie)
'n Bevelsin is die sin wat jy skryf of sê om vir iemand 'n opdrag te gee. ⇨ *Die volgende sin is 'n voorbeeld van 'n bevelsin [order]: "Moenie water mors nie!" ⇨ In sommige bevelsinne [orders] word asseblief gebruik: "Gee asseblief die plante water!"*

Wanneer 'n mens 'n bevelsin skryf, is daar gewoonlik 'n **uitroepteken (!)** aan die einde van die sin.
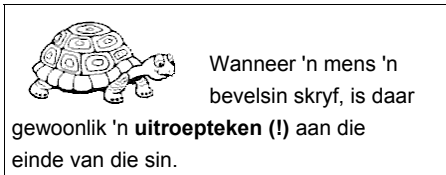
kyk **vraagsin, stelsin** en **uitroep**

Die volgende artikels het sterker kommunikasiegerigte funksies:

uit<u>haal</u>   werkwoord ■
(het uitgehaal)
❶ As 'n mens iets uithaal, bring jy dit buitentoe. ⇨ *Haal asseblief die suiker uit [take out] die kas. ⇨ Die onderwyser raas elke dag met Vuyo omdat hy nie sy hande uit sy sakke haal [take out] nie.*
❷ As 'n mens baie hard probeer, dan haal jy uit. ⇨ *Ons sal vandag moet uithaal [exert ourselves] om die span te klop. ⇨ As julle nie alles uithaal [exert yourselves] nie, sal julle nie slaag nie.*

mak   byvoeglike naamwoord ∎
                        en bywoord

(mak; makker, die makste)

As 'n dier mak is, reageer hulle as mense
vir hulle bevele gee en sal hulle nie
sommer vir 'n mens byt nie. ⇨ *Haar
hond is so mak* **[tame]**, *hy doen alles wat sy
sê.* ⇨ *Jy kan maar aan hierdie leeu vat, hy
is mak* **[tame]**.

◕ teenoor **wild**

Weens die teikengebruiker se onvertroudheid met woordeboeke is daar besluit
op duidelik onderskeibare soekvelde met eksplisiete tipografiese en/of nie-
tipografiese struktuurmerkers, selfs in die geval van soekvelde wat die beteke-
nisparafrase en voorbeeldmateriaal bevat.

Alhoewel die aanduiders wat in NWSG die data-items weergee, nie afwyk
van dié wat in talle ander woordeboeke gevind word nie, speel die struktuur-
merkers 'n belangrike rol om veral die interne toegangstruktuur te verbeter en
'n gebruikersvriendelike soekroete te bewerkstellig. Dit is in regstreekse reaksie
op die behoeftes van die gebruikers. Die mikrostrukturele inhoud en die woorde-
boekstrukture tree saam op om die gewenste leksikografiese funksies te vervul.

9.   Ten slotte

Die aard en omvang van die leksikografiese proses van enige woordeboek mag
nooit onderskat word nie. In hierdie proses moet daar voorsiening gemaak
word vir 'n verskeidenheid fases en aspekte en moet daar gelet word op 'n ak-
tiewe wisselwerking tussen verskillende komponente van die beplande woor-
deboek, maar ook 'n interafhanklikheid wat tussen sommige van hierdie kom-
ponente bestaan. Waar die vroeë leksikografiese navorsing sterk gefokus het op
die inhoud van woordeboeke, die latere navorsing op woordeboekstrukture en
onlangse navorsing op leksikografiese funksies, moet die hedendaagse leksi-
kograaf aan al drie hierdie komponente aandag gee in die beplanning en op-
stelling van 'n woordeboek. Vanuit 'n gebruikersgedrewe benadering is dit nood-
saaklik om reeds vroeg te besluit watter leksikografiese funksies die woorde-
boek moet vervul. Die datakeuse moet in ooreenstemming hiermee gedoen
word en daar moet op woordeboekstrukture en hulle aard besluit word waarin
die data op 'n sodanige manier aangebied kan word dat dit die teikengebruiker
in staat stel om 'n optimale inligtingsonttrekking te verwesenlik. Om dit moont-
lik te maak, is die dataverspreidingstruktuur en die totstandbrenging van 'n
soekgebiedstruktuur en 'n mikro-argitektuur binne die woordeboekartikels van
deurslaggewende belang.

Bibliografie

Bergenholtz, H. en S. Tarp. 2002. Die moderne lexikographische Funktionslehre. Diskussions-beitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände ver-stehen. *Lexicographica* 18: 253-263.

Carstens, A. 1995. Language Teaching and Dictionary Use: An Overview. *Lexikos* 5: 105-116.

Dolezal, F. 1986. How Abstract is the English Dictionary? Hartmann, R.R.K. (Red.). 1986. *The History of Lexicography*: 47-55. Amsterdam: John Benjamins.

Gouws, R.H. 2001. Der Einfluß der neueren Wörterbuchforschung auf einen lexikographischen Gesamtprozeß und den lexikographischen Herstellungsprozeß. Lehr, A. et al. (Reds.). *Sprache im Alltag*: 521-531. Berlyn: De Gruyter.

Gouws, R.H. 2004. Milestones in Metalexicography. Van Sterkenburg, P.G.J. (Red.). 2004. *Linguistics Today — Facing a Greater Challenge*: 187- 205. Amsterdam: John Benjamins.

Gouws, R.H. 2004a. Outer Texts in Bilingual Dictionaries. *Lexikos* 14: 67-88.

Gouws, R.H. en D.J. Prinsloo. 2005. *Principles and Practice of South African Lexicography.* Stellen-bosch: African Sun Media.

Gouws, R.H., M. Stark en L. Gouws. 2004. *Nuwe Woordeboek sonder Grense.* Kaapstad: Maskew Miller Longman.

Hartmann, R.R.K. en G. James. 1998. *Dictionary of Lexicography.* Londen: Routledge.

Hausmann, F.J. 1977. *Einführung in die Benutzung des neufranzösischen Wörterbücher*. Romanistische Arbeitshefte 19. Tübingen: Max Niemeyer.

Hausmann, F.J. en H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. (Reds.). 1989–1991: 328-360.

Hausmann, F.J. et al. (Reds.). 1989–1991: *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie.* Handbücher zur Sprach- und Kommunikationswissenschaft 5.1–5.3. Berlyn/New York: Walter de Gruyter.

Kammerer, M. en H.E. Wiegand. 1998. Über die textuelle Rahmenstruktur von Printwörter-büchern: Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-237.

Kromann, H.-P., T. Riiber, P. Rosbach. 1984. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. Wiegand, H.E. (Red.). *Studien zur neuhochdeutschen Lexikographie*: 159-238. Hildesheim: Georg Olms Verlag.

Louw, P.A. 1997. *Kriteria vir ʼn standaard vertalende woordeboek.* Ongepubliseerde MA-skripsie. Stellenbosch: Universiteit van Stellenbosch.

Mihindou, G.-R. In voorbereiding. *A Theoretical Model for a Yipunu–English–French Explanatory Dictionary of Medical Terms.* Ongepubliseerde D.Litt.-verhandeling. Stellenbosch: Universiteit van Stellenbosch.

Nielsen, S. 2003. Changes in Dictionary Subject Matter. Wiegand, H.E. (Red.). 2003. *Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache I*: 109-114. Tübingen: Max Niemeyer Verlag.

Ščerba, L.V. 1940. Towards a General Theory of Lexicography. *International Journal of Lexicography* 8 (4): 315-350, 1995.

Steyn, M. en L. Gouws. 2005. *Nuwe Woordeboek sonder Grense*: A Typological and Communicative Bridge. *Lexikos* 15: 164-178.

Swanepoel, P. 2001. Dictionary Quality and Dictionary Design: A Methodology for Improving the Functional Quality of Dictionaries. *Lexikos* 11: 160-190.

Tarp, S. 2000. Theoretical Challenges to LSP Lexicography. *Lexikos* 10: 189-208.

Tarp, S. 2004. Basic Problems of Learner's Lexicography. *Lexikos* 14: 222-252.

Tarp, S. en R.H. Gouws. 2004. Wie leer wat uit Afrikaanse (aan)leerderwoordeboeke. *Tydskrif vir Geesteswetenskappe* 44(4): 276-298.

Wiegand, H.E. 1989. Der gegenwärtige Status der Lexikographie. Hausmann, F.J. et al. (Reds.). 1989–1991: 246-280.

Wiegand, H.E. 1989a. Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. Hausmann, F.J. et al. (Reds.). 1989–1991: 409-462.

Wiegand, H.E. 1989b. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: Alphabetische Anordnungsformen und ihre Probleme. Hausmann, F.J. et al. (Reds.). 1989–1991: 371-409.

Wiegand, H.E. 1989c. Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. Hausmann, F.J. et al. (Reds.). 1989–1991: 462-501.

Wiegand, H.E. 1996. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. Wiegand, H.E. (Red.). *Wörterbücher in der Diskussion II.* 1989–1991: 1-82. Tübingen: Max Niemeyer Verlag.

Wiegand, H.E. 1998. *Wörterbuchforschung.* Berlyn: De Gruyter.

Zgusta, L. 1971. *Manual of Lexicography.* Den Haag: Mouton.

Lexicography and Its Interdisciplinary Contacts, with Special Reference to Linguistics and Onomasiology[*]

R.R.K. Hartmann, *School of English, University of Exeter, Exeter, United Kingdom (R.R.K.Hartmann@exeter.ac.uk)*

Abstract: The article attempts a first conspectus of what is known about interdisciplinary contacts between the fields of Lexicography and Linguistics, and in the process asks a range of fundamental questions on whether and how the subject matter is shared, and how we can improve mutual relations. Firstly, some opinions and arguments in the literature are reviewed, with particular attention to (a) what criteria determine the status of a discipline, (b) how various relations between Lexicography and Linguistics can be modelled, (c) how practice and theory interact, and (d) what parallels there are between the two disciplines, in terms of such notions as description, codification and remediation. Secondly, views from the position of Linguistics are categorised historically and systematically, first by tracing the development from historical-comparative or diachronic linguistics (philology) to structural-descriptive (synchronic) linguistics as well as applied linguistics, and then by combining five of the linguistic 'levels' with the three semiotic 'dimensions', and relating them to various interdisciplinary or problem-solving fields, such as sociolinguistics and computational linguistics. Thirdly, several points of view in Lexicography are presented, and an explicit framework is developed for displaying its place relative to its 'mother', 'sister' and 'daughter' disciplines as well as some of its 'data-supplying' disciplines, paying special attention to the problem of methodology. Fourthly, as a special case study, the territory of 'onomasiology' is discussed to illustrate the limitations and possibilities of various approaches to the analysis, description and lexicographic presentation of synonym and antonym sets of vocabulary. The conclusion is reached that ways must be found to understand the interdisciplinary nature of Lexicography, and a plea is made to move from mutual neglect to deliberate collaboration. (Several diagrams are used to display networks of relations; bibliographical references focus on recent work and include representative reference works for the major disciplines mentioned.)

Keywords: LEXICOGRAPHY, METALEXICOGRAPHY, THEORY AND PRACTICE, DESCRIPTION, CODIFICATION, REMEDIATION, APPLICATION, INTERDISCIPLINARY CONTACTS, LINGUISTICS, ONOMASIOLOGY, MOTHER DISCIPLINES, SISTER DISCIPLINES, DAUGHTER DISCIPLINES, DATA-SUPPLYING DISCIPLINES, METHODOLOGY, REFERENCE SCIENCE, SEMIOTICS

**Opsomming: Leksikografie en sy interdissiplinêre skakels, met spesiale ver-
wysing na linguistiek en onomasiologie.** Die artikel probeer 'n eerste oorsig gee van
wat bekend is oor interdissiplinêre skakels tussen die gebiede van die Leksikografie en die Lin-
guistiek, en vra in die proses 'n reeks grondliggende vrae oor of en hoe die vakinhoude gedeel
word, en hoe ons onderlinge verwantskappe kan verbeter. Eerstens word 'n aantal opvattings en
argumente in die literatuur beskou, met besondere aandag aan (a) watter kriteria die status van 'n
dissipline bepaal, (b) hoe verskillende verwantskappe tussen die Leksikografie en die Linguistiek
gevorm kan word, (c) hoe praktyk en teorie in wisselwerking is, en (d) watter parallelle daar tussen
die twee dissiplines bestaan ingevolge sulke begrippe soos beskrywing, kodifikasie en remedi-
ering. Tweedens word beskouings uit die hoek van die Linguistiek histories en sistematies gekate-
goriseer, eers deur die ontwikkeling na te spoor vanaf die histories-vergelykende of diachroniese
linguistiek (filologie) tot die struktureel-beskrywende (sinchroniese) linguistiek sowel as die toege-
paste linguistiek, en dan deur vyf van die linguistiese "vlakke" met die drie semiotiese "dimensies"
saam te voeg, en hulle in verband te bring met verskillende interdissiplinêre of probleemoplos-
sende gebiede, soos die sosiolinguistiek en die rekenaarlinguistiek. Derdens word verskeie gesigs-
punte in die Leksikografie aangebied, en 'n noukeurige raamwerk ontwikkel om sy plek te toon
met betrekking tot sy "moeder"-, "suster"- en "dogter"-dissiplines sowel as sommige van sy "data-
verskaffende" dissiplines deur spesiale aandag aan die probleem van metodologie te gee. Vierdens,
as 'n spesiale gevallestudie, word die terrein van die "onomasiologie" bespreek om die beperkinge
en moontlikhede van verskillende benaderings tot ontleding, beskrywing en leksikografiese aan-
bieding van sinoniem- en antoniem-woordeskatgroepe toe te lig. Daar word tot die gevolgtrekking
gekom dat maniere gevind moet word om die interdissiplinêre aard van die leksikografie te
verstaan, en 'n pleidooi word gelewer om van onderlinge verwaarlosing tot doelbewuste samewer-
king te beweeg. (Verskeie diagramme word gebruik om verwantskapsnetwerke te toon; bibliogra-
fiese verwysings fokus op onlangse werk en sluit verteenwoordigende naslaanwerke vir die ver-
melde hoofdissiplines in.)

Sleutelwoorde: LEKSIKOGRAFIE, METALEKSIKOGRAFIE, TEORIE EN PRAKTYK,
BESKRYWING, KODIFIKASIE, REMEDIËRING, TOEPASSING, INTERDISSIPLINÊRE SKAKELS,
LINGUISTIEK, ONOMASIOLOGIE, MOEDERDISSIPLINES, SUSTERDISSIPLINES, DOGTER-
DISSIPLINES, DATAVERSKAFFENDE DISSIPLINES, METODOLOGIE, VERWYSINGSWETEN-
SKAP, SEMIOTIEK

1.   Introduction

After many years of being involved with both lexicography and linguistics, I
have an active interest in the double issue of what kinds of disciplines they are
and how they relate to each other (love and hate? ivory tower and coal-face?).
Other fields of potential relevance that will also be mentioned here, in passing,
are lexicology, terminology, language learning, translation, and information
technology.

The aims of this article are as follows: (a) to review the strengths and
weaknesses of the literature (especially from the points of view of both lexico-
graphy and linguistics), (b) to examine mutual relations between them, with

special attention to 'models' explaining these, (c) to illustrate all this with reference to the special case of onomasiology, and (d) to discuss the implications of these topics for interdisciplinary collaboration. In the process, a number of basic reference works claiming to cover the field(s) will also be highlighted.

2.   Opinions, arguments and models

What kinds of relations can be observed between lexicography and linguistics? In an area where personal opinion still tends to overshadow solid scholarly evidence, we need to sift through the statements made from various positions. Such statements are, in any case, rather limited and selective; they depend on the theoretical framework within which respective authors work, e.g. are they concerned with the 'lexicon' as the linguistic study of vocabulary or the product of lexicographic compilation? Are they interested in the 'lexeme' as a basic unit of linguistic analysis or as one of the prototypical information carriers in the dictionary? And although the activities in question may have very long traditions, e.g. the descriptive grammar of Panini in ancient India (*Lexicon Grammaticorum* ed. Stammerjohann 1996: 692-697) or the even older bilingual thematic word-lists produced in ancient Mesopotamia (McArthur 1986: 24-25), the literature for both fields is relatively recent (for linguistics only five or six generations, for lexicography only two or three generations back), and publications on the interdisciplinary relations between them are still rather recent and rare indeed. But it is still worth looking for models that might help us throw light on this topic; four are presented here.

(1) The first aspect for which we need diagrammatic enlightenment is the double question of what constitutes a 'discipline', and what are the criteria for its scholarly status. One of the few authorities who have addressed these issues head-on is the German linguist and metalexicographer Wiegand who (in 1998, citing Posner) stipulated five such standards for the delineation of 'dictionary research'. Fig. 1 displays the implications of this approach for the two subject fields of linguistics and lexicography.

Fig. 1:   Criteria for establishing an academic discipline

	Linguistics	*Lexicography*
subject-matter	language analysis …	dictionary compilation …
set of perspectives	variety, change …	typology, history …
research methods	informant testing …	data-gathering …
body of knowledge	linguistic units …	lexical definition …
modes of discourse	terminology …	terminology …

Lexicography, according to Wiegand, has some way to go until it can claim to be a fully developed scholarly discipline, based on a coherent subject-matter

(e.g. compiling reference works), consistent perspectives (e.g. history and ty-pology of reference works), replicable research methods (e.g. collecting data from informants or texts), a specific body of knowledge (e.g. information cate-gories and how they are presented in a reference work), and convincing modes of discourse (e.g. conferences and journals). Interestingly, Lutzeier (2002: 6) has recently made similar comments about the less than universal acceptance of lexicology as a branch of linguistics, stipulating six criteria: subject matter, investigative interest, appropriate methodology, tools utilised, active partici-pants, and public relations.

(2) A second model is required for explaining the whole range of positions from one extreme of complete dependence (Fig. 2a: lexicography considered as part of linguistics, but not usually linguistics as part of lexicography) through relative interdependence (Fig. 2b: mutual relations) between the two subjects to relative independence (Fig. 2c) of each of them.
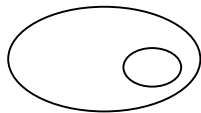
Fig. 2a: Inclusion   **Fig. 2b:** Mutuality   **Fig. 2c:** Independence
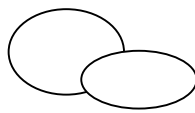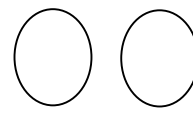
The first view (Fig. 2a) is suggested by a minority of linguists who have de-fined lexicography as "a branch of applied linguistics" (Svensén 1993: 1) or even as "a branch of applied lexicology" (*Dictionary of Language and Linguistics* comps. Hartmann and Stork 1972: 129 and also *A First Dictionary of Linguistics and Pho-netics* comp. Crystal 1980: 210). The second (Fig. 2b) is held by those who are in favour of healthy interdependence, e.g. "lexicography and linguistics are now inextricably mixed" (Béjoint 1994: 177) for the reason that "precision and quality of work" in the former often has to rely on evidence provided by the latter (Al-var Ezquerra 1995: 193-194 and 2003: 366). The third position (Fig. 2c) is taken either by those lexicographers who are suspicious of academic linguists (Lan-dau 1984/2001) or by theoretical and applied linguists who manage to ignore lexicography (and lexicology) altogether (*Encyclopedic Dictionary of Applied Lin-guistics* eds. Johnson and Johnson 1998).

(3) To complicate matters further, both lexicography and linguistics exhibit two sides or aspects of their field, 'practice' and 'theory'. How can this fact be mod-elled? Linguistics and lexicography indeed have a lot in common in terms of practical activities and theoretical perspectives, but traditionally the former has always had a more academic image, while the latter, as we have already noted, has sometimes been called (merely) an 'art and craft' or an 'applied field'. While there are now many more textbooks, monographs, journals, and conference proceedings available for lexicography, the number of dictionary research cen-

tres is still very small, and there are even fewer university professorships (most of them in neighbouring subjects such as linguistics and modern languages). Fig. 3 depicts the potential relations between practice and theory in both disciplines.

Fig. 3:   Linguistic/lexicographic practice and metalinguistic/metalexicographic theory

Linguistics/lexicography

Practice   Theory

(a) – (c)   --------------≫ meta   (i) – (vi)

feedback ≪---------
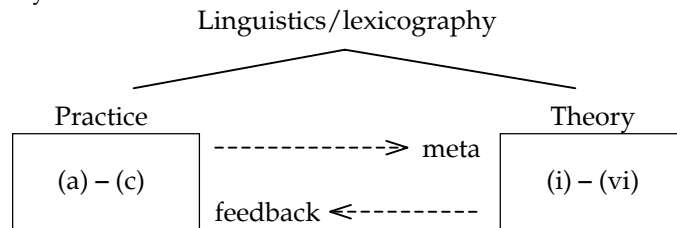
Practical activities include such multi-stage operations as (a) observation or data-gathering (in linguistics said to include both 'fieldwork' and 'intuition', in lexicography 'recording' and 'corpus collection'), (b) description (called 'analysis' in linguistics and 'editing' in lexicography), and (c) presentation (called 'system of rules' in linguistics and 'publishing' in lexicography). The theoretical concerns lead to such perspectives as (i) criticism (prescription/dictionary criticism), (ii) classification (linguistic variety/dictionary typology), (iii) history (language change/dictionary history), (iv) structure (linguistic levels/information categories), (v) use (linguistic usage/dictionary use) and (vi) automation (computational linguistics/dictionary IT).

The path from practice to theory is usually one of model-building and generalisation (metalinguistics/metalexicography), while the opposite route, from theory to practice, may be termed 'feedback' and assumed by implication to lead to improvements in the linguistic/lexicographic process(es). This direction is sometimes called 'application', but the term can also have several other connotations (see below).

(4) There are certainly parallels between linguistics and lexicography, then, as indicated in the diagrams above, but are there suitable models for their systematic interaction? One attempt I made myself many years ago (Hartmann 1979: 2) was to suggest that there seems to be a gradual progression from (linguistic) theory and description to codification, even remediation and practical problemsolving. Another contribution to the topic of linguistics vis-à-vis other disciplines concerned with codification (such as orthography, grammar and lexicography) was written by Haugen from the point of view of language planning. In a paper first published in 1980 and reprinted as part of a collection, Haugen (1987: 60) stated that "to some degree linguistics owes its existence to the practical services linguists could offer as codifiers of language", and that "the typi-

cal product of codification has been a prescriptive orthography, grammar, and dictionary".

Fig. 4: From (linguistic) description to (lexicographic) codification (and remediation)

theory	*description*	*codification*	*remediation/problem-solving*
discourse	textology	textography	style; rhetoric
sentence(s)	grammar	grammatography	composition; language learning
vocabulary	lexicology	lexicography	definition; translation
letter(s)	graphology	orthography	writing; spelling
sound(s)	phonology	orthoepy	pronunciation; speech therapy

From left to right, the columns in Fig. 4 indicate these three kinds of utilisation of linguistic expertise: 'description' (observation of individual and collective linguistic usage), 'codification' (production of prescriptive works, e.g. usage guides, grammars, and dictionaries) and 'remediation' (application of linguistic knowledge to problem-solving areas such as speech therapy, spelling, and translation). From bottom to top these are arranged in order of structural levels, from small/simple to large/complex units of linguistic communication (and also roughly in the historical order in which they have become established).

This tabulation is not complete, however. Thus, the line starting with 'vocabulary' could be supplemented, for instance, by (technical) terminology, terminography and standardisation; more needs to be said, too, about the role of vocabulary in language learning and teaching. Some of the special fields mentioned in Fig. 4 have become labels for professional disciplines represented in and promoted by national and international societies. Thus, in a recent submission to a Higher Education exercise (May 2004), the British Association for Applied Linguistics (BAAL) has stressed that (a) "Applied Linguistics cannot truly be considered as a branch of Linguistics", (b) "the two disciplines often have markedly different research practices", and (c) the field of applied linguistics embraces quite a range of 'subject areas' from child language to sign language, including lexicography and lexicology (Carter 2004: 2-4). Similarly, within the International Association of Applied Linguistics (AILA) and its triennial congresses, some 25 'scientific commissions' have been active for a number of years, including Lexicography and Lexicology, Language Planning, Rhetoric and Stylistics, Interpreting and Translation, Contrastive Linguistics and Error Analysis, Literacy, and Second Language Acquisition.

Most of these disciplines, and some of the mutual contacts and interactions between them, are covered in two book series, first *Current Trends in Linguistics* (14 volumes published by Mouton in the 1960s and 1970s) and then the comprehensive *Handbücher zur Sprach- und Kommunikationswissenschaft* (HSK), some in multi-volume sets, issued by Walter de Gruyter since the early 1980s. 25 HSK volumes have been published, and 11 more are in preparation, on such

fields as semiotics (4 volumes), (the history of) linguistics, morphology, syntax, semantics, lexicology (2 volumes), lexicography (3 volumes), onomastics (2 volumes), languages for special purposes (2 volumes), sociolinguistics (2 volumes), psycholinguistics, computational linguistics, media studies, and translation (some of these will be referred to below).

Two other handbook series covering some of these disciplines, but on a much more limited scale, are available from Oxford University Press and Blackwell Publishing. In one of the Blackwell volumes, *The Handbook of Applied Linguistics* (eds. Davies and Elder 2003), the useful distinction is made between 'linguistics applied' (L–A) and 'applied linguistics' (A–L), the former being concerned with real-world language data for understanding language and evaluating linguistic theory (e.g. in language description, lexicography, corpus linguistics and stylistics), while the latter covers problem-based approaches for understanding language use and remedying social problems (e.g. in language learning, language for special purposes, language testing and language planning).

Most of these subjects are comprehensively covered in the *Kenkyusha Dictionary of Applied Linguistics* (ed. Koike 2003). It is worth noting, from the perspective of this article, that the 2nd edition of the *Encyclopedia of Language and Linguistics* (first published 1994), currently in preparation, is intended to have a much broader and deeper coverage of the field of Lexicography. For more on the linguistic standpoint, see Section 3 below; for that of lexicography, see Section 4; for the special case of onomasiological lexicography, see Section 5.

3.   The view from Linguistics

What kind of discipline is linguistics? As has been noted above, it can tend either more towards theoretical model-building or practical problem-solving, but both of these are grounded in a long tradition, as so-called 'diachronic linguistics' and 'synchronic linguistics'. The history of the subject is well documented in Arens (1955/1969) in terms of (a) prescientific endeavours associated with Ancient Greece and Rome and the Medieval and Modern periods, (b) 19th-century philological scholarship and (c) various phases and schools of 20th-century structural, descriptive and cognitive linguistics. One of the most important influences on the development of lexicography was exerted by comparative-historical linguists who, according to Collison (1982), stimulated much of the work on so-called historical dictionaries for many of the Indo-European languages, such as the *Deutsches Wörterbuch* by the Grimm brothers and the *Oxford English Dictionary* by James Murray et al.

There has always been interaction with other academic disciplines, of course, e.g. biology (especially during the 19th century), philosophy (during the early 20th century), and more recently the so-called 'hyphenated disciplines' of sociolinguistics (*The Handbook of Sociolinguistics* ed. Coulmas 1998), psycholinguistics (*Psycholinguistik/Psycholinguistics. An International Handbook*

eds. Deutsch et al. 2003), ethno- and anthropological linguistics, and 'lexicomputing' in the form of corpus linguistics and computational linguistics (Dodd 1989, Ooi 1998, and the *Oxford Handbook of Computational Linguistics* eds. Mitkov et al. 2003).

Over the years, the hierarchical relationships between various linguistic levels and units, as specified in Fig. 4 above, have been further developed into component branches of the discipline: from sounds to phonetics and phonology (and orthoepy), from writing to graphetics and graphology (and orthography), from words to morphology and lexicology (and lexicography), from sentences to grammar (and grammatography), from discourse to textology (and stylistics and rhetoric), from language variety to dialectology, from language affiliation to linguistic typology, etc. Throughout this time, some linguists have also shown a regular concern for making themselves useful by 'applying' their knowledge to fields like language planning, language teaching and translation, which in turn has given rise to new specialisations, e.g. contrastive linguistics and languages for special purposes (LSP).

In Fig. 5 (based on a diagram in Hartmann 1991: 2855), five of the levels of linguistic structure are combined with the three dimensions distinguished in semiotics, resulting in 15 divisions of descriptive (and contrastive) linguistics that can then be used in the codification of usage information in monolingual or bilingual dictionaries. Conversely, information presented in dictionaries can be used by linguists for verifying usage data.

Fig. 5:   Levels of linguistic analysis

paradigmatic dimension	*syntagmatic dimension*	*pragmatic dimension*
text semantics	text syntax	text pragmatics
morphology	syntax	pragma-grammar
lexical semantics	semotactics	pragma-lexicology
segmental graphology	graphotactics	pragma-graphology
segmental phonology	phonotactics	pragma-phonology

For example, at the level of lexicology it may be possible to distinguish the different meanings of a polysemous word in terms of its membership of a particular lexical field, collocation or context, although there is no guarantee that the resulting linguistic explanation is automatically the best way for presenting its senses in the dictionary. Neither can we be sure that words considered by linguists as having similar meanings within the language (intralingual synonyms) or as being translatable into another language (interlingual equivalents) are necessarily the ideal candidates for lexicographic treatment in a thesaurus. Thus, within the field labelled 'contrastive lexicology' by Wikberg (1983), the vocabulary handled in the practical bilingual dictionary may not correspond either to the 'theoretical lexicon' as defined in linguistics or the 'mental lexicon' as studied in psycholinguistics. On the other hand, Martin (2003) has suggested

that lexicology and lexicography could be brought together by systematically linking the translation equivalents found in bilingual dictionaries for several language pairs.

Linguists can therefore be shown to be both 'donors' to and 'users' of the work of lexicographers. Thus, it is possible not only to note whether and in what way(s) dictionary compilers depend on previous linguistic studies, e.g. for deciding what kinds of semantic information to present in the dictionary (Hartmann 1972, Svensén 1993), but also to verify how linguists in turn utilise lexicographic information for checking the evidence of the meaning range (e.g. in definitions and contextual examples) or the etymological background (e.g. origin and semantic change) of certain vocabulary items (Harras 1989, Hanks 2003).

4.   The view from Lexicography

What kind of field is lexicography? In its practical manifestation of compilation or dictionary-making, it has a very long tradition which has already been documented in a number of monographs: McArthur (1986) has pursued the universal anthropological background, Boisson et al. (1991) have investigated the early beginnings of both the Western and Eastern traditions of monolingual and bilingual lexicography, Starnes and Noyes (1946/1991) have traced the story of English dictionary-making from Cawdrey to Johnson, and Van Hoof (1995) has shown that the compilation of bilingual dictionaries is often tied to the practice of translation. The currently best textbook on the compilation process is by Landau (1984/2001), who in its revised and updated 2nd edition covers most aspects of planning, managing and editing dictionaries, including the changes brought about by computer technology. The most ambitious attempt so far to survey the whole field is the three-volume encyclopedia *Wörterbücher/Dictionaries/Dictionnaires* (eds. Hausmann et al. 1989–1991).

Lexicography as theory, on the other hand, is a very recent phenomenon. In one of the most critical overviews of the status of metalexicography, Wiegand (1998) has sketched out several of its perspectives or branches, most notably the study of dictionary use. In a three-volume reader (Hartmann 2003), I have myself used these headings for grouping the 70 texts reprinted there into nine parts: compiler perspectives, critical perspectives, user perspectives, historical perspectives, regional perspectives, linguistic perspectives, typological perspectives, structural perspectives, and interdisciplinary perspectives. In one of the contributions to the final (ninth) part based on the chapter on methodology in another recent book, I proposed a model to show how lexicography is related to a number of other disciplines. In Fig. 6 below, which is based on that diagram (Hartmann 2001: 123), these are variously referred to as 'mother', 'sister' and 'daughter' disciplines as well as 'data-supplying' disciplines.

Fig. 6: Interdisciplinary contacts of lexicography

```
                   ┌─────────────────────────────────────┐
                   │        MOTHER DISCIPLINES           │
                   │ Linguistics   Semiotics  (Reference │
                   │            Science)                 │
                   └─────────────────────────────────────┘
                                   │
                   ┌─────────────────────────────────┐
                   │    Dictionary Research          │
                   │ its modes of discourse:         │
                   │    pub. organs, terminology …   │
┌──────────────────┤ its methods:                    ├──────────────────┐
│ SISTER           │    survey, text analysis …      │ DATA-SUPPLYING   │
│ DISCIPLINES      │ its perspectives:               │ DISCIPLINES      │
│ Information       │    historical, critical, struc- │ Linguistics      │
│ Technology        │    tural …                     │ Philosophy       │
│ Library science   │ its body of knowledge:         │ History, Law     │
│ Language teaching │    dictionary history, diction-│ Social sciences  │
│ Terminology       │    ary use …                   │ Art, Music       │
│ Onomastics        │ its subject matter:            │ Medicine         │
│ Media studies     │    Lexicography and its        │ Sciences         │
│   (etc.)          │    working process             │ Technology       │
└───────────────────┤                                │   (etc.)         │
                   └─────────────────────────────────┘──────────────────┘
                                   │
                   ┌─────────────────────┐
                   │      DAUGHTER        │
                   │    DISCIPLINES       │
                   │ Indexing             │
                   │ Word-processing      │
                   │ Printing             │
                   │ Publishing           │
                   │ Marketing            │
                   │ Translating          │
                   │   (etc.)             │
                   └─────────────────────┘
```

Lexicography, both as practical compilation process and as theoretical dictionary research, is displayed in the central box, taking account, in reverse order, of Wiegand's five criteria for disciplinarity as indicated in Fig. 1 (Section 2) above.

The *mother disciplines* comprise at least three:

(1)   Linguistics, since Meier (1969) argued, well before the rise of applied linguistics as a fashionable new field, that the sceptical love-hate relationship between linguists and lexicographers should be replaced by a more sensible attitude of constructive give and take, e.g. in terms of such issues as determining the word-list in the historical dictionary, discriminating senses within entries, and sifting and treating specific etymological, grammatical and lexical information. More recently, some of us have made the case for more bridge-building in these respects (Hartmann 1996, Hartmann 2001/2003);

(2)   Semiotics, since Rey-Debove (1971) showed that lexicographers have to work within a wider frame-work of language, communication and information. Some more progress has been achieved along these lines in the intervening years (see the handbook *Semiotik/Semiotics* eds. Posner et al. 1998–2004);

(3)   'Reference science', since McArthur (1998: 218) saw it as an overarching, but not yet fully established specialisation and defined it as "all aspects of organizing data, information, and knowledge in any format whatever, for any purpose whatever, using any materials whatever".

At least seven *sister disciplines* are relevant:

(1)   Information technology, a powerful force that has extended the frontiers of lexicography as well as many linguistic topic areas, involving all systems and tools for the collection, processing and presenting of reference data (Wilks et al. 1996);

(2)   Library science, especially those aspects of it that concern the provision of 'reference services' (Katz 1998);

(3)   Language teaching and related problem-solving areas, which are explored in the *Longman Dictionary of Language Teaching and Applied Linguistics* (comps. Richards et al. 1992) and the *Routledge Encyclopedia of Language Teaching and Learning* (ed. Byram 2000), although their entries on lexicographic notions tend to be rather superficial, ignoring recent work on the user perspective;

(4)   Terminology, in relation (and sometimes in opposition) to technical or special-purpose lexicography, is treated in the *Handbook of Terminology Management* (eds. Wright and Budin 1997/2001) and the international handbook *Fachsprachen/Languages for Special Purposes* (eds. Hoffmann et al. 1998/1999);

(5)   Onomastics, the study of personal and place names, remains an often neglected specialisation, well surveyed in *Name Studies. An International Handbook of Onomastics* (eds. Eichler et al. 1995);

(6)   Media studies, a collective term for the field which investigates newspapers, radio, television and film, often together with the technologies on which they are based (*The Complete A–Z Media and Communication Studies Handbook* comp. Price 1997);

(7)   Translation, the field concerned with the theory of interlingual mediation and the training of translators (*The Routledge Encyclopedia of Translation Studies* ed. Baker 2001).

Whether the contact is with 'mother' or 'sister' disciplines, given sufficiently frequent and intensive collaboration, new bridge subjects may develop. At the intersection between language teaching and lexicography, for instance, we have seen the emergence of 'pedagogical lexicography'; at the point of contact with information technology, the new field of 'computational lexicography' (or dictionary IT) has become established, which in turn interacts with linguistics under such banners as computational lexicology and corpus linguistics. Other areas of potential interdisciplinary linkage already referred to include sociolinguistics, psycholinguistics, translation studies, LSP/terminology and language testing, each of them sometimes occurring in combination with applied linguistics and/or information technology.

To illustrate the potential links with psycholinguistics, the English linguist Crystal (1986) has argued that in our search for better dictionaries, we need more observational evidence on what constitutes better dictionary makers as well as better dictionary users. In relation to LSP and terminology, Kretzenbacher (2002) has demanded more attention to interdisciplinary and intercultural dimensions in the study of technical vocabulary, particularly in respect of the issue of whether lexical relations within and between scientific disciplines are 'open or 'closed'. And at the borderline between pedagogical and computational lexicography, Campoy Cubillo and Safont Jordà (2004) have explored the exciting potential of electronic dictionaries in foreign-language learning.

At the bottom of Fig. 6 appear those *daughter disciplines* which provide a service function for the various organisational and technological processes that lexicographers are charged to perform, such as typography, text manipulation (including pictorial illustrations) and — especially for bilingual dictionaries — translation. Finally, on the right of the page there is an incomplete and potentially infinite list of *data-supplying disciplines* which are basic for dictionary projects, especially those devoted to the production of technical (encyclopedic or terminological) lexica. Some of these, like law and medicine, have very long traditions, others are more recent, such as linguistics (for usage dictionaries and dictionaries of linguistics), history (for dictionaries of historical events and personalities), even lexicography itself (for dictionaries of lexicography, e.g. the *Dictionary of Lexicography* ed. by Hartmann and James 1998/2001).

A good example of how these various mother, sister, daughter and data-supplying disciplines interact is the dissertation on bilingual technical lexicography by Wang (2001): it lies at the crossroads of lexicography, LSP and contrastive linguistics and shows how German–English dictionaries of electrical engineering are classified, structured and used, providing empirical details previously unavailable. What this study — and some of the ones cited in the previous paragraphs — demonstrate is the need for more information on how lexicography can be improved, not least by appealing to corpus evidence. This process is, of course, even more complex if aimed at the production of bilingual reference works, which would require parallel text corpora in addition to an understanding of the interdisciplinary links with translation. A case study of

this topic is provided by Kwong et al. (2004) who show how the sister disciplines of terminology and IT can be brought to bear on the extraction of technical terms from two languages (English and Chinese), the data-supplying discipline being law. However, it could be argued (Hartmann 2004) that the results might be even more reliable if the text corpus was not based on the sentence-by-sentence alignment of parallel texts which are translated court judgements, but on the expert comparison of terms found in independently formulated discourse which is similar both in content and context.

We must also take note of the less obvious fact, already hinted at in Sections 2 and 3 above, that interdisciplinary links are often mutual. Just as the lexicographer typically processes data based on the subject matter of other disciplines, the practitioners of these disciplines regularly use dictionaries as information sources, whether as scholars while checking facts or as trainees while learning basic concepts. For an example of various forms of collaboration not only in the direction from linguistics (as 'applied theory') to lexicography, but also from lexicography (as 'common sense' practice) to linguistics and other 'basic' and 'neighbouring' (or mother and sister) disciplines, see Herbst et al. (2004), one of several volumes in 'Lexicographica Series Maior' devoted to interdisciplinary bridge-building.

5.   The special case of Onomasiology

What kind of field is onomasiology? It lies right between lexicography and linguistics, since it is concerned with the problem of 'meaning' and its treatment in the dictionary, both in terms of intralingual sense relations between vocabulary items within a language and of interlingual equivalence relations between vocabulary items in two or more languages. This is usually put in the form of the bipolar distinction between 'semasiology' (or sign interpretation) and 'onomasiology' (or name allocation). The traditional general monolingual dictionary is semasiological in the sense that it provides an explanation of the meaning(s) of a given word by means of definitions or examples (from word to meaning — for helping native speakers 'decode' or comprehend). More specialised onomasiological dictionaries such as thesauruses, on the other hand, start with a given concept or meaning and provide a range of lexical choices for expressing that notion (from meaning to word — for assisting with 'encoding' or composition).

It is not surprising, then, that the linguistic disciplines of semantics and lexicology as well as the whole of lexicography have claimed this important territory, in order to understand, model and improve the presentation of sense relations, for the benefit of native speakers and foreign learners. Various theories and models have been proposed to describe such lexical properties (componential analysis, semantic fields, frame semantics and prototype semantics), and various reference works have tried to codify their essential features (nom-

enclators, promptoria, usage guides and vocabulary builders), none with fully satisfactory completeness. One recurring example is the domain of 'seats' (*chair*, *bench*, *couch*, *sofa* etc.), which has been investigated, decomposed and tabulated by structural, anthropological and cognitive linguists (like Gipper, Pottier, Fillmore, Wierzbicka and Burkhanov), and referred to by the (meta-)lexicographer Van Sterkenburg (1992: 524) as an "evergreen from structuralism". But long before all this tentative modelling in linguistic semantics, lexicographers from Girard to Roget had to cope with the notion of overlapping meanings in their synonym dictionaries and thesauruses. The history of Roget's *Thesaurus* and some of his predecessors, contemporaries and successors is documented in Hüllen (2004), but a comparative study of such dictionaries in various languages is still outstanding.

A number of observations can be made on the development and nature of such onomasiological reference works. Most are monolingual, but for the more popular languages bilingual synonym dictionaries exist, too; there is even a polyglot comparative *Dictionary of Selected Synonyms in the Principal Indo-European Languages* (Buck 1949), which deals with the word field of 'seats' in terms of one entry, *chair* (within the notion 'Dwelling, House, Furniture', between *bed* and *table*), listing its equivalents in Greek, Latin, French, English, German and 20 other languages, but the comment on those selected says that "not included are words which may denote any kind of 'seat'…".

For English, a recent survey (Hartmann forthcoming) elicited 158 titles for the 20th century alone. Most of these are alphabetic, but a minority in the thesaurus tradition are arranged thematically into taxonomic clusters, e.g. 'couches and sofas' would be treated within the class 'the home' and the sub-class 'furniture'. Some early synonym dictionaries attempted to discriminate individual members of conceptually related groups of words (they are therefore called 'distinctive'), but more recent products, partly under the influence of Roget's *Thesaurus*, tend to merely enumerate them in ('cumulative') list form, thus making it almost impossible to know exactly what distinguishes, say, a couch from a sofa. Sometimes the boundary between a semasiological dictionary and an onomasiological thesaurus is deliberately blurred, e.g. in the French tradition by the so-called 'analogical' dictionary or in the English tradition by the increasingly popular 'dictionary and thesaurus' hybrid, especially since it has become relatively easy to combine the two genres electronically (Calzolari 1988).

During the last few years, there has been a convergence of such disciplines as (case) grammar and (frame) semantics, both within 'cognitive' linguistics and in relation to lexicography (and metalexicography), all supported by IT. This trend is documented in at least three recent thematic issues of the *International Journal of Lexicography* 3(4), 13(4), 16(3) and 17(2). Thus, summarising developments since *WordNet*, a psycholinguistically inspired lexicon of semantic associations which specifies synonyms, hypernyms, hyponyms, holonyms, meronyms and other coordinate terms for most of the more frequent vocabu-

lary items, Fontenelle (2000: 231) introduces four papers "on the representation of collocational, thesauric and semantic knowledge and on the extraction of such knowledge from [electronic] language resources". More recently, Atkins et al. (2003: 356-357), assessing the potential uses of *FrameNet*, another American approach to computer-aided semantic analysis, conclude that "(t)he potential value of this resource for lexicographic practitioners is surely greater than that of any currently available way of accessing corpus data". And in the current issue of the *International Journal of Lexicography*, four (Dutch) models of multilingual databases are discussed in their relevance to the compilation of bilingual dictionaries, e.g. Janssen (2004) on the potential uses of 'concept lattices' as illustrated in Fig. 7.

Fig. 7:   A synonym network
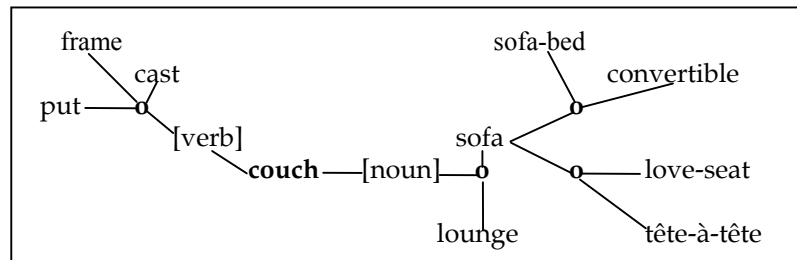
The synonym cluster associated with the word *couch* is shown here in the kind of semantic constellation diagram that can be found in the displays of the *Visual Thesaurus*[†], first as verb with the meaning 'to formulate', leading on to at least three possible synonyms (*put*, *frame* and *cast*) and then as a noun with several distinct meanings, one of which ('upholstered seat for more than one person') leads on to the synonyms *lounge* and *sofa*, and from the latter further to 'upholstered seat for more than one that can be changed into a bed' with two synonyms (*sofa-bed* and *convertible*) and to 'small sofa for two' with two more (*love-seat* and *tête-à-tête*).

    This is one of several modes of presentation that are becoming available in the early 21st century, although it is difficult to keep track of the emerging formats and genres. The volume edited by Campoy Cubillo and Safont Jordà (2004) makes a contribution to the terminological classification, qualitative evaluation and pedagogical integration of these digital reference media. Thus, Tono (2004) distinguishes five types of electronic interfaces (CD-ROM, Hyperlink, Pop-up mode, Parallel format, and Pocket e-dictionary); other contributions survey the advantages and limitations of specific digital and Web-based products. The new generation of CD dictionaries, for example the *'Smart' Thesaurus* associated with the *Cambridge Advanced Learner's Dictionary* and the *'Genie' Thesaurus* linked to the *Oxford Advanced Learner's Dictionary* and the *Oxford Collocations Dictionary for Students of English*, can be used for word searches, text searches

and sound searches in both decoding and encoding contexts, e.g. during word-processing activities. Thus, for the word *couch* a range of synonyms and other related words will appear in the former, from *armchair* and *chaise longue* to *lounge* and *sofa*, together with definitions and examples. In the latter, collocational information is also included, as well as links to illustrations and study pages.

For users with appropriate IT skills, these reference experiences are promising indeed. However, instruction and persistence is often required to exploit them fully on a practical level, and ideally the 'new' material needs to be tested against both user practices for their effectiveness and corpus evidence for their reliability. To verify exactly what denotational and connotational differences there are between words like *couch* and *sofa*, for example, would in turn require more interdisciplinary contacts with such fields as semantics, sociolinguistics and computational linguistics.

6. Conclusion

A number of links between lexicography and linguistics have been explored and quite a few issues have been raised in this article, e.g. what does 'applied' mean in terms of disciplinary status, methodology, problem-solving and improving practice? Many of these are yet to be addressed, but already we have seen some possibilities of theoretical and practical advances, with special attention to onomasiological lexicography and information technology.

A number of desiderata have emerged. The emphasis has been on English and the European context, but a global and comparative view is also needed (cf. Huang 1994); we further need to re-evaluate older methodologies and explore new ones, e.g. corpus technology; several neglected areas must be developed, e.g. linguistic versus encyclopedic knowledge, general versus technical vocabulary; lexicographic training and dictionary instruction must be improved; most of all, we need to move from mutual ignorance and distrust to more intensive interdisciplinary collaboration.

All this will have important implications in the future for academic institutions, dictionary publishers and dictionary users.

Endnote

†   www.visualthesaurus.com, published by Plumb Design (1998–2004) as an electronic adaptation of *WordNet*, developed at Princeton University by Miller et al. (1990). In the *Visual Thesaurus*, the basic meanings indicated graphically by small circles are correlated with definitions which appear in boxes when pointed at by the mouse; they can also be called up and compared with related ones in a marginal column on the right side of the screen.

Bibliography

Dictionaries and other reference works

Cambridge Advanced Learner's Dictionary ed. by Patrick Gillard. 2003. Cambridge: Cambridge University Press [with CD-ROM 'Smart' Thesaurus].

(The) Complete A–Z Media and Communication Studies Handbook comp. by Stuart Price. 1997. London: Hodder Arnold.

Deutsches Wörterbuch comp./ed. by Jacob Grimm and Wilhelm Grimm (et al.). 1852–1960. (16 volumes) Berlin/Göttingen: Hirzel.

Dictionary of Language and Linguistics comp. by R.R.K. Hartmann and F. Colin Stork. 1972. Barking: Applied Science Publishers.

Dictionary of Lexicography comp. by R.R.K. Hartmann and Gregory James. 1998/2001. London: Routledge.

(A) Dictionary of Selected Synonyms in the Principal Indo-European Languages comp. by Carl D. Buck. 1949. Chicago: University Press.

Encyclopedia of Language and Linguistics ed. by R.E. Asher. 1994. (10 volumes) Oxford: Pergamon. [2nd edn by E. Keith Brown in preparation; 14 volumes projected by Elsevier.]

Encyclopedic Dictionary of Applied Linguistics ed. by Keith Johnson and Helen Johnson. 1998. Oxford: Blackwell.

Fachsprachen/Languages for Special Purposes. An International Handbook of Special Language and Terminology Research ed. by Lothar Hoffmann et al. 1998, 1999. HSK Vols. 14.1, 14.2. Berlin: W. de Gruyter.

(A) First Dictionary of Linguistics and Phonetics comp. by David Crystal. 1980. The Language Library. London: A. Deutsch.

(The) Handbook of Applied Linguistics ed. by Alan Davies and Catherine Elder. 2003. Oxford: Blackwell Publishing.

(The) Handbook of Sociolinguistics ed. by Florian Coulmas. 1998. Oxford: Blackwell Publishing.

Handbook of Terminology Management ed. by Sue Ellen Wright and Gerhard Budin. 1997, 2001. (2 volumes) Amsterdam: J. Benjamins.

Kenkyusha Dictionary of Applied Linguistics/Oyogengogaku Jiten ed. by Ikuo Koike. 2003. Tokyo: Kenkyusha.

Lexicon Grammaticorum. Who's Who in the History of World Linguistics ed. by Harro Stammerjohann. 1996. Tübingen: M. Niemeyer.

Lexikologie/Lexicology. An International Handbook on the Nature and Structure of Words and Vocabularies ed. by D. Alan Cruse et al. 2002. HSK Vol. 21.1. Berlin: W. de Gruyter.

Longman Dictionary of Language Teaching and Applied Linguistics 2nd edn comp. by Jack C. Richards et al. 1992. Harlow: Longman.

Name Studies. An International Handbook of Onomastics ed. by Ernst Eichler et al. 1995/1996. HSK Vols. 11.1, 11.2. Berlin: W. de Gruyter.

Oxford Advanced Learner's Dictionary of Current English 6th edn by Sally Wehmeier. 2000. Oxford: Oxford University Press [with CD-ROM 'Genie' Thesaurus]

(The) Oxford English Dictionary comp./ed. by James A.H. Murray (et al.). 1884–1928. (1st edn 10 volumes) 1933. (Supplement); comp./ed. by Robert Burchfield. 1972–1986. (4-volume Supplement); comp./ed. by John Simpson and Edmund Weiner. 1989. (2nd edn 20 volumes) Oxford: Clarendon Press.

(The) Oxford Handbook of Computational Linguistics ed. by Ruslan Mitkov. 2003. Oxford: Oxford University Press.

Oxford Collocations Dictionary for Students of English ed. by Margaret Deuter et al. 2003. Oxford: Oxford University Press [with CD-ROM 'Phrasebuilder'].

Psycholinguistik/Psycholinguistics. An International Handbook ed. by Werner Deutsch et al. 2003. HSK Vol. 24. Berlin: W. de Gruyter.

Routledge Encyclopedia of Language Teaching and Learning ed. by Michael Byram. 2000. London: Routledge.

(The) Routledge Encyclopedia of Translation Studies ed. by Mona Baker. 1998. London: Routledge.

Semiotik/Semiotics. A Handbook on the Sign-Theoretic Foundations of Nature and Culture ed. by Roland Posner et al. 1997, 1998, 2003, 2004. HSK Vols. 13.1, 13.2, 13.3, 13.4. Berlin: W. de Gruyter.

Thesaurus of English Words and Phrases comp. by Peter Mark Roget. 1852 (and many later editions). London: Longman.

Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie ed. by Franz J. Hausmann et al. 1989, 1990, 1991. HSK Vols. 5.1, 5.2, 5.3. Berlin: W. de Gruyter.

Other literature

Alvar Ezquerra, Manuel. 1995. Los diccionarios del español en su historia. *International Journal of Lexicography* 8(3): 173-201. [Reprinted as 'Dictionaries of Spanish in their historical context' in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. II: 343-374.]

Arens, Hans. 1955/1969. *Sprachwissenschaft. Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*. Orbis Academicus. Freiburg: K. Alber.

Atkins, Sue et al. 2003. The Contribution of FrameNet to Practical Lexicography. *International Journal of Lexicography* 16(3): 333-357.

Béjoint, Henri. 1994/2000. *Tradition and Innovation in Modern English Dictionaries*. [2nd edn published as *Modern Lexicography. An Introduction.*] Oxford: Clarendon Press.

Boisson, Claude et al. 1991. Aux origines de la lexicographie: les premiers dictionnaires monolingues et bilingues. *International Journal of Lexicography* 4(4): 261-315.

Calzolari, Nicoletta. 1988. The Dictionary and the Thesaurus Can Be Combined. Evens, Martha Walton (Ed.). 1988. *Relational Models of the Lexicon. Representing Knowledge in Semantic Networks.* Studies in Natural Language Processing. Cambridge: Cambridge University Press.

Campoy Cubillo, María Carmen and M. Pilar Safont Jordà (Eds.). 2004. *Computer-mediated Lexicography in the Foreign Language Learning Context.* Collecció Estudis Filològics 18. Castelló de la Plana: Publicacions de la Universitat Jaume I.

Carter, Ronald. 2004. Report from the Chair. *BAAL News* 77: 2-4.

Collison, Robert L. 1982. *A History of Foreign-Language Dictionaries.* The Language Library. London: A. Deutsch/Oxford: Blackwell. [Chapter 11, 'The Contribution of Historical and Comparative Linguistics', reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. II: 52-64.]

Crystal, David. 1986. The Ideal Dictionary, Lexicographer and User. Ilson, Robert (Ed.). 1986. *Lexicography: An Emerging International Profession*: 72-81. The Fulbright Papers Vol. I. Manchester: Manchester University Press. [Reprinted in Hartmann R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 319-327.]

Dodd, W. Steven. 1989. Lexicomputing and the Dictionary of the Future. James, Gregory (Ed.). *Lexicographers and Their Works*: 83-93. Exeter Linguistic Studies 14. Exeter: University of Exeter Press. [Reprinted in Hartmann R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 351-362.]

Fontenelle, Thierry. 2000. Introduction: Dictionaries, Thesauri and Lexical-semantic Relations. *International Journal of Lexicography* 13(4): 229-231 [followed by 4 papers on the theme of the issue].

Fontenelle, Thierry. 2003. Introduction. *International Journal of Lexicography* 16(3): 231 [followed by 5 papers et al. on the theme of the issue].

Hanks, Patrick. 2003. Lexicography [Chapter 3]. Mitkov, Ruslan (Ed.). 2003. *The Oxford Handbook of Computational Linguistics*: 48-69.

Harras, Gisela. 1989. Wörterbücher als Hilfsmittel der linguistischen Forschung. Hausmann, F.J. et al. (Ed.). 1989. *Wörterbücher/Dictionaries/Dictionnaires*. Vol. I: 159-163.

Hartmann, Reinhard R.K. 1972. Über den Einfluß der linguistischen Semantik auf die englisch-amerikanische Lexikographie. *Linguistik und Didaktik* 3: 197-208.

Hartmann, Reinhard R.K. (Ed.). 1979. *Dictionaries and Their Users. Papers from the 1978 BAAL Seminar on Lexicography.* Exeter Linguistic Studies. Exeter: University of Exeter Press.

Hartmann, Reinhard R.K. 1991. Contrastive Linguistics and Bilingual Lexicography. Hausmann, F.J. et al. (Ed.). 1991. *Wörterbücher/Dictionaries/Dictionnaires*. Vol. III: 2854-2859.

Hartmann, Reinhard R.K. 1994. The Onomasiological Dictionary in English and German. A Contrastive Textological Perspective. Hüllen, W. 1994: 137-149.

Hartmann, Reinhard R.K. 1996. Lexicography as an Applied Linguistic Discipline. Hartmann, R.R.K (Ed.). *Solving Language Problems*: 230-244. Exeter Linguistic Studies. Exeter: University of Exeter Press.

Hartmann, Reinhard R.K. 2001. *Teaching and Researching Lexicography.* Applied Linguistics in Action Series. Harlow: Pearson Education. [Chapter 8, 'Methods in dictionary research', reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 429-444.]

Hartmann, Reinhard R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. London: Routledge.

Hartmann, Reinhard R.K. 2004. Lexicography and Translation. Chan, Sin-Wai (Ed.). *Translation and Bilingual Dictionaries*: 7-21. Lexicographica Series Maior 119. Tübingen: M. Niemeyer.

Hartmann, Reinhard R.K. Forthcoming. Onomasiological Dictionaries in 20th-century Europe. *Lexicographica. International Annual for Lexicography.*

Haugen, Einar. 1987. *Blessing of Babel. Bilingualism and Language Planning. Problems and Pleasures*. Berlin: W. de Gruyter.

Herbst, Thomas et al. (Eds.). 2004. *Lexikografie, ihre Basis- und Nachbarwissenschaften. (Englische) Wörterbücher zwischen 'common sense' und angewandter Theorie.* Lexicographica Series Maior 118. Tübingen: M. Niemeyer.

Huang, Jianhua. 1994. Chinese and Western Metalexicography. Flowerdew, L. and A.K.K. Tong (Eds.). *Entering Text*: 228-238. Hong Kong: HKUST Language Centre. [Reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 391-404.]

Hüllen, Werner (Ed.). 1994. *The World in a List of Words.* Lexicographica Series Maior 58. Tübingen: M. Niemeyer.

Hüllen, Werner. 1999. *English Dictionaries 800-1700. The Topical Tradition*. Oxford: Oxford University Press.

Hüllen, Werner. 2004. *A History of Roget's Thesaurus*. Oxford: Oxford University Press.

Janssen, Maarten. 2004. Multilingual Lexical Databases, Lexical Gaps, and SIMuLLDA. *International Journal of Lexicography* 17(2): 137-154.

Katz, Bill. 1998. *Cuneiform to Computer. A History of Reference Sources.* History of the Book Series 4. Lanham MD: The Scarecrow Press. [Chapter 2, 'Reference Books from Cuneiform to Computer', reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 405-421.]

Kretzenbacher, Heinz L. 2002. Wissenschaftsbezogene Wortschätze [Article 114]. Cruse, D.A. et al. (Ed.). 2002. *Lexikologie/Lexicology*: 919-925.

Kwong, Oi Yee et al. 2004. Alignment and Extraction of Bilingual Legal Terminology from Context Profiles. *Terminology* 10(2): 81-99.

Landau, Sidney. 1984/2001. *Dictionaries. The Art and Craft of Lexicography*. New York: C. Scribner/Cambridge: Cambridge University Press.

Lutzeier, Peter R. 2002. Der Status der Lexikologie als linguistische Disziplin [Article 1]. Cruse, D.A. et al. (Ed.). 2002. *Lexikologie/Lexicology*: 1-14.

Martin, Willy J.R. 2003. Lexicography, Lexicology, Linking and the Hub-and-spoke Model. Botha, Willem (Ed.). *'n Man wat beur. Huldigingsbundel vir Dirk van Schalkwyk*: 268-285. Stellenbosch: Buro van die WAT.

McArthur, Tom. 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablets to the Computer*. Cambridge: Cambridge University Press.

McArthur, Tom. 1998. What Then *Is* Reference Science?' [Chapter 17]. Tom McArthur. 1998. *Living Words. Language, Lexicography, and the Knowledge Revolution*: 215-222. Exeter: University of Exeter Press. [Reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 422-428.]

Meier, Hans H. 1969. Lexicography as Applied Linguistics. *English Studies* 50: 141-151. [Reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts*. Vol. III: 307-318.]

Miller, George A. et al. 1990. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography* 3(4): 235-244.

Ooi, Vincent B.Y. 1998. *Computer Corpus Lexicography.* Edinburgh Textbooks in Empirical Linguistics. Edinburgh: Edinburgh University Press.

Rey-Debove, Josette. 1971. *Étude linguistique et sémiotique des dictionnaires français contemporains.* Approaches to Semiotics 13. The Hague: Mouton.

Starnes, DeWitt T. and Gertrude E. Noyes. 1946/1991. *The English Dictionary from Cawdrey to Johnson 1604–1755*. Chapel Hill NC: University of North Carolina Press. [2nd edn by Gabriele Stein. Amsterdam: J. Benjamins. Chapter 17, 'Thomas Dyche and William Pardon's *A New General English Dictionary* (1735)', reprinted in Hartmann, R.R.K. (Ed.). 2003. *Lexicography. Critical Concepts.* Vol. II: 15-28.]

Svensén, Bo. 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.

Tono, Yukio. 2004. Research on the Use of Electronic Dictionaries for Language Learning: Methodological Considerations. Campoy Cubillo, M.C. and M.P. Safont Jordà (Eds.). 2004: 13-28.

Van Hoof, Henri. 1995. Translators and the Writing of Dictionaries. Delisle, Jean and Judith Woodsworth (Eds.). *Translators through History*: 229-242. Benjamins Translation Library 13. Amsterdam: J. Benjamins.

Van Sterkenburg, Piet G.J. 1992. Electronic Onomasiology: Van Dale Greater Dictionary of Synonyms. Tommola, Hannu et al. (Ed.). *EURALEX '92 Proceedings*: 519-526. Studia Translatologica A.2. Tampere: Tampereen Yliopisto.

Wang, Weiwei. 2001. *Zweisprachige Fachlexikographie. Benutzunsforschung, Typologie und mikrostrukturelle Konzeption.* Angewandte Sprachwissenschaft 8. Frankfurt: P. Lang.

Wiegand, Herbert E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Berlin: W. de Gruyter.

Wikberg, Kay. 1983. Methods in Contrastive Lexicology. *Applied Linguistics* 4(3): 213-221.

Wilks, Yorick A. et al. 1996. *Electric Words: Dictionaries, Computers and Meanings*. Cambridge MA: MIT Press.

Ostensiewe illustrasies as mikrostrukturele items in verklarende skoolwoordeboeke gerig op leerders in die junior sekondêre fase[*]

Phillip Louw, *Buro van die Woordeboek van die Afrikaanse Taal, Stellenbosch en Departement Afrikaans en Nederlands, Universiteit van Stellenbosch, Stellenbosch, Republiek van Suid-Afrika (pal@sun.ac.za).*

Opsomming: 'n Verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase moet, in die Suid-Afrikaanse konteks, gebruikersgerigte praktyke navolg om op kommunikatiewe en pedagogiese vlak waarde toe te voeg vir beide moedertaal- en niemoedertaalsprekende leerders. Een so 'n praktyk is die opname van ostensiewe illustrasies as mikrostrukturele items om die oordrag van pragmatiese en semantiese data te bevorder. In hierdie artikel word die waarde van ostensiewe illustrasies en hulle relevansie binne 'n verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase beoordeel, waarna sekere voorbehoude rakende hulle gebruik behandel word. 'n Besluit om wel ostensiewe illustrasies op te neem, noodsaak verder 'n bespreking van waar om hierdie items in die mikrostruktuur te plaas sowel as 'n bespreking van verskillende tipes illustrasies wat vir opname in die woordeboektipe onder bespreking geskik sal wees. Vervolgens word 'n kritiese evaluering van die hantering van ostensiewe illustrasies in 'n eg Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase gegee en voorstelle gemaak vir 'n verbeterde mikrostrukturele behandeling wat aan die eise van die gebruikersperspektief voldoen. Laastens word die kwessie van noodsaaklike mediostrukturele ondersteuning vir hierdie datatipe in die vorm van presiese, omvattende kruisverwysings, ter sprake gebring. Hierdie artikel probeer dus om kriteria vir die suksesvolle leksikografiese hantering van ostensiewe illustrasies aan te bied.

Sleutelwoorde: BYSKRIFTE, BYVOEGLIKE NAAMWOORD, GROEPILLUSTRASIE, JUNIOR SÊKONDERE ONDERRIGFASE, KOMMUNIKATIEWE WAARDE, KRUISVERWYSINGS, MIKROSTRUKTURELE ITEMS, OSTENSIEWE ILLUSTRASIES, PEDAGOGIESE WAARDE, SELFSTANDIGE NAAMWOORD, STRUKTUURMERKERS, VERKLARENDE SKOOLWOORDEBOEK, VOORBEELDMATERIAAL, WERKWOORD

Abstract: Ostensive Illustrations as Microstructural Items in Explanatory School Dictionaries Targeting Learners in the Junior Secondary Phase. An

---

[*] Hierdie artikel is 'n vertaling en verwerking van 'n hoofstuk uit 'n D.Litt.-proefskrif *Criteria for a Multifunctional, Monolingual Dictionary in Junior Secondary Education*, wat in April 2004 deur die Universiteit van Stellenbosch, Stellenbosch, Republiek van Suid-Afrika, aanvaar is.

explanatory school dictionary targeting learners in the junior secondary phase must, in the South African context, pursue user-oriented practices in order to add value for both mother-tongue- and non-mother-tongue-speaking learners on a communicative and a pedagogical level. One such practice would be the inclusion of ostensive illustrations as microstructural items in order to enhance the transfer of pragmatic and semantic data. In this article, the value of ostensive illustrations and their relevance in an explanatory school dictionary targeting learners in the junior secondary phase are evaluated, after which certain reservations regarding their use are addressed. A decision on whether to include ostensive illustrations demands a further discussion of where to position these items in the microstructure, as well as a discussion of different types of illustrations that will be suitable for inclusion in the type of dictionary profiled. Thereafter a critical evaluation of the treatment of ostensive illustrations in a genuinely South African explanatory school dictionary targeting learners in the junior secondary phase is provided and suggestions made for an improved microstructural treatment that meets the demands set by the user perspective. Lastly, the issue of necessary mediostructural support for this data type in the form of precise, comprehensive cross-references will be raised. This article aims to present criteria for the successful lexicographical treatment of ostensive illustrations.

Inleiding

Die junior sekondêre onderrigfase is sonder twyfel deurslaggewend in die taalverwerwingsproses, hetsy in die verbreding van die beheersing van die moedertaal, of, veral in die Suid-Afrikaanse konteks, in die vaslegging van die tweede taal (dikwels Engels) en uitbreiding van die taalvermoë daarin om die leerder in staat te stel om daardie tweede taal as onderrigmedium effektief te ontsluit. In hierdie proses speel die gebruik van innoverende, effektiewe taalhulpmiddels 'n rol van kardinale belang. 'n Verklarende woordeboek wat pasgemaak is om aan die behoeftes van leerders (beide moedertaal- en niemoedertaalsprekers) in die junior sekondêre fase te voldoen, is een so 'n hulpmiddel.

Buiten hierdie belangrike kommunikatiewe funksie sal so 'n woordeboek ook noodwendig 'n refleksiewe leksikografiese funksie moet vervul. Vir baie leerders sal hierdie fase die eerste wees waartydens die oordrag van woordeboekgebruiksvaardighede deel uitmaak van spesifieke programme. So 'n woordeboek moet dus bruikbaar wees in 'n klaskameropset om die leerders voor te berei om lewenslange gebruikers van standaard- verklarende woordeboeke te wees. Om aan hierdie kriterium te voldoen, moet innovasie opgeweeg word teen leksikografiese tradisie en 'n balans verkry word sonder om die gebruikersperspektief uit die oog te verloor.

Een leksikografiese praktyk waarvan die kommunikatiewe en pedagogiese waarde nie onderbeklemtoon kan word nie, is die benutting van ostensiewe illustrasies (sien die afdeling "Die waarde van ostensiewe illustrasies" hieron-

der). Hierdie praktyk neem ook die gebruikersperspektief in ag. 'n Steekproef wat onlangs deur die skrywer in drie multikulturele skole met 107 moedertaal- en niemoedertaalsprekende leerders van Engels in grade agt en nege onder- neem is, het sterk daarop gedui dat hierdie leerders verwag dat 'n verklarende woordeboek wat op hulle gerig is, ostensiewe illustrasies sal insluit om die verbale inligtingsoordrag aan te vul (74% van die respondente was ten gunste daarvan). Dit is dus vreemd dat slegs een Suid-Afrikaanse verklarende skool- woordeboek, naamlik die *Chambers–Macmillan South African Dictionary Junior Secondary* (voortaan SADJS), wat spesifiek op leerders in hierdie onderrigfase gerig is, wel ostensiewe illustrasies op 'n konsekwente basis benut, veral gegewe die waarde wat hierdie items kan toevoeg.

Die waarde van ostensiewe illustrasies

Ostensiewe illustrasies kan ingesluit word as mikrostrukturele items en kan 'n baie belangrike rol vervul deur "visual support for the description of the meaning content of linguistic units" (Svensén 1993: 167) te verskaf. Illustrasies is van besondere belang in 'n skoolwoordeboek, selfs in een wat leerders in die junior sekondêre fase as teikenmark het. Svensén (1993: 167-168) beskryf die opvoedkundige kriterium vir insluiting soos volg: "A picture reinforces learn- ing, and also a picture may often, quite differently from a definition, appeal to the reader's previous experience of the world, and enable him to achieve fresh insights." Omdat illustrasies oor die vermoë beskik om 'n vroeëre ervaring met 'n bepaalde voorwerp, konsep, e.s.m. beter as 'n item wat 'n betekenisparafrase- ring bied, te heraktiveer, voeg hy (1993: 168) by dat "illustrations are therefore more important in children's dictionaries than in dictionaries for adults".

Voorbehoude rakende die gebruik van ostensiewe illustrasies

Daar is sekere voorbehoude rakende die insluiting van illustrasies in 'n verkla- rende skoolwoordeboek wat op leerders in die junior sekondêre fase gerig is. Daar kan vernaamlik geredeneer word dat "they are rather space consuming" (Bogaards 1999: 122) en daarmee saam dus duur. Moontlike teenargumente is dat illustrasies egter die gebruikersvriendelikheid en opvoedkundige waarde van 'n skoolwoordeboek beduidend verbeter en dat, soos Svensén (1993: 167) aanvoer, hulle ruimte kan spaar "in many cases where a purely verbal descrip- tion would have to be very detailed".

Nog 'n argument wat dikwels teen die insluiting van illustrasies aange- voer word, is dat hulle neig om die lading van ensiklopediese data in die woor- deboek te verhoog. Gouws (1994: 69-71) het egter aangetoon dat hierdie kritiek hoofsaaklik ontstaan het weens die ondeurdagte gebruik van illustrasies in sommige verklarende woordeboeke, waar hulle nie dien om die verbale inlig- tingsoordrag aan te vul nie, maar eerder om dit te herhaal. Hy (1994: 69) wys

duidelik daarop dat 'n mate van ensiklopediese data nie noodwendig onaan-
vaarbaar is nie en dat die gebruik van illustrasies teen die agtergrond van die
"toelaatbare ensiklopediese belading" van die spesifieke woordeboek gesien
moet word. Daar kan geredeneer word dat illustrasies in 'n verklarende skool-
woordeboek gerig op leerders in die junior sekondêre fase ingesluit kan word,
nie net weens hulle opvoedkundige waarde nie, maar ook weens die doeltref-
fende wyse waarop hulle meer van die ensiklopediese voorwerp-konstitueren-
de kennis van 'n spesifieke leksikale item reflekteer. Sorg moet egter gedra
word om nie te veel van die data wat reeds in die betekenisparafrasering van 'n
item en in die voorbeeldmateriaal aangebied is, te herhaal nie.

Bogaards (1999: 122) noem nog 'n potensiële nadeel van illustrasies wat
leksikograwe noop om versigtig te werk te gaan, naamlik "it is difficult to find
pictures which are clear to learners with very different cultural backgrounds
and which at the same time conserve a flavour of authenticity". Die samestel-
lers van 'n Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in
die junior sekondêre fase het wel in hierdie opsig 'n voordeel, naamlik dat
hulle woordeboek Suid-Afrikaanse leerders teiken. Daar is moontlik 'n hoë
mate van kulturele diversiteit in hierdie teikengroep en meer navorsing moet
gedoen word om vas te stel of hierdie kulturele groeperings die tipes voorwer-
pe, handelings en beskrywende items wat ostensief geïllustreer word, kan kon-
septualiseer. Daar is egter 'n sterk saak uit te maak dat daar 'n gedeelde Suid-
Afrikaanse ervaring en gevolglik 'n gedeelde Suid-Afrikaanse kulturele identi-
teit bestaan, wat deur die samestellers van 'n Suid-Afrikaanse verklarende
skoolwoordeboek gereflekteer moet word. SADJS blyk 'n gebalanseerde bena-
dering te volg deur generiese illustrasies en illustrasies wat in besonder rele-
vant vir Suid-Afrikaanse leerders is, af te wissel. Die insluiting van illustrasies
van baie diere wat uitsluitlik of veral in Suid-Afrika voorkom en die insluiting
van 'n illustrasie van 'n AK 47-aanvalsgeweer in die groepillustrasie by **gun**
bied voorbeelde van die samestellers se verbintenis daartoe om die illustrasies
relevant vir Suid-Afrikaanse leerders te maak.

Die plasing van ostensiewe illustrasies in verklarende skoolwoorde-boeke gerig op leerders in die junior sekondêre fase

Indien besluit is om wel illustrasies te verskaf, moet ernstige oorweging ge-
skenk word aan waar hierdie illustrasies geplaas moet word. Die plasing van
illustrasies word hoofsaaklik deur hulle aard en die tipologiese profiel van die
woordeboek bepaal. Groot prente met 'n komplekse stel byskrifte word goed
benut in die agterwerk van die *Reader's Digest Afrikaans–Engelse Woordeboek/
English–Afrikaans Dictionary*, 'n tweerigting- vertalende gesinswoordeboek, en
blyk veral nuttig te wees om die struikelblok van insluiting in beide kante van
'n tweetalige woordeboek te oorkom. In 'n eentalig verklarende skoolwoorde-
boek, waarin groot prente onprakties en die nadruk op die onmiddellikheid
van inligtingsoordrag is, kan geredeneer word dat kleiner illustrasies wat inge-

sluit word as mikrostrukturele items met óf een, óf 'n paar bepaalde lemmas as adres(se), sowel bekostigbaar as doeltreffend is.

In die setproses moet egter noukeurig daarop gelet word dat hierdie illustrasies naby die lemma (of betekenisonderskeiding) wat die adres daarvan is, geplaas word, en, indien hulle meer as een lemma as adres het, moet die samestellers verseker dat toereikende artikel-eksterne kruisverwysings snel-toegang na beide die teikeninskrywing en, van daar af, na die relevante illustrasie bewerkstellig.

'n Klassifikasie van tipes ostensiewe illustrasies

Die besluit om kleiner illustrasies in die mikrostruktuur op te neem, verg versigtige beplanning rakende watter tipes ostensiewe illustrasies om in te sluit. Dit val nie binne die reikwydte van hierdie artikel om 'n omvattende tipologie van ostensiewe illustrasies te bied nie. Aspekte van sodanige tipologieë (soos in o.a. Svensén (1993: 171-178) en Ilson (1987: 193-199) vervat) sal egter benut word in die volgende kort bespreking oor gekose en geskikte illustrasietipes in 'n Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase.

'n Moontlike beginpunt in die bepaling van watter tipes illustrasies om te benut, sou wees om te besluit watter woordsoorte ostensiewe beligting sou benodig. Hierdie potensiële klassifikasie is bloot van praktiese waarde vir die leksikograaf, aangesien dit 'n maklik klassifiseerbare vertrekpunt in die woordeboekontwerpproses bied. Die effektiwiteit van ander klassifikasiestelsels kan in die aanloop tot die woordeboekontwerpfase ondersoek word.

Aangesien selfstandige naamwoorde die primêre woordsoort is wat vir ostensiewe beligting gemerk is, moet 'n meer gedetailleerde bespreking van selfstandige naamwoorde volg. Tog kan ostensiewe illustrasies ook benut word by ander woordsoorte, veral by byvoeglike naamwoorde en werkwoorde.

Byvoeglike naamwoorde wat 'n stel antonieme of 'n komplementêre paar uitmaak (lank x kort, dik x dun, e.s.m.) kan goed deur middel van 'n prent geïllustreer word, mits die illustrasie maklik dekodeerbaar is en daar toereikende kruisverwysings is na die korrekte lemmas of betekenisonderskeidings waarop die illustrasie van toepassing is. Die sukses van die illustrasie van byvoeglike naamwoorde hang egter (ten spyte van die leksikograaf se beste pogings om duidelikheid te verseker) in 'n sekere mate af van die gebruiker se vermoë om die abstrakte verhouding tussen die prent en die leksikale item(s) wat dit probeer illustreer, te interpreteer, bv. dat die illustrasie van twee seuns van wisselende lengtes wat langs mekaar staan, die verskil tussen *lank* en *kort* kan verklaar. Ilson (1987: 198) maak dus 'n geldige punt: "a verbal explanation is a necessary complement to the picture."

Daar is ook "a risk of misunderstanding" (Svensén 1993: 178) in gevalle waar werkwoorde van ostensiewe illustrasies voorsien word. Waar die presiese verwysing van die illustrasie nie voldoende duidelik gemaak word nie, kan die

gebruiker maklik die fout begaan om aan te neem dat een van die selfstandige naamwoorde die fokuspunt in die illustrasie is, eerder as die handeling wat uitgebeeld word.

Selfstandige naamwoorde sal die fokuspunte van die meeste illustrasies uitmaak en kan op verskeie maniere weergegee word. In 'n relatief klein skoolwoordeboek kan 'n mens verwag dat die meeste ostensiewe illustrasies in die mikrostruktuur net "one object of the kind concerned, without comment" sal vertoon. 'n Mens moet dan aanneem dat "the one shown is representative of its class" (Svensén 1993: 171). Ten spyte van die voor die hand liggende beperkings op ruimte, is daar tog ook drie ander illustrasietipes wat effektief kan optree as mikrostrukturele items in 'n kompakte Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase. Svensén (1993: 172-174) identifiseer hierdie tipes as illustrasies wat "several objects of the same class", dié wat "an object in its surroundings", en dié wat "objects in operation" uitbeeld. Al vier hierdie illustrasietipes is teenwoordig in SADJS.

Die gebruik van ostensiewe illustrasies in SADJS

Terwyl 'n mededingende woordeboek soos *The South African Oxford School Dictionary* (wat dieselfde gebruikersgroep teiken) op 'n minder visuele benadering besluit het en die ruimte daarin vir 'n meer omvattende makrostruktuur en meer verbale data in die mikrostruktuur benut, sluit SADJS wel ostensiewe illustrasies as mikrostrukturele items in. Die verbale inligtingsoordrag word by sommige byvoeglike naamwoorde en werkwoorde deur ostensiewe illustrasies aangevul, maar sulke illustrasies word gewoonlik vir selfstandige naamwoorde gereserveer.

Illustrasies word selde in SADJS by byvoeglike naamwoorde geplaas. Een so 'n buitengewone voorkoms is 'n illustrasie wat tussen die lemmas **alternate** en **alternative** geplaas is om die verskil tussen hierdie twee byvoeglike naamwoorde duidelik te maak. Hierdie voorstelling van 'n "contrast set" (Ilson 1987: 194) deur middel van 'n illustrasie hou tred met die organisasieteorie van SADJS soos uitgestippel in die "Introduction", waar gestel word dat "numerous illustrations throughout the text help the reader … to distinguish between things that could easily be confused" (SADJS, p. v).

Die benadering om 'n illustrasie wat uit twee of meer kleiner prente bestaan, te benut om 'n onderskeid tussen verwarbare woorde te tref, word ook by sommige werkwoorde gebruik. Daar is byvoorbeeld illustrasies by **bring** (om 'n onderskeid tussen *bring*, *take* en *fetch* te tref), **lay** (om te onderskei tussen *lay* en *lie*) en **mow** (om te onderskei tussen *mow*, *harvest/reap* en *trim*). SADJS slaag hoofsaaklik in die doel van ondubbelsinnigmaking as gevolg van twee belangrike besluite wat deur die samesteller(s) geneem is.

Eerstens, illustrasies tree in wisselwerking met die ingevoegde binnetekste (gebruiksnotas) in getinte tekskassies op wat die "verbale" komplementering verskaf wat Ilson (1987: 198) as 'n voorvereiste in sulke gevalle stel. Hierdie in-

gevoegde verbale binnetekste is op hulle eie alreeds 'n kragtige medium om tussen verwarbare leksikale items te onderskei, maar indien hulle saam met die illustrasies gebruik word, word die gebruikersvriendelikheid van die woordeboek verbeter, veral vir minder bedrewe gebruikers.

Tweedens maak SADJS van verwyderde artikel-eksterne kruisverwysings gebruik om te verseker dat dié gebruiker wat byvoorbeeld **lie** opsoek, vinnig en doeltreffend na die illustrasie en meegaande gebruiksnota by **lay** gestuur word. Hierdie kruisverwysings word ook in getinte tekskassies aangebied, en hoewel hulle nie heeltemal doeltreffend is nie (hulle waarde word onder meer verminder deurdat dieselfde tipe kruisverwysing ook gebruik word om die gebruiker te rig op 'n paar verbale ingevoegde tekste wat 'n ondubbelsinnigmakende funksie het, en hulle dus nie unieke verwysing bied nie), verskaf hulle waardevolle hulp aan die teikengebruiker (sien die afdeling oor "Illustrasies en kruisverwysings" hieronder vir 'n verdere bespreking van hierdie stelsel se tekortkominge).

Dieselfde benadering word gevolg wanneer 'n onderskeid tussen twee of meer selfstandige naamwoorde getref moet word (sien die illustrasies by **scissors**, **plait** en **nut**). Tog is daar gevalle waarin illustrasies bloot "the reader (help) to identify objects" (SADJS, p. v). In hierdie verband bied Svensén (1993: 171-174) se vier illustrasietipes 'n nuttige raamwerk om SADJS se illustrasies te evalueer.

Sommige enkelvoorwerpe word in SADJS geïllustreer, maar daar word nooit verduidelik wat die kriteria vir insluiting is nie. Talle bokspesies, asook ander diere soos 'n gorilla en tarentaal word by die relevante lemmas uitgebeeld, maar dit is nie die geval met ander diere soos 'n sjimpansee, kameelperd, olifant, leeu en pou nie. Hierdie inkonsekwentheid is ook teenwoordig in die uitbeelding van ander tipes voorwerpe en getuig van 'n gebrek aan versigtige beplanning.

'n Verdere moontlike punt van kritiek rakende SADJS se illustrasies van enkelvoorwerpe is dat die illustrasies in sommige gevalle te gedetailleerd is, met die gevolg dat dit nie altyd vir alle gebruikers duidelik is wat eintlik uitgebeeld word nie. In die illustrasie van **dashboard** word 'n stuurwiel ook aangetoon en by **dinghy** (sien Illustrasie 1) is 'n roeispaan ook aanwesig.

Illustrasie 1

'din·ghy /'dɪŋɡɪ/ *noun* (**dinghies**) a small sailing boat or rowing boat: *The boy rowed his rubber dinghy across the dam.*

dinghy
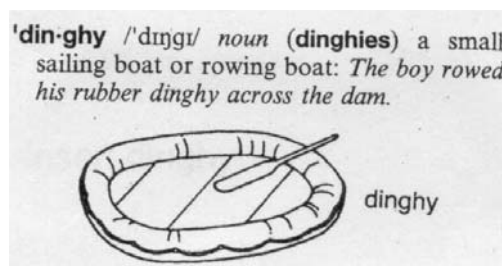
Gedetailleerde byskrifte of 'n verbale komplement moet benut word om te verseker dat daar geen risiko van 'n misverstand is nie.

Bykomend tot illustrasies van enkelvoorwerpe sluit SADJS ook groepillustrasies van verskeie hiponieme in, gewoonlik by 'n geskikte superordinaat. By **instrument** (sien Illustrasie 2), byvoorbeeld, bied 'n volbladgroepillustrasie van verskeie musiekinstrumente met 'n byskrif by elk van die komponente, 'n goeie oorsig oor die "klas" waartoe hierdie voorwerpe behoort. Hier vertoon die illustrasie kenmerke van die sinopsisartikel, aangesien dit vinnige toegang tot hiponieme bied deur hulle by die superordinaat uit te beeld.

Illustrasie 2
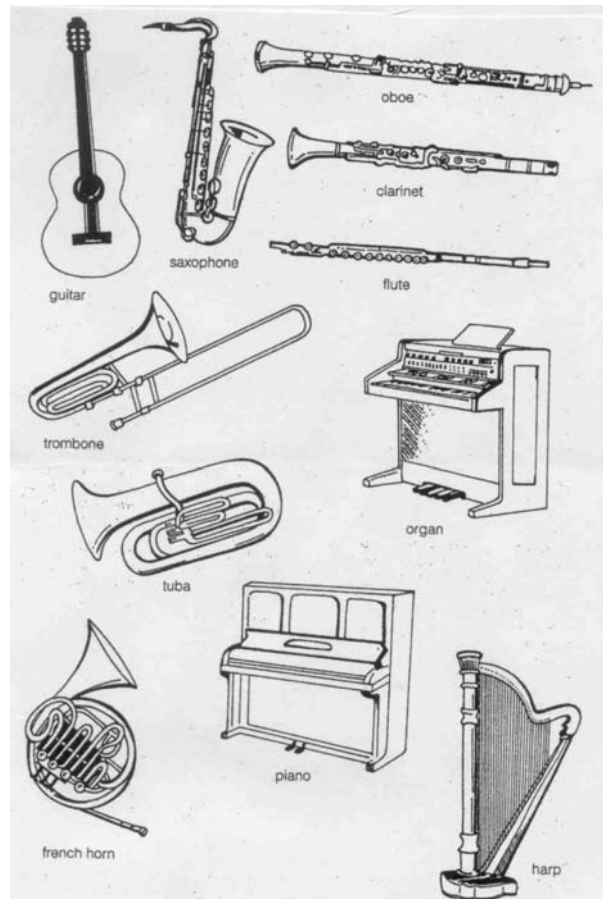
Daar is egter 'n tekortkoming in SADJS se behandeling van hierdie illustrasies. Anders as wat die geval was met die benadering tot illustrasies wat daarop gemik is om tussen verwarbare leksikale items te onderskei, skiet die deurslaggewende mediostrukturele ondersteuning by **instrument** tekort. Daar is geen

kruisverwysings met die groepillustrasie as teiken by enige van die lemmas wat die hiponieme (bv. **harp**) wat in hierdie groepillustrasie uitgebeeld word, verteenwoordig nie. Dit verlaag die waarde van SADJS se gebruik van illustrasies as 'n effektiewe hulpmiddel by die leksikografiese behandeling van hiponimie.

SADJS sluit ook illustrasies van "an object … shown in its normal environment" (Svensén 1993: 173) in. By **iris** (sien Illustrasie 3) word byvoorbeeld 'n illustrasie aangebied wat die iris van ander dele van die oog (bv. die kornea en die pupil) en van omliggende kenmerke soos die wenkbrou onderskei.

Illustrasie 3
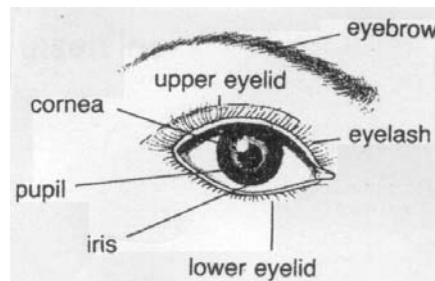
Hierdie illustrasie is besonder suksesvol omdat die byskrifte en hulle verwysingspunte duidelik gegee word en omdat daar gedetailleerde kruisverwysings is by elk van die lemmas wat met die beskrywings ooreenstem. By **horn** (sien Illustrasie 4) is die prosedure egter minder deeglik. Geen byskrifte word aangebied om die horings van die nek en kop (en samestellende dele daarvan) van die bok wat hier geïllustreer is, te onderskei nie.

Illustrasie 4

Dit strook duidelik nie met Svensén (1993: 173) se hoofkriterium vir die behandeling van sulke gevalle nie, d.w.s. dat die "object … must be plainly distinguishable".

Laastens bied SADJS 'n paar illustrasies van voorwerpe in werking. Hulle word sorgvuldig gekies en gewoonlik goed saamgestel (bv. die illustrasie van 'n vulkaniese uitbarsting by **volcano**). In sommige gevalle is die illustrasies egter te klein en te gedetailleerd (bv. by **rapid** (sien Illustrasie 5)), wat, ten spyte van 'n goeie stelsel van byskrifte, tot verwarring aanleiding kan gee.

Illustrasie 5
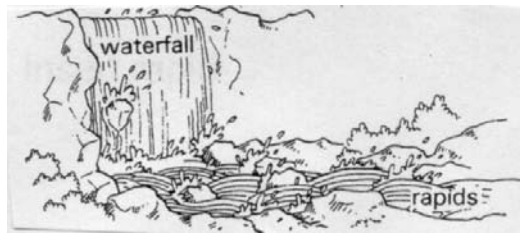
Illustrasies en kruisverwysings

Soos reeds vermeld, bied die plasing van illustrasies in die agterwerk oplossings vir heelwat van die probleme van die samesteller(s) van 'n tweerigting-vertalende woordeboek, maar in die geval van 'n verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase is, soos in die afdelings hierbo voorgestel, insluiting van illustrasies by of naby aan die relevante lemmas 'n meer koste-effektiewe opsie wat inligting onmiddellik beskikbaar stel. Die sukses van so 'n stelsel hang egter grootliks van die mediostruktuur af en in dié verband is daar bykomende lesse te leer uit die suksesse en foute van SADJS. Eerstens moet die byskrifte van illustrasies waarvan die teikenlemma polisemies is, met groot sorg hanteer word. By sommige gevalle in SADJS (sien bv. **bed**) word die probleem van verwyderde adressering opgelos deur die illustrasie net na die relevante betekenisonderskeiding in te sluit. Dit maak die gebruik van kruisverwysings onnodig en behoort, waar moontlik, ook in 'n Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase benut te word. SADJS pas egter nie hierdie benadering konsekwent toe nie (sien bv. **earth** (Illustrasie 6)).

Illustrasie 6
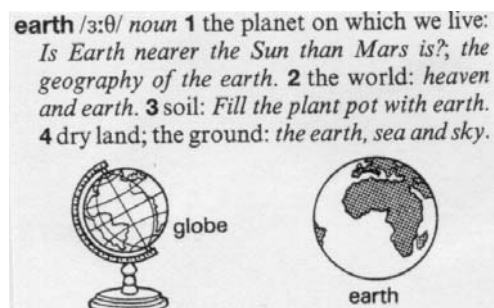
'n Benadering wat onmiddellike adressering bied, is nie in alle gevalle toepaslik nie. Die illustrasie kan te groot wees (soos in die volgende paragraaf uitgewys sal word) en kan derhalwe die kohesie van die mikrostruktuur onderbreek en word daarom aan die einde (en soms in die nabyheid) van die relevante artikel aangebied. In gevalle waar die lemma polisemies is, is dit belangrik dat die byskrifte as presiese kruisverwysings optree wat die gebruiker lei na die presiese betekenisonderskeiding wat in die illustrasie uitgebeeld word. Sulke presiese kruisverwysings word nie in SADJS toegepas nie, wat veroorsaak dat komplekse, verwyderde adresseringsprosedures gevolg moet word, sonder dat voldoende items wat sekondêre adresse aandui, as gidse verskaf word.

Om ruimte te bespaar en onomasiologiese verhoudings tussen lemmas te bewerkstellig, word groot groepillustrasies dikwels by 'n lemma aangebied wat 'n superordinaat verteenwoordig. Soos duidelik behoort te blyk uit die vorige paragraaf, behoort die byskrifte baie presiese kruisverwysings te wees wat die gebruiker lei na die relevante betekenisonderskeiding van enige van die lemmas wat in die illustrasie uitgebeeld word. Hierdie kruisverwysings moet egter deur 'n ewe presiese en duidelike verwysingsinskrywing in die mikrostrukuur van die relevante artikel aangevul word.

SADJS bied wel kruisverwysings om met die byskrifte ooreen te stem. Hulle word in getinte tekskassies teen die einde van elke relevante betekenisonderskeiding of artikel aangebied en kom in die vorm van 'n vet gedrukte kruisverwysing voorafgegaan deur die mediostrukturele merker "See" voor. Daar kan egter gevra word of hierdie kruisverwysings hulle rol bevredigend vertolk. In hulle huidige formaat word hulle tussen die ingevoegde binnetekste wat pragmatiese en grammatikale data bied, genes. In 'n lang artikel soos **each**, is die kruisverwysing een van vier getinte tekskassies en daar kan geargumenteer word dat die tekskassie as struktuurmerker nie prominent genoeg is om vinnige toegang tot die kruisverwysing te verseker nie. In hierdie verband is die unieke merker wat deur CIDE benut word, miskien 'n beter opsie. By die lemma **hip** fruit word die volgende item wat die kruisverwysing gee, aangebied: PIC Berries. Tipografiese en nietipografiese struktuurmerkers kombineer om 'n suksesvolle kruisverwysing te bewerkstellig na die collage wat naby **berry** geplaas is en waarin die prentjie van *rose hip* voorkom. Die pyl (nietipografies) en die hoofletters PIC as 'n afkorting van "picture" kombineer met die aanbieding van **Berries** in vet druk en 'n ander druktipe om 'n prominente, unieke en maklik identifiseerbare verwysingsinskrywing te skep. Hierdie uitstekende metode, of soortgelyke markering, kan met vrug benut word in 'n Suid-Afrikaanse verklarende skoolwoordeboek gerig op leerders in die junior sekondêre fase.

Samevatting

Illustrasies sal 'n belangrike rol in die ondersteuning van die oordrag van pragmatiese en selfs semantiese data in 'n Suid-Afrikaanse verklarende skoolwoor-

deboek gerig op leerders in die junior sekondêre fase speel. Groter sorg moet egter bestee word as wat die geval was in SADJS. Alle tipes illustrasies in so 'n skoolwoordeboek moet voldoen aan die beginsels van duidelikheid, omvattende behandeling en presiese kruisverwysing, soos in die voorafgaande evaluering van SADJS uitgestippel is.

Bibliografie

Woordeboeke

Chambers–Macmillan. 1996. *Chambers–Macmillan South African Dictionary Junior Secondary*. Manzini: Macmillan Boleswa Uitgewers.

Grobbelaar, P. (Red.). 1987. *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Dictionary*. Kaapstad: The Reader's Digest Association.

Hawkins, J.M. 1996. *The South African Oxford School Dictionary*. Kaapstad: Oxford University Press.

Procter, P., et al (Eds.). 1995. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press.

Ander bronne

Bogaards, P. 1999. Access Structures of Learners' Dictionaries. Herbst, T. en K. Popp (Reds.). 1999. *The Perfect Learner's Dictionary (?)*: 113-130. Tübingen: Max Niemeyer Verlag.

Gouws, R.H. 1994. Ostensiewe adressering in vertalende woordeboeke. *Lexikos* 4: 61-85.

Ilson, R. 1987. Illustrations in Dictionaries. Cowie, A. (Red.). 1987. *The Dictionary and the Language Learner*: 193-212. Tübingen: Max Niemeyer Verlag.

Svensén, Bo. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.

On Issues of Labelling in the
Dictionnaire Français–Mpongwé:
A Case-study

P.A. Mavoungou, *Department of Linguistics, Omar Bongo University,*
Libreville, Gabon (moudika2@yahoo.fr)

Abstract: Dictionaries contain lexicographic data whose occurrence is restricted to certain geographical areas, subject fields, professions, etc. It is part of the duties of the lexicographer to give an account of such deviations to ensure a successful retrieval of the information on the part of the user. This contribution presents a discussion on labelling issues in the *Dictionnaire Français–Mpongwé*. Although the main focus is on the presentation of different types of labelling as well as problems in labelling, textual condensation procedures and mediostructural representations (together with some aspects of the user perspective) are also critically evaluated. It is shown that these procedures reveal some inconsistencies which are not accounted for in the outer texts (front matter and back matter texts) of the dictionary. Finally suggestions are made for the improvement of the access structure of this dictionary.

Keywords: LABELS, TEXTUAL CONDENSATION, MEDIOSTRUCTURAL REPRESENTATION, ACCESS STRUCTURE, BORROWING, OUTER TEXTS, FRONT MATTER, BACK MATTER, USER PROFILE, METALEXICOGRAPHY

Résumé: **À propos de quelques questions d'étiquetage dans le** *Dictionnaire Français–Mpongwé*: **un cas à l'étude.** Les dictionnaires contiennent des données lexicographiques dont l'occurrence est restreinte à une région géographique donnée, un champ de connaissance, une profession, etc. Il est du devoir du lexicographe de prendre en compte ces déviations afin d'assurer un transfert d'information positif au niveau de l'utilisateur. La présente contribution présent une discussion sur les problèmes d'étiquetage dans le *Dictionnaire Français–Mpongwé*. Bien que le principal centre d'intérêt se situe au niveau de la présentation des différentes catégories d'étiquettes ainsi que les problèmes d'étiquetage, les procédés de condensation textuelles et les représentations mediostructurelles (accompagnées de quelques aspects relatifs aux utilisateurs du dictionnaire) sont également critiquement évalués. Il est démontré que ces procédés révèlent quelques inconsistances qui ne sont pas rendues au niveau des textes externes (prétextes et posttextes) du dictionnaire. Finalement, des suggestions sont faites dans le sens de l'amélioration de la structure d'accès de ce dictionnaire.

Mots-clés: ETIQUETTES, CONDENSATION TEXTUELLE, REPRÉSENTATION MEDIOSTRUCTURELLE, STRUCTURE D'ACCÈS, EMPRUNT, TEXTES EXTERNES, PRÉ-TEXTES, POSTTEXTES, PROFILE DE L'USAGER, METALEXICOGRAPHIE

1.   Introduction

Mpongwe[1] is one of the six dialects of Omyene, the home language of five per

cent of the population of Gabon. It is spoken in Libreville as well as on the island of Pointe-Dénis.

The *Dictionnaire Français–Mpongwé* by Raponda-Walker is a monodirectional publication with French as source language and Mpongwe as target language. It was published for the first time during 1930–1934 by the Imprimerie de la Libre Lorraine in Metz. In 1961, 27 years later, the dictionary was published in the reverse direction French**–**Mpongwe by the Imprimie Saint Paul in Brazzaville. This version of the dictionary has since then been reprinted in 1995 under the auspices of the Raponda-Walker Foundation. The last version comprises some 8 000 articles arranged alphabetically according to the word tradition. The first part is the central word list, the dictionary proper, while the back matter or the second part of the dictionary contains an exposition of Mpongwe grammar (cf. Nyangone Assam and Mavoungou 2000).

The edition of 1961 was released when the compiler Raponda-Walker was already 90 years old. J. Adam, in the preface to this edition, is correct when stating that one needs much courage, good health and a good memory to carry out such a task.

The compilation of any dictionary has to take the needs and reference skills of the intended target user groups into account (cf. Hartmann 1989: 103). Although the introductory notes of the *Dictionnaire Français–Mpongwé* does not contain explicit reference to its target users, it is aimed at the average member of the Mpongwe speech community. It can be assumed that this target user group will not have a sound knowledge of dictionary use and therefore will need more assistance to cope successfully with the data presented in the dictionary. Moreover, the extent to which these data should be treated has to be determined by the needs and reference skills of the target users of the dictionary. From these assumptions one of the lexicographer's biggest tasks will be to make the product user-friendly.

2. Labelling, inconsistent labelling and the absence of outer texts in the frame structure

Dictionaries should be a reflection of the lexicon of a given language. Moreover, dictionaries contain lexicographic data whose occurrence is restricted to certain geographical areas, subject fields, professions, etc. For example, in any bilingual or multilingual dictionary with English as the treated language, the focus should be on one variety of English, e.g. British or American. But when including lexical items restricted to a variety exclusively used by the target users of the dictionary, the lexicographer should indicate this restricted usage by employing a system of geographical labels. With regard to lexicographic studies that have so far dealt with labelling issues, one can roughly classify lexicographers' viewpoints into three approaches. Some believe that lexicographic labels may be divided into four main categories, namely stylistic, geo-

graphical, temporal, and sphere of usage labels. Labels in the first category are used to indicate style levels, e.g. *formal*, *popular* or *slang*. The second category deals with the so-called diatopic or regional labels such as *American English*, *British English* or *South African English*. Labels in the third category have a diachronic function. They are temporal labels, i.e. *archaic*, *historical* or *obsolete*. Sphere of usage labels are labels indicating use in a special field of activity. Labels such as *lexicography*, *physics* and *law* fall in this category.

Further types include labels indicating attitude or connotation (e.g. *derogatory*, *offensive* or *obscene*), labels indicating frequency (e.g. *seldomly used*) and labels of borrowing (e.g. *loanword*) (Bureau of the WAT 1999: 10-13). Contrary to this viewpoint, Landau (1991: 217-218) distinguishes eight groups of lexicographic labels, namely: labels indicating currency or temporality (*old-fashioned, dated, archaic, obsolete, old use*), labels indicating regional or geographic variations (*U.S., British, Canadian, Australian, New Zealand, South African,* etc.), labels indicating technical or specialized terminology (*astronomy, chemistry, physics, sports,* etc.), labels indicating restricted or taboo sexual and scatological usage (*offensive, taboo, vulgar, obscene, rude,* etc.), labels for insulting terms (*offensive, insult, disparaging, derogatory, disapproving, contemptuous, sexist, racist*), labels for slang language (*slang*), labels indicating style (level), functional variety, or register (*formal, written, informal, spoken, colloquial,* (now rarely used) *literary, historical, poetic, humorous, facetious, jocular, approving, euphemistic, baby talk* or *child's word*), and labels indicating status or cultural level (*nonstandard* or *not standard*).

Another group of lexicographers (cf. Benson, Benson and Ilson 1986: 215-216) indicates that labels under the heading 'stylistic labels' include among others *derogatory, pejorative* and *offensive*. This constitutes quite a different opinion compared to the WAT's classification in which labels such as *derogatory, pejorative* and *offensive* appear under the heading 'labels indicating attitude or connotation'.

Norri (1996: 1-29) addresses the issue of labelling in some British and American dictionaries. More recently, the same topic has again been discussed by Norri (2000: 71-106). In this new approach, she groups the 145 words of her corpus into seven headings, namely: (1) Labelling for words for nationalities, (2) Labelling for words denoting racial and cultural groups, (3) Labelling for words for people considered to be lacking in intelligence, (4) Labelling for words for deceitful people, (5) Labelling for words for sexual orientation, (6) Labelling for some derogatory words for women, (7) Labelling for some derogatory words for men.

Although this article aims at addressing labelling issues in the *Dictionnaire Français–Mpongwé*, textual condensation procedures and mediostructural representations (with the user-perspective as background) will necessarily come to the fore. The emphasis will be on labels of borrowing, part of speech labels, stylistic labels, labels for special fields of activity and labels conveying information on pragmatics, syntax and semantics.

2.1   Labels of borrowing

Tables 1 and 2 contain examples of labels of borrowing used in the *Dictionnaire Français–Mpongwé*.

entries	labels	source language form	source language
ekwèrè	(fang)	unknown	Fa<u>n</u>
kado	(néol.)	cadeau	French
kalavati	(néol.)	cravate	French
keretyè	(néol.)	chrétien	French
kopi	(néol.)	copie	French
kurusa	(néol.)	cruz	Portuguese
krusu	(néol.)	cruz	Portuguese
lakóli	(néol.)	colle	French
lakrwa	(néol.)	croix	French
lamèsi	(néol.)	messe	French
laso	(néol.)	chaux	French
mangazè	n.l.	magasin	French
marabu	n.l.	marabout	French

entries	labels	source language form	source language
mero	(néol.)	numéro	French
molata	n.l.	mulato	Portuguese
pese	(néol.)	péché	French
póketi	(néol.)	poche	French
pósi	(néol.)	poche	French
semizi	(néol.)	chemise	French
sigarèti	(néol.)	cigarette	French
sigaro	(néol.)	cigare	French
sokola	(néol.)	chocolat	French
su	(néol.)	choux	French
tizani	(néol.)	tisane	French
tsatsópu	(néol.)	unknown	unknown

Convention: n.l. *is short for* not labelled.

Table 1: Examples of labels of borrowing used in the *Dictionnaire Français–Mpongwé*.

entries	labels	source language form	source language
bèlèti	(ang.)	bullet	English
bóyi	–	boy	–
fatere	–	factory	–
fiva	–	fever	–
kapina	–	carpenter	–
keki	–	cake	–
korèyi	–	corail	–
kóltar	–	coltar	–
meli	–	mail-steamer	–

entries	labels	source language form	source language
paléti	–	pilot	–
patóni	–	patron	–
pikóki	–	peacock	–
pusi	–	pussy	–
silipasi	–	slipper	–
siti	–	sheet	–
sitowa	–	store	–
sópu	–	shop	–
tikèti	–	kettle	–

Convention: *a dash indicates identity with the form given in the preceding square.*

Table 2: Examples of labels of borrowing used in the *Dictionnaire Français–Mpongwé*.

According to the system applied in Tables 1 and 2, it seems that the compiler wants to give an account of the item indicating the source language of the entry treated only for English. Many questions are likely to arise: Are the entries labelled (néol. = néologisme/neologism) not accurate enough as compared to the one labelled (ang. = anglais/English)? Are these entries too recent in the system of Mp<u>o</u>ngw<u>e</u> to be labelled (*from French*)? A close look at Raponda-Walker (1995: 138) leads us to the answer to these questions. In Raponda-Walker's view, the term *néologisme* is a comprehensive name for a variety of linguistic phenomena, viz. corrupted European words, borrowed words and indigenous words having acquired a new meaning.

From this definition, lexical items in Table 1 mainly fit into the categories of corrupted forms and loanwords. But this terminology is ambiguous and may confuse the user. As a matter of fact, Raponda-Walker first used the term *néologisme* in 1933. When the dictionary was published in 1961 the same terminology was again used. It is therefore also not surprising that nothing has been changed in the edition of 1995 under the auspices of the Raponda-Walker Foundation, for this is simply a reprint and not a revision. It is, however, confusing for the user because these lexical items which were labelled *néologismes* in 1961 may not be factually correct today. Objections can also be made with regard to the absence of labels at entries such as *molata*, *mangazè* and *marabu* as well as the lack of explanations in the user's guidelines as far as the system the lexicographer applied is concerned. The absence of a label is just as important as its presence. In other words, the absence of labels at these entries incorrectly indicates that they are indigenous to the Mpongwe dialect (cf. Bureau of the WAT 1999: 10 and 14).

2.2   Part of speech labels

In Table 3, the labels in capitals indicate parts of speech in the *Dictionnaire Français–Mpongwé*:

part of speech labels	significations	part of speech labels	significations
N.	Nom	V. PR.	Verbe pronominal
N. & ADJ	Nom et Adjectif	V. IMP.	Verbe impersonnel
ADJ.	Adjectif	V. D.	Verbe dérivé
ADJ. & N.	Adjectif et Nom	PR. IND.	Pronom indéfini
ADJ. NUM.	Adjectif numéral	PR. DEM.	Pronom démonstratif
ART.	Article	PRO. Démons.	Pronom démonstratif
A. N. C.	Adjectif numéral cardinal	PRÉP.	Préposition
ADJ. N. C.	Adjectif numéral cardinal	PR. PERS.	Pronom personnel
A. N. O.	Adjectif numéral ordinal	PR. PSS.	pronom possessif
ADJ. POSS.	Adjectif possesif	CONJ.	Conjonction
V.	Verbe	ADV.	Adverbe
V. A.	Verbe actif	ADV. LOC.	Adverbe locutionel
V. A. & N.	Verbe actif et Nom	L. ADV.	Locution adverbiale
V. N.	Verbe nominal	LOC. ADV.	Locution adverbiale
V. N. & A.	Verbe nominal et adjectif	INTERJ.	Interjection
V. N. & N.	Verbe nominal et Nom		

Table 3: Part of speech labels used in the *Dictionnaire Français–Mpongwé.*

Apart from some minor inconsistencies (the part of speech labels *Adjectif numéral cardinal*, *Pronom démonstratif* and *Locution adverbiale* have two variants each), the main objection with regard to these labels is the absence of a metatext in the introductory section listing and explaining the part of speech labels to the target user.

2.3   Stylistic labels

Stylistic labels mainly include labels indicating style levels (e.g. *formal/informal*, *slang* or *vulgar/elevated*, *colloquial*) and labels indicating frequency of usage (e.g. *mostly used in the plural form or singular form*, *rare,* etc.). This is illustrated by the examples under (1).

(1)   **machin**   N. (*pop.*) *J'ai vu machin*, my'adyena wónó. *Un machin à nettoyer les bouteilles*, mandè yi sogin'imbute.
        méli-mélo   N. (*fam.*) bindaka-bindaka.

In the examples under (1), the lexicographer uses the labels (*pop.*) and (*fam.*) to warn the user against the application of the lexical items *machin* and *méli-mélo* in formal conversation. In the treatment of the article of the lemma **machin**, in particular, the user is provided with co-text examples illustrating its typical use in local French. Each co-text example is followed by its translation equivalent in Mpo̱ngwe̱.

2.4   Labels for special fields of activity

Labels in this category identify the special area of knowledge to which a lemma or an entry applies:

(2)   **joint**   N. (*maçonnerie*) iñanga-ñanga; (*menuiserie*) nguga, ndatizo.
        brasse   N. eguwa. *Une brasse d'étoffe*, eguwa z'onamba; *demi-brasse*, erene z'eguwa; ntsandè y'onamba. (*Natation*) *nager à la brasse*, dyóg'erongè.

Immediately after the lemma sign (in bold) and the part of speech indicator appears the label (*maçonnerie*) (masonry), followed by the translation equivalent (in italics) and another label (*menuiserie*) (joinery) plus its translation equivalents. Through the application of these labels, the user is immediately informed about the professional fields to which the lemma belongs. In the treatment offered in the article of the lemma **brasse** (cf. Hausmann and Wiegand 1989: 353), the user is provided with both macro- and microstructural data. The first data types are referred to as the comment on form (e.g. the lemma sign **brasse** and the item giving the part of speech, N.), whereas the second data categories are known as the comment on semantics (paraphrase of meaning/translation equivalent). The latter is divided into two sections. The first section deals with the typical use of the lemma **brasse** plus its co-text examples. The second section is introduced by a sphere of usage label (*Natation*) making the user aware of the specialised field (swimming) to which the lemma belongs.

2.5   Labels conveying information on pragmatics, syntax and semantics

The *Dictionnaire Français–Mpongwé* also makes use of italicised words in parenthesis after the lemma or an entry in order to convey information on pragmat-

ics, syntax and semantics. This is illustrated by the examples in Table 4 and under (3).

	Lemmata	Contextual data
1	amphisbène	N. (*serpent dit «double-tête»*) mbumba-ntse.
2	ampoule	N. (*cloque*) iwowi. (*Electrique*) intsó ñi lambi.
3	ascaride	N. (*ver intestinal*) ogola.
4	candeur	N. (*ingénuosité*) mbondwè; (*pureté d'âme*), nkèrè y'orèṁa.
5	cangue	N. (*supplice*) ompango.
6	canicule	N. (*époque des grandes chaleurs*) enongo-nongo.
7	cimeterre	N. (*sabre recourbé*) ambóndóngó m'okwara.
8	cleptomanie	N. (*manie du vol*) nkanie y'idyufa.
9	colonie	N. (*possession*) ntse yi re bóngo-pa. (*Population*) mbèïy'iinóngó.
10	digue	N. (*chaussée*) ogomba w'ado; (*barrage de pêche*) nkumba.
11	éructation	N. (*renvoie*) mbyogo.
12	famille	N. (*sens large*) mbuwe; (*sens restreint*) ngwèmo, inu.
13	incessamment	ADV. (*sans délai*) du go tètètè vènó.
14	mélopée	N. (*chant monotone*) orangwa-rangwo, ...
15	patriarche	N. (*chef de famille*) onèró wi mbuwe; ...
16	pou	N. (*de tête*) mbina; (*de corps*) mbanda; (*de pubis*) ndjili
17	profane	N. (*non initié*)

Table 4: Italicised words in parenthesis used as labels in the *Dictionnaire Français–Mpongwé.*

(3)   **affection**   N. itónda; (*en médecine*) nkani.
       affranchir   V.A. (*un esclave*) dandun'osaka; (*une lettre*) postu'ezango.
          S'affranchir, dandwa; *un affranchi*, olanduno.

Considering the data presented in Table 4, one can question the use of these contextual data in the form of labels since an explanation (cf. Zgusta 1987: 5) as given in the example in Table 5 below is particularly necessary if the translation equivalent offered is not an absolutely exact one. An entry like *mbyogo*, Table 4 number 11, is perhaps a unique translation equivalent of the French word *éructation*. It is an insertible equivalent (cf. Zgusta 1987: 5) and *a priori* does not need any gloss. The italicised explanation given in parenthesis is meant in the first place for speakers of French to help them disambiguate the meaning of the lemma **éructation**. But the entry *mbyogo* itself has both a translational and an explanatory power because it can be inserted into any co-text example of contemporary Mpongwe. The same line of argumentation holds true for the other members of Table 4. With the entry (*serpent dit «double-tête»*), Table 4 number 1, the user is referred to the denotatum. All these entries help the user both to contextualise and co-textualise. They have both an encoding and a decoding function. Further examples include those under (4).

(4)   **blé**   N. (*farine de.*), fariña y'atanga.
       blesser   V.A. bóndja. *Blesser quelqu'un*, bóndj'oma evora. (*Fig.*) *Ce mot l'a blessé*, igamba mè-nónó ñ'atuṁa ...

> **borne**   N. (*de pierre.*) ido ñ'okao, elingilyo z'okao. (*Par ext.*) *dépasser les bornes*, pasiza.

In the article of the lemma **blé**, this is the information the user is expected to retrieve from the macro- and microstructural data presented: **blé**, noun whose typical syntagmatic pattern is *farine de blé* and whose translation equivalent in Mp<u>o</u>ngw<u>e</u> is *fariña y'atanga*. In the treatment of the articles of the lemmata **blesser** and **borne**, the user is provided with two kinds of semantic data. Subsection one contains the dominant sense of the lemmata (the one that would first occur to the target user) plus their co-text examples. Subsection two is introduced by the labels (*Fig.*) and (*Par ext.*) (also (*sens restreint*) and (*sens large*) in the article of the lemma **famille**, Table 4 number 12) denoting distinctions in meaning of the treatment units. Despite the absence of typographical structural markers (i.e. the diamond ♦ or, as it seems to be the case in the dictionary, the dash —) or ordering devices, as far as this dictionary article is concerned, these labels also create divisions in the lexicographic text (cf. Hausmann and Wiegand 1989: 338-339). As such, they form part of the inner rapid access structure of a dictionary because the user is merely interested in the metaphorical use of the lemma and immediately gets the information about the meaning that is looked for and thus does not have to work through the entire article. This point is illustrated in the treatment of the article of the lemma **partir**:

(5)   **partir**   V. N. kènda; dubwa. *Partir à Kango*, kènda go Kango. *Partir de Néndé*, dubwa go Néndé. *Etant parti de là*, ayè lubu gogo; ayè gogo *Partir quelques minutes*, kènd'ogèndó wa gore; *pour revenir le même jour*, Kènda kili. *A partir d'aujourd'hui*, avilo nèno winó, avil'owendja wa nènó. — *L'Arambo part du Mont-bouët*, olówi w'arambo wi pakilya g'Indóku. *Le coup de fusil est parti tout seul*, ndjali yakamba ou yabaṁina yomè. *La foudre part des nuages*, ndjalitowa yi pila go vindi.

From the typographical exposition applied in the example under (5), the user can retrieve information belonging to two different search areas. The first search zone is concerned with co-text examples in which the lemma typically occurs, whereas in the second search zone, the user is provided with additional competence examples taken from daily conversation.

　　Apart from using the dash as a typographical structural marker, others also come into play. Lemmata are printed in bold face, items giving the translation equivalents are printed in roman type and examples and meaning explanations in the source language (French) are given in italics. The presentation of the lemmata set apart typographically from the rest of the lexicographic text has much improved the access structure of the dictionary. Along with typographical structural markers non-typographical structural indicators are used to achieve a metacommmunicative function. Commas are used to separate translation equivalents of the same polysemous sense of the lemma, whereas semicolons separate translation equivalents representing different polysemous senses of the lemma (cf. Gouws 1999a: 12). As already mentioned, both typographical and non-typographical structural indicators have the same genuine purpose (as

explained by Wiegand 1999: 299) of assisting the user in retrieving the relevant information as quickly as possible. Unfortunately, they are not explicitly explained in the introductory notes. Judging from these data, one would conclude that the dictionary is compiled for well-informed users, persons well familiar with dictionary-using skills. But since there is no metatext or outer text advising the target user how to use the dictionary, the well-devised microstructure of the dictionary is likely to be unknown to him/her.

From the system applied in the dictionary, it seems that the hyphen has a double function. Firstly it is used as a typographical structural marker as it occurs in the example under (5). But it is also used to substitute the lemma in the illustrative examples, functioning as a place-keeping symbol. Compare the following example from the dictionary:

(6)   **aborder**   V. N. suṁa, dyèza. *Aborder à Ovéndo — à l'île Perroquet*, suma g'Ovéndo; dyèza go Mbini. *Aborder une personne*, susa gor'oma.

This use of textual condensation procedures also reveals some inconsistencies. Consider the following example in this respect:

(7)   **bien**   N. mbya. *Un homme de bien*, om'ombya. *Rendre le bien pour le mal*, finizembya go mbe. *Le bien public*, mbya y'anaga waodu. (*Richesses*) *tous ses biens*, aniv'imè modu, sik'iyè yodu. *Le bien d'autrui*, ya yi ngani.
ADV. mbyambye, kwèkwè. *Tu as bien agi*, o denda mbyambye, o denda kwèkwè. *Très bien*, mbyambye polo. *Mon fardeau est bien lourd*, irwano ñami ñ'adiri polo. *Bien des gens*, anag'awenge, ntango anaga. *Il y a bien un an*, atweni ga ompuṁa. *Tout va bien*, zel'osaṁu. *Ou bien*, ntsó.
L. CONJ. *bien que*, wanga go re, wanga go. *Bien qu'il soit jeune*, wanga go re yè oṁwango. *Bien qu'il pleuve*, wanga go nóge ningo.
INTERJ. *hé bien! qu'en dis-tu?* o buya sè rèti! *Eh bien! soit*, mbuku!

The lemma **bien** is polyfunctional for it can be used as a noun, an adverb, a conjunction and an interjection. The following represents a partial, annotated structural graph of this article:

(8)
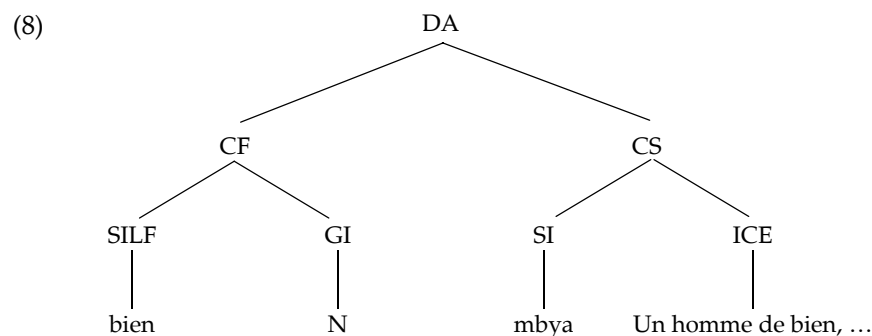
Abbreviations: DA = dictionary article; CF = comment on form; CS = comment on semantics; SILF = semantic item giving the lemma form; GI = grammatical item; SI = semantic item; ICE = item giving the competence example.

What is interesting here is that the target language examples (in italics) function in a non-lemmatic addressing procedure. In other words, they are not addressed at the lemma but at the source language competence examples. Thus the latter become treatment units within the article. Furthermore, the system applied in the dictionary is dominated by the user perspective. In fact, the compiler does not make use of textual condensation devices, which consist of decreasing the data types in order to save space (cf. Gouws and Prinsloo 1997: 51-55). Employing a textual condensation approach, the compiler should have used the tilde ~ or the dash — to substitute the lemma in the illustrative examples. By avoiding textual condensation procedures, the compiler has taken the needs and reference skills of the user of the dictionary into account. As a matter of fact, the user is not well familiar with dictionary-using skills. The way in which the compiler applied this system of textual condensation could have added to the predictability, systematicity and scientificity of the dictionary if it had been done consistently. Another area where the user profile also comes into play is the use of mediostructural representations. According to Gouws (1999a: 43), the mediostructure is "the system of cross-referencing which leads a user from a reference position to reference address". In the *Dictionnaire Français–Mpongwé*, it seems that the compiler only uses the external reference address. The following examples illustrate this point:

(9) (a) **celui, celle, ceux, celles**   PR. DEMONS. oma, mongi. *Celui ou celle qui a menti*, oma w'awanoka. *Ceux ou celles qui ont a menti*, oma w'awanoka. *Ceux ou celles qui ont volé*, mongi w'adyufa. *Celui-ci*, oma winó; *celui-là*, oma wónó; *ceux-ci*, mongi sinó; *ceux-là*, móngi sónó. *V. Grammaire*.

   (b) **être**   V. DEF. (*Se rend de différentes façons*). Re *Dieu est bon*, Añambyè are ombya. *Ils sont là-bas*, wi re gonó.
     — Bya. *La journée sera mauvaise*, owendja wi bya ombe. *Ces enfants seront des paresseux*, awana winó wi bya agera.
     — Ni. *C'est toi le voleur*, ofe ni wè; o dyufa ni wè. *Quelle est ta maison?* n'agw'ó ni ye?
     — Duo. *Il était ici ce matin*, aduo gunu ibanga. *Je serai là-bas demain*, m'be duo gogo mènè.
     — Dwana. *Ce village sera grand*, nkala yinó yi be dwane mpolo. *Sois tranquille*, lwana kende-kende.
     — Pegaga. *J'étais couché*, my'apegaga bu. *Ils étaient à Sisè hier*, w'avegagi dyao go Sisè.
     — (*est parfois sous-entendu*). *Je suis un Mpongwè. Vous êtes des imbéciles*, myè om̂wo-Mpongwè anwè lèngèlè y'anaga. *Cf. Grammaire …*

In example (9)(a), the user is provided with the item giving the part of speech as well as the translation equivalents. Then follow the co-text examples in French (in italics) and their translation equivalents in Mpongwe (in roman). The article ends with a cross-reference entry in which the user is referred to the mini-grammar of the dictionary where grammatical data are discussed in more detail. The reference entry *V. Grammaire* (Voir Grammaire/See Grammar) con-

sists of a reference marker (*V.*) and an entry indicating the reference address (*Grammaire*) (cf. Gouws 1999b: 7). Here one also finds a new type of macrostructural ordering, namely a nested lexemic paradigm[2]. As far as this point is concerned, lexicographers usually make use of niching or nesting procedures. An example of the use of a sinuous lemma file which is very user-friendly[3] occurs in *Cassell's French–English/English–French Dictionary* by J.C. Manchon (1951[5]). Manchon has interpreted the lemma sign **celui** as it occurs under (10) as a macrostructure with grouping.

(10)   **celui** (sə'lɥi)   *demonst. pron. m.* (*fem.* **celle** *pl.* **ceux, celles**) He, him; she, her; they, them; that, those. *Ceux qui ont vécu avant nous*, those who lived before us. **celui-ci, celle-ci,** This one, the latter; **ceux-ci, celles-ci,** these, the latter. **celui-là, celle-là,** That, the former; **ceux-là, celles-là,** those, the former. *Aimez-vous mieux celui-ci?* do you like this best? *celui-ci est meilleur que celui-là,* this is better than that; *celui-là n'est pas si beau*, that; one is not so fine. [*Celui-ci* relates to an object near the speaker; *celui-là*, to an object distant from him; or after two nouns already expressed, *celui-ci* refers to the last, *celui-là* to the first mentioned.]

Sublemmata **celle**, **ceux** and **celles** are listed according to their gender and number. Later in the article, the ordering of **celui-ci, celle-ci**, **ceux-ci, celles-ci,** and **celui-là, celle-là, ceux-là, celles-là** has not been done at random. Morphological considerations play an important role in the arrangement within this lemma cluster. In this nest, the compiler makes a distinction between demonstrative pronouns related to an object near the speaker and those related to an object distant from the speaker. The first category of demonstrative pronouns is given first. In the listing, priority is given to demonstrative pronouns with a masculine gender. Moreover, the singular form of the demonstrative pronoun dealt with comes first in the list, then followed by the plural form. The comment on usage at the bottom of the article given in square brackets is an example of a good lexicographic practice which contains useful information for the user.

In example (9)(b), the reference entry *Cf.* is used instead of *V*. In addition, the article also displays features of textual condensation devices. Criticisms have already been made with regard to the absence of a user-driven approach as well as inconsistencies regarding the system the compiler applied. Front matter texts and back matter texts are not the sole reference address as far as mediostructural representations are concerned. In fact, mediostructural procedures usually include three categories, namely the internal reference address, the external reference address and the dictionary external reference address (cf. Gouws and Prinsloo 1998: 20-22). The article-internal cross-referencing functions within the boundaries of an article whereas the article-external cross-referencing refers a user to an entry in another article or another text in the dictionary (cf. Gouws 1999a: 43). The dictionary external reference address is especially useful to refer a user from a mini-grammar to a source outside the dictionary where a comprehensive account on grammatical issues can be found.

3.   Improving the access structure of the dictionary

As far as the access structure is concerned, a distinction is usually made between the outer and the inner access structure. The outer search path proceeds vertically through the carriers of guiding elements (i.e. from A to Z) while the inner search path starts at the lemma and proceeds horizontally through the dictionary article (cf. Hausmann and Wiegand 1989: 338).

Changes can be made at the level of the access structure. For example, Gouws (1996: 21) points to the following: "The *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Dictionary* (hereafter abbreviated as RD) owes its poly-accessible character to the inclusion of additional outer access structures ... The original macrostructure of RD is presented in two columns occupying the centre of each page. An additional column is included on either side of the central word list." This metalexicographic innovation could be used to convey cultural or pragmatic information to the user as illustrated in the treatment of the article of the lemma **marché** in Table 5.

Ikasa, primitivement pont, jetée, a pris ultérieurement la signification de marché public, parce que le premier marché établi dans la colonie était installé sur la jetée de Libreville (A. Raponda-Walker: *Les langues du Gabon*, 1998, 148).	**marché** N. (*lieu public*) ikasa, igolino. *Affluer au marché*, suzana g'ikasa, bengana g'ikasa. (*Transaction*) igolo. *Conclure un marché*, teniz'igolo, kol'igolo. *A bon marché, igolo iwaya, igolo ilèle.*

Table 5: The treatment of the article of the lemma *marché* as it occurs in the *Dictionnaire Français–Mpongwé*.

This is the kind of presentation that may be used if changes to the central list of the existing dictionary is not possible in a revised reprint or reissue. A so-called "secondary macrostructure" may be introduced. Next to the previous macrostructure new data complementing or updating previous data may be added as demonstrated in the two columns of Table 5: the right column showing the previous macrostructure and the left column giving the new added macrostructure.

Instead of using metalexicographic innovations of the kind of the RD, one can also improve the access structure of the dictionary by means of inserted inner text blocks. Compare the treatment of the same lemma as it occurs in Table 6:

marché
Ikasa, primitivement pont, jetée, a pris ultérieurement la signification de marché public, parce que le premier marché établi dans la colonie était installé sur la jetée de Libreville (A. Raponda-Walker: *Les langues du Gabon*, 1998, 148).

Table 6: The treatment of the lemma *marché* by means of an inserted inner text block.

Apart from the main text or central list of the dictionary, the entering of **marché** as part of the inserted inner text will give the work the feature of a poly-accessible dictionary. The user seeking the meaning of the lemma **marché** will proceed along the inner search path in order to find information regarding both the comment on form (mainly the item giving the part of speech) and the comment on semantics (e.g. translation equivalents and examples). However, if the user needs information of a pragmatic nature, he/she will have to consult the inserted inner text, which in this case is a secondary macrostructure or middle matter (cf. Hausmann and Wiegand 1989: 338). In addition, the treatment of lemmata such as **moustiquaire**, **fourmis**, **parapluie**, **nappe de table**, etc. can be improved by providing the user not with inserts but by making provision for short comments on pragmatics. Such comments on pragmatics can be found in Raponda-Walker (1933). This is relevant because these additional texts may help the user to disambiguate the meaning of the treated lemmata. They also provide the user with morphological and syntactical clues. In this way, the user will be well aware of the cultural value as well as the semantic load of the lexical items that he/she uses in his/her daily routine.

4.   Conclusion

The intention of this article was to give an overview of issues relating to labelling in the *Dictionnaire Français–Mpongwé*. Lexicographic data whose occurrence is restricted to certain geographical areas, subject fields, professions, etc. are generally acknowledged in the dictionary. The use of contextual entries as well as labels in the dictionary articles does not only add to the predictability of the dictionary but it is also a user-driven approach. Apart from some shortcomings, the target user of the dictionary is consistently aided by guidelines enabling him/her to choose and use the right translation equivalent in the right context. However, Raponda-Walker's work has failed to supply lists of labels in the outer texts of the dictionary. The dictionary displays a well-devised access structure with a Mpongwe speaker in mind. The use of typographical and non-typographical structural markers, largely predictable, have improved the quality of the dictionary. There are still some shortcomings especially in the areas of textual condensation (lack of uniformity in the treatment) and mediostructural representations. When dealing with borrowed words, data categories regarding the donor language (i.e. the item giving the source language as well as the one giving the source language form) should be acknowledged in future (revised) editions. For example, if the word *mangazè* is proven to be a contemporary example of oral usage of the Mpongwe lexicon, then it has to be labelled as *corrupted form of the French word* magasin.

For all lexical items labelled in the dictionary as *néologisme* and *anglais*, research has to be conducted through fieldwork at grassroots level. A survey of these lexical items may improve the dictionary's standard with regard to its label policy. Labelling is the lexicographer's way of dealing with lexical items which deviate from the standard language. Therefore there is a need for accu-

rate labelling. In addition, problems relating to labelling also include the absence of a clearly defined and consistent policy on labelling. Attempts should be made toward the standardisation of labels across dictionaries (cf. Bureau of the WAT 1999: 10 and 14).

Endnotes

1. The spelling used here is the official Gabonese spelling and not the traditional semiphonetic spelling employed by Guthrie, Jacquot, Kwenzi-Mikala and others. The underlined e̱ and ṉ represent [ɛ] and [ŋ] respectively.
2. This information comes from Prof. R. H. Gouws (lectures).
3. It is user-friendly in the sense that the compiler has made the typographical presentation of the lexicographic text in such a way that it is clearly perceived by the user. The clustering is made "visual" by the vertical line to the left of the text. Wiegand (1989: 392) has also pointed out a similar example from the *Dictionnaire du français vivant* (DFV) by Davau et al. See also Smit (1996: 178-179).

Bibliography

Dictionaries

Davau, Maurice, Marcel Cohen and Maurice Lallemand. 1981. *Dictionnaire du français vivant*. New edition, completely revised and enlarged. Paris: Bordas.

Manchon, J.C. 1951[5]. *Cassell's French–English/English–French Dictionary*. New York: Funk and Wagnalls Company.

Raponda-Walker, A. 1930–1934. *Dictionnaire mpongwè–français, suivi d'éléments de grammaire*. Metz: Imprimerie de la Libre Lorraine.

Raponda Walker, A. 1961. *Dictionnaire français–mpongwè*. Brazzaville: Imprimerie St Paul.

Raponda-Walker, A. 1995[2]. *Dictionnaire Français–Mpongwé*. Libreville: Éditions Raponda-Walker. Classiques Africains.

Other literature

Benson, M., E. Benson and R.F. Ilson. 1986. *Lexicographic Description of English.* Amsterdam/Philadelphia: John Benjamins.

Bureau of the WAT. 1999. *Study Guide of the Woordeboek van die Afrikaanse Taal* (WAT). Unpublished Course Notes.

Gouws, R.H. 1996. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 15-31.

Gouws, R.H. 1999a. *Equivalence Relations in Translation Dictionaries*. Unpublished M.A. Course Notes.

Gouws, R.H. 1999b. Mediostructural Representation, Textual Condensation and User-orientation in the WAT X. *Lexicographica* 15: 4-37.

Gouws, R.H. and D.J. Prinsloo. 1998. Cross-Referencing as a Lexicographic Device. *Lexikos* 8: 17-36.

Hartmann, R.R.K. 1989. Sociology of the Dictionary User: Hypothesis and Empirical Studies. Hausmann, F.J. et al. (Eds.). 1989-1991: 102-111.

Hausmann, F.J. and H.E. Wiegand. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. (Eds.). 1989–1991: 328-360.

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: De Gruyter.

Landau, S.I. 1991. *Dictionaries: The Art and Craft of Lexicography*. New York: C. Scribner's Sons/Cambridge: Oxford University Press.

Norri, J. 1996. Regional Labels in Some British and American Dictionaries. *International Journal of Lexicography* 9(1): 1-29.

Norri, J. 2000. Labelling of Derogatory Words in Some British and American Dictionaries. *International Journal of Lexicography* 13(2): 71-106.

Nyangone Assam, B. and P.A. Mavoungou. 2000. Lexicography in Gabon: A Survey. *Lexikos* 10: 252-274.

Raponda-Walker, A. 1998. *Les langues du Gabon*. Libreville: Éditions Raponda-Walker.

Raponda-Walker, A. 1933. Les néologismes dans les idiomes gabonais. *Journal de la société des Africanistes* 3(2): 305-314.

Smit, M. 1996. *Wiegand's Metalexicography as a Framework for a Multilingual, Multicultural, Explanatory Music Education Dictionary for South Africa*. Unpublished D.Litt. Thesis. Stellenbosch: University of Stellenbosch.

Wiegand, H.E. 1989. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: Alphabetische Anordnungsformen und ihre Probleme. Hausmann, F.J. et al. (Eds.). 1989-1991: 371-409.

Wiegand, H.E. 1999. *Semantics and Lexicography. Selected Studies (1976-1996)*. Edited by Anntje Immken and Werner Wolski. Tübingen: Max Niemeyer.

Zgusta, L. 1987. Translational Equivalence in a Bilingual Dictionary. *Dictionaries* 9: 1-43.

The Compilation of the Shona–English Biomedical Dictionary: Problems and Challenges[*]

Nomalanga Mpofu *(nomalanm@yahoo.com)* and
Esau Mangoya *(emangoya@arts.uz.ac.zw), African Languages Research Institute (ALRI), University of Zimbabwe, Harare, Zimbabwe*

Abstract: The bilingual Shona–English dictionary of biomedical terms, *Duramazwi reUrapi neUtano,* was compiled with the aim of improving the efficiency of communication between doctor and patient. The dictionary is composed of terms from both modern and traditional medicinal practices. The article seeks to look at the methods of production of the dictionary, the presentation of entries in the dictionary and the problems and challenges encountered in the compilation process, namely, developing Shona medical terminology in the cultural context and especially the aspect of equivalence between English and Shona biomedical terms.

Keywords: BIOMEDICAL, ADOPTIVES, ENTRIES, SYNONYMS, CROSS-REFERENCES, IDIOMS, CIRCUMLOCUTION, STANDARDISATION, HEADWORD, EQUIVALENCE, VARIANTS, DEFINITION, CULTURE, EUPHEMISMS, MODERN, TRADITIONAL, MONOLINGUAL, BILINGUAL, CORPUS, BORROWING, SHONA, COMMUNICATION

Opsomming: Die samestelling van die Sjona–Engelse biomediese woordeboek: Probleme en uitdagings.   Die tweetalige Sjona–Engelse woordeboek van biomediese terme, *Duramazwi reUrapi neUtano,* is saamgestel met die doel om die effektiwiteit van kommunikasie tussen dokter en pasiënt te verbeter. Die woordeboek bestaan uit terme van sowel moderne as tradisionele geneeskundige praktyke. Die artikel wil die metodes van die totstandkoming van die woordeboek beskou, die aanbieding van die inskrywings in die woordeboek en die probleme en uitdagings wat in die samestellingsproses teëgekom is, naamlik, die ontwikkeling van Sjonamediese terminolgie binne die kulturele konteks en veral die aspek van ekwivalensie tussen Engelse en Sjona- biomediese terme.

Sleutelwoorde: BIOMEDIES, LEENWOORDE, INSKRYWINGS, SINONIEME, KRUISVERWYSINGS, IDIOME, OMSKRYWING, STANDAARDISASIE, TREFWOORD, EKWIVALENSIE, WISSELVORME, DEFINISIE, KULTUUR, EUFEMISMES, MODERN, TRADISIONEEL, EENTALIG, TWEETALIG, KORPUS, ONTLENING, KOMMUNIKASIE, SJONA

---

[*]   This article is a combination of two papers presented by the authors 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.

1.   Introduction

This article aims to highlight the problems and challenges faced by a team of four researchers of the African Languages Research Institute (ALRI) at the University of Zimbabwe in the compilation of the first biomedical dictionary in Shona, *Duramazwi reUrapi neUtano*. This project, started in August 2001, was undertaken in collaboration with the Institute for Continuing Health Education (ICHE) at the University of Zimbabwe's Medical School.

The dictionary is composed of terms from both modern and traditional medical practices. It was compiled with the aim of providing a tool for communication between doctors of mainly the younger generation and patients. There was a need for better communication between doctors and patients so that patient expectations could be fulfilled after a visit to the doctor. The situation that has acted as a barrier to communication between doctor and patient in Zimbabwe, is the training of doctors in English while the majority of people who will consult them, will use indigenous languages, in this case Shona.

Entry into university to study medicine is based on the marks obtained at A-level, proficiency in the indigenous languages not being one of the requirements. The situation is made more complex when young doctors are deployed to rural areas where only Shona is used as compared to urban areas where there may be instances of the use of English or cases of code-switching. It was foreseen that there would automatically be a communication problem because of the different languages and levels of language use. Quite often, there is also a generation gap between doctor and patient, since some of the doctors are young and fresh from medical school. As a result, there are cultural nuances in the language that are usually missed by the younger generation of doctors. Kotzé (1999: 91) states that language as a phenomenon forms part of the cultural stock of a community. Hence a cultural gap between any two people can bring about communication problems.

Language use also differs according to factors such as age, social status, educational background and geographical location. For example, an elderly patient may come to a young doctor and try to explain his/her sickness using veiled language. Because the young doctor does not have proficiency in Shona and as a result also lacks the cultural background, there is automatically a communication problem between them. The purpose of the *Duramazwi reUrapi neUtano* is therefore to try and address all the differences in communication that may be found among different individuals, especially those from different age groups. This dictionary thus wants to help address the needs of doctors to understand the terms and expressions used by patients, especially those of the older generation. It also wants to help standardise terms used by different age groups in different parts of the country. The targeted users of the dictionary are therefore the medical fraternity and the general public.

2.   Selecting the Headwords

The scope of the words selected for inclusion in this dictionary are biomedical,

that is to say, they are both biological and medical. The compilers tried to include as many terms as possible from the major areas of medicine, namely general medicine, gynaecology, ophthalmology, paediatrics, surgery and urology, to mention but a few. Terms are also both technical and general. Maintaining a balance in the inclusion of terms from these areas of medicine was one of the challenges faced. Preference was given to those terms mainly used in everyday medical practice and consultation.

The headword selection process was preceded by field research. Four people in the nursing field were engaged by ALRI to carry out field research. The data gathering technique mainly used by these field researchers was that of targeted questions on diseases and their symptoms. These discussions were tape-recorded and later transcribed. The informants included medical doctors, nurses, other medical personnel and the general public.

Four traditional healers from the Zimbabwe Traditional Healers' Association (ZITHA) were also used as reference group for traditional medical terms and practices. They were consulted for terms from the traditional medical perspective referring to diseases, their symptoms and treatments. They were also helpful as far as beliefs associated with or linked to some ailments are concerned. From the Shona traditional perspective, beliefs were seen to play an important role in the understanding and treatment of diseases.

Finding headwords for *Duramazwi reUrapi neUtano* also included library research. This involved searching through medical and other general books such as dictionaries. One of the books that was helpful in finding headwords for the dictionary, was A.E. Strover's *A Shona–English Phrase Book* (1965), which contains questions in English addressed by the doctor to the patient. In this scenario, the doctor is not a native Shona speaker and he/she needs a translator to be able to communicate with the patient. The questions range from the patient's medical family history to gastrointestinal ailments and gynaecological problems, roughly translated and in some cases transliterated into Shona. No explanations or definitions of terms relating to ailments are given in this phrase book.

Shona bilingual and monolingual dictionaries, namely *Standard Shona Dictionary* (1974), *Duramazwi: A Shona–English Dictionary* (1981), *Duramazwi reChiShona* (1996) and *Duramazwi Guru reChiShona* (2001), were also used as sources for headwords. English medical books such as *Blackwell's Dictionary of Nursing* (1994) were consulted, especially during the defining stage as these provide guidelines pertaining to the causes, symptoms and effects of ailments.

3.   Presenting the Information

3.1   Entries

Duramazwi reUrapi neUtano is a bilingual dictionary. It is divided into two sections. The first section comprises Shona headwords with each headword having an English equivalent. The entries in this section consist of the headword in

Shona, followed by a variant or variants where applicable. After the variant comes the English equivalent, followed by the definition(s). After the definition(s), synonyms of the headword are given. Two illustrations of entries from the dictionary are presented under (1):

(1)   **chipfunga** [chipfungwa] bilharzia. Chirwere chinokonzerwa nekupindwa neutachiwana hwemumvura isina kuchena, chinoonekwa nekuita weti kana manyoka ane ropa. Zvimwe zviratidzo zvinoti kutemwa nemusoro nekurwadziwa nemudumbu. FAN *bharaziya, muhamba*.

(An illness caused by bacteria found in contaminated water that results in passing bloody urine or faeces. Other symptoms include headache and stomach pains. SYN *bharaziya, muhamba*.)

dhepo depoproveria. Jekiseni rinobayiwa vanhukadzi senzira yekudzivirira kubata pamuviri.

(An injection given to women as a form of contraceptive.)

Not all entries in this section have definitions. Synonyms and variants are cross-referenced by using *ona* (see) to the more frequently used word. Examples of this treatment are given under (2):

(2)   **chipfungwa** bilharzia. ONA *chipfunga*.

chizvaro womb, uterus. ONA *chibereko*.

-fufudzika miscarry. ONA *-pfupfudzika*.

The second section of the dictionary is a reversal of the articles in section one. Here the English headword is the main entry followed by the Shona equivalent(s). As illustrated under (3), the reversal is in alphabetic order:

(3)   **diabetes mellitus** chirwere cheshuga, shuga.

lymphogranuloma venereum mverapi, mvirapo.

malaria marariya, marungu, musarara, rumungwe.

male puberty dzvito.

xeroderma -oma ganda.

As can be seen from the **malaria** example, all the Shona words referring to *malaria* are given as one entry separated by commas. This method of presentation was used to avoid repeating the same English headword for every Shona term and its synonyms. No definitions are given in the English–Shona section because it was generally felt that there are many medical books and dictionaries in English giving information on and explanations of these terms.

3.2   Illustrations

The dictionary is also illustrated. The illustrations are in the middle section of the dictionary between the Shona–English and the English–Shona sections. The main parts of the human anatomy were selected for illustration, these including the skeleton, the head and its different parts, the reproductive organs and the leg. In all, 14 pages of the dictionary are devoted to illustrations. These illustrations are captioned in Shona.

4.   Solving the Problems and Accepting the Challenges

4.1   Equivalence

In bilingual translation dictionaries, the source language headword has to be rendered by an equivalent word in the target language. In this dictionary, it entailed translating Shona medical terms into English in the first section and vice versa in the second section. Some terms may be called universal or culture free, causing no problems as far as finding equivalents for them in another language is concerned. Among these can be classed general terms such as *dhokotera* (doctor), *chirwere* (disease) and *chipatara* (hospital); parts of the human body such as *ziso* (eye), *musoro* (head), *gokora* (elbow) and *datira* (hamstring); and terms referring to general diseases such as *gomarara* (cancer), *gwirikwiti* (measles) and *manyoka* (diarrhoea).

Problems of equivalence between Shona and English biomedical terms occur when some medical conditions are culture bound. Because of the cultural differences which these two languages often show as a result of different belief systems, there is not a one-to-one relation as far as medical terms are concerned. Both languages therefore present term gaps in the translation of some words. In the Shona–English section, the result of these term gaps was the filling of the slot for the English terms by an explanatory equivalent rather than just a corresponding translation. In other instances, a near equivalent had to be given. For instance, in the case of the term *runyoka* which is, according to Shona custom, a charm believed to cause punishment to an unfaithful partner, an explanation had to be given because no such practice is known in English society. Therefore, it is not only a problem of the lack of equivalent terms between the two languages but also differences in the belief of the causes of certain medical conditions.

4.1.1   Circumlocution

The conceptualisation of medical terms may vary across languages and cultures. Simple and standardised English terms often do not have direct equivalent terms in Shona. As a result circumlocutions sometimes had to be given as

Shona equivalents to precise and detailed single English terms, as can be seen from the examples under (4):

(4)   **allergy** kusawirirana kwomuviri nechimwe chinhu (literally, failure of the body to agree with something)

appetite kuda kudya (literally, to want to eat)

haematuria weti ine ropa (literally, urine with blood)

jaundice ruvara rweyero mumaziso (literally, yellow colour in the eyes)

menorrhagia -buritsa ropa rakawandisa pakutevera (literally, to lose a lot of blood during menstruation)

miscarry -bva pamuviri (literally, to lose pregnancy)

quadriplegia -remara maoko nemakumbo (literally, to be crippled or deformed in both arms and legs)

In these examples, the Shona equivalents are given in the form of longish phrases which the doctor and patient have to memorise. In a way, this defeats the idea of trying to create a standardised Shona terminology. However, the given phrases were the best possible equivalents the compilers could find.

Circumlocution must therefore often be used for finding suitable Shona equivalents for English medical terms. The goal of compiling the dictionary being to enhance communication between doctor and patient, the longer translations do not result in direct and natural communication. Every time the longer phrasal terms are used, they must be adapted to the form of the sentence. These longer version translations are subject to a variety of conceptualisations which in turn might misdirect and distort the sense contained in the source language terms.

Mtintsilana and Morris (1988) state that the disadvantage of circumlocution is that instead of one word the term consists of two or more words. This is especially cumbersome when a phrase has to take the place as headword. Circumlocution, though, cannot be avoided when dealing with two languages such as Shona and English that do not have a one-to-one correspondence for all the words, making a one-to-one equivalence impossible. Kotzé (1999: 91) states that in such cases one has to settle for equivalence of whatever kind between the two languages.

The best translations are those that would also give the equivalence in the same grammatical category. One would expect the equivalent of a noun to be a noun and of a verb to be a verb if the texts are to be equivalent. In the above examples, the Shona equivalents are in phrase form and the phrases involve words belonging to various grammatical categories. The given Shona version of the English term "miscarry" which is a verb, combines the verb stem *-bva* with the noun *pamuviri*. The verb stem *-bva* is therefore further qualified by the noun *pamuviri* in an attempt to capture the concept of miscarriage. In these phrases, words in particular grammatical categories are qualified by words

from different classes and grammatical categories. The phrasal equivalents given are evidence of terminological gaps in Shona. The introduction of these phrases plays a double role. Firstly it exposes these terminological gaps and secondly it encourages the finding of more suitable terms and the elimination of the initial inapt attempts.

In the dictionary, term creation was confined to a minimum. Only for very few selected entries where it was felt that their introduction would not bring about confusion, terms were created. This involved the application of Shona terms that are already being used in other spheres or fields, as can be seen from the examples under (5):

(5)   **anaesthetise** -fendesa (literally, make someone have a temporary sleep)

    blood bank bhangi reropa (literally, place where blood is preserved)

    cell chivako (literally, building brick)

Further examples of created terms for specific conditions are given under (6):

(6)   **cephalopelvic disproportion** mabhonzo echirume, mapfupa echirume (literally, male bones)

    colour blindness, achromatism -saona ruvara (literally, failure to see colours)

    laparotomy -ongororo yemudumbu (literally, examination of the inside of the stomach)

    night blindness -saona usiku (literally, failure to see at night)

    osteoporosis -pera mabhonzo (literally, wearing out of the bones)

    radiotherapy -rapwa uchipiswa nemagetsi (literally, to be treated by being burnt with electricity)

Certain diseases or disorders had to be described as a disease or disorder of a specific organ or part of the body. Hence there is a section with headwords that begin with *chirwere che-* (disease of …), e.g. rheumatoid arthritis is given as *chirwere chemagodo* (disease of the bones), arthritis as *chirwere chemakumbo* (disease of the legs), stomatopathy as *chirwere chemumuromo* (disease of the mouth) and sexually transmitted disease as *chirwere chepabonde* (disease of the reed mat).

Circumlocution results from the fact that, since modern medicine is a foreign field, medical concepts might not be in existence in a particular cultural group. The terminological gaps between English and Shona result from the fact that medicine in Zimbabwe is still largely practised in English and that doctors are also trained in English. The foreign concepts from specialised fields like medicine therefore remain unfamiliar to the general people who would under normal circumstances quickly create matching terms in their own languages. Where concepts exist in both cultural groups and are conceptualised in the same fashion the process becomes easy and natural.

That symmetry does not exist between languages communicating different cultures was acknowledged in this dictionary. What was aimed at was a symmetry in the content of the messages in the two languages involved (Fourie 1993: 82), and not a one-to-one correspondence between all the terms. The compilers also aimed at coining terms that would be acceptable to the users of the dictionary.

4.1.2   Traditional Medical Beliefs

Some medical terms are based on cultural and societal beliefs. Even the causes of some diseases are based on such beliefs. Barrenness in a woman is most commonly attributed to witchcraft and sorcery, which causes some people to consult traditional and faith healers. In instances where there were absolutely no equivalents, an explanatory equivalent was given, as can be seen from the examples under (7):

(7)   **kamhandara** medicine to tighten vaginal muscles …

   mukutura medicine for cleansing uterus …

   mutimwi charm tied around waist of child …

   sokora woman who has passed childbearing age ….

From these examples it can be seen that traditional healers have their own terminology based on their understanding and methods of diagnosing and treating ailments or disorders different from Western medical practice. Before the introduction of Western medicine sick people consulted traditional healers about ailments or disorders. As a consequence, traditional medicine has its own terms for diseases and their causes based on these beliefs. Terms from traditional medicine were incorporated in this dictionary to enable young doctors to enhance their understanding of the differences between traditional and Western medicine. The definitions given in the dictionary had to capture all these differences, some being pronounced while others being quite subtle. These differences are often the result of varying perceptions in the diagnosis and cause of a particular disease. For example, menstrual pain is *jeko* in Shona. Traditional medical practice distinguishes between *jeko gadzi* (literally, menstrual pain that is female in nature) and *jeko gono* (literally, menstrual pain that is male in nature). *Jeko gadzi* is said to be the milder type of menstrual pain that can be treated and does not cause fertility problems. *Jeko gono,* on the other hand, is the more painful type that is difficult to treat and is believed to be the cause of barrenness in women. Traditional healers, however, believe that *jeko gono rinogona kuputswa kuti riite gadzi*, that is, "breaking" or "taming" *jeko gono* into the milder type *jeko gadzi* is a way of treating it. These differences in understanding the cause and cure or management of an ailment or condition posed difficulties in finding equivalent terms between Shona and English biomedical terminology. Because of the lack of equivalent terms in English for *jeko gono* and

jeko gadzi, the compilers called them both menstrual pain or dysmenorrhoea and then highlighted their differences in the definitions.

Other diseases or disorders are also believed to be brought about by an enemy or a jealous person. Hence there are terms such as *chikandwa, chipotswa* and *chitsinga* whose near equivalent "form of rheumatism" was given. In Shona, it is believed that *muvengi anenge akukandira mushonga unokukuvadza* (an enemy used harmful medicine affecting your limbs), i.e. someone bewitched you. Other examples are given under (8):

(8)   **chipindira** failure by infant to suckle. Chirwere chinoita kuti mwana atadze kuyamwa mukaka uye chinonzi chinokonzerwa nokuti anenge abatwa nemunhu ane mhepo dzakashata.

(An illness that makes an infant fail to suckle, said to be brought about by the handling of the infant by a person with bad spirits.)

mubobobo magical sexual powers. Mushonga wokuromba unonzi kana munhurume ainawo anokwanisa kupindira chero munhukadzi waanoda achirara naye asi iye mukadzi wacho asingazvizivi.

(Magical powers that are said to enable a man who possesses them to go in and sleep with any woman he desires without the woman's knowledge.)

Because belief systems differ, the defining language made it clear that what is being explained is "said to be" so. It has been tried to demarcate between what is laid down as fact and what is commonly believed. The words *unonzi* (which is said to) in the **mubobobo** example and *chinonzi* (which is said to) in the **chipindira** example were therefore used to show that what follows is a cultural belief.

4.2   Adoptives

Some English terms and concepts used in the specialised field of medicine have never been utilised in Shona. This became a challenge to the compilers in that they had to introduce new concepts and completely new terms as well. The best solution was to borrow both the term and the concept. In this process, terms were adopted from English. These adoptives had to be adapted and naturalised to be identified, written and articulated as Shona words.

Examples of adoptives from English entered in their Shona form into the biomedical dictionary are shown under (9):

(9)   **andibhayotiki** antibiotic

areji allergy

bhiipii blood pressure, hypertension

gauti gout

> **karisiyamu** calcium
>
> **kondomu** condom
>
> **maigireni** migraine
>
> **meninjaitisi** meningitis

Entering these adoptives in the dictionary contributes to standardising them. Trudgill (1983: 161) views standardisation as the establishment of the agreed orthography in which some set of forms are written. Their inclusion helps to have them accepted and legitimatised as Shona terms usable in the medical field. Demonstrating the way they should be written will enable young practitioners to confidently use and write them and as a result be able to adequately communicate with their patients.

There are also cases where indigenous Shona terms exist alongside some of the adoptives. This posed a challenge to the lexicographers, for the adoptives have gained a more substantial use than the indigenous Shona equivalents. In these cases, it was considered that the adoptives would mostly be familiar to both the practitioner and the patient. The dictionary user would more likely search for the adoptive first before looking up the indigenous term. In such cases, the adoptives were entered as headwords. However, the purpose of compiling the dictionary being to enhance communication in Shona, the compilers agreed that there should be a bias towards the Shona terminology. As a result, despite the familiarity of the adoptives, the Shona equivalents were made to carry the definitions. Examples of such entries in which *ona* means "see" are given under (10):

(10)   **eidzi** acquired immunodeficiency syndrome. ONA *mukondombera*

        hepisi herpes. ONA *madzvausiku*

In the first example, *eidzi* is the adoptive which has been cross-referenced to the Shona equivalent, *mukondombera.* The Shona term *mukondombera* is now defined. So when users look up the common adoptive the dictionary refers them to the Shona term where the definition is given. Such is also the case with the second example where *hepisi* is cross-referenced to *madzvausiku.*

The idea is to start from the commonly used terms which in these cases are adoptives. Then the adoptives are cross-referenced to the not so commonly used Shona terms and in this way the users of the dictionary benefit by being introduced to the Shona terms. Their understanding is enhanced as the Shona terms now carry the meanings they share with the commonly used adoptives. The incentive for users to go to the Shona entries are the definitions carried by the Shona terms. Consulting the dictionary, users are cross-referenced from such adoptives to more information at the Shona terms.

The dictionary had to be as informative as possible. The biomedical dictionary project being a pioneering work of its kind, most of the Shona terms

have not yet been standardised. This is unlike the English situation where amongst many possible English terms, only the standard ones are now used in medical practice. In many instances, Shona has more than one term referring to the same concept. Again, for purposes of giving full information, cross-referencing was used.

4.3   Synonyms and Variants

Another challenge to the compilers was that, as was the case with adoptives and indigenous terms, Shona synonyms and variants also were in competition. In the absence of standardised Shona terms, a decision had to be made as to which entries would be defined and which ones would be cross-referenced. The problem was compounded by the fact that, in some instances, adoptives were competing with Shona synonyms. Some of the synonyms are the result of the existence of different dialects in Shona. The dialects that compose Shona are Karanga, Zezuru, Ndau, Manyika and Korekore. Some of these dialects have terms more peculiar to them than those of other dialects. This resulted in several Shona terms where English only has a single standard term. In the examples of synonymous terms from the various dialects under (11), adoptives from English are marked with an asterisk.

(11)   **cancer** mhuka, nhuta, nyamakazi, gomarara,*kenza

goitre humbu, *goita, sokorodzi, bukumuro

uterus chibereko, chizvaro, chirambo, nyoka

One of the major problems was what to include and what to exclude. If all of them were included, the size of the dictionary would be enlarged but the content value would be small. This would also compromise the quality of the dictionary because so many entries would have to be given for the same concept. On the other hand, there would be a limitation of communication if either the patient or the doctor encountered such terms but cannot find them in the dictionary. Young doctors would not be aware of the existence of these various Shona terms. It was therefore decided that these terms would be entered but for the purpose of saving space only one of them would be defined.

The decision to have all the synonyms entered created yet another problem. Criteria for choosing the headword which would be given the definition had to be found. A number of options were considered. These included corpus searches to find out which term featured most in the concordances. Despite the fact that the Shona corpus consists of about 3,5 million running words, the concordances were found not to be fully reliable, for the material used in the building of the corpus did not directly focus on the field of medicine and other scientific disciplines. The attempt to use the corpus showed that some medical terms are actually lacking. This is due to the fact that no medical material had been introduced into the corpus, the main reason being that the Shona corpus

has been established mainly for language and linguistic research. In its present state, the corpus cannot fully fulfil scientific needs.

One way of making a selection was to choose the term that had been popularised through both the electronic and print media. This would be a more standardised term used in all the dialects. As a result of this procedure, from the synonyms that denote "cancer" in the above example, the entry *gomarara* was chosen to be defined. This term has been extensively used during cancer awareness campaigns. In such cases, the popularised term would automatically be given the definition. As regards the synonyms in the other examples, those terms that were mostly used in all the dialects were considered suitable to carry the definitions. With this criterium, it was found that *chibereko* and *humbu* were the most familiar of the terms denoting "uterus" and "goitre" respectively. So the most familiar ones were defined and the rest of the synonyms were cross-referenced to them.

In the same way, the compilers had to deal with more than one term in cases where there were variants. Variants are alternative forms of words in a given context (Crystal 1980: 370) which are close to each other in articulation and as a result in spelling. This is even more so with Shona which has a phonetic spelling system. The variants are quite close in articulation and spelling as can be seen from the examples under (12):

(12)   **bilharzia** chipfunga, chipfungwa

      menstrual pain jeko, jekwa

      symptom chiratidzi, chiratidzo

Variants are given as separate entries. As was the case with synonyms, one of the variants had to carry the definition and again a criterium for choosing which entry to define had to be found. The definition was given at the most familiar of the variant terms. So the compilers had to consider all the variants to find the most familiar one which should be defined. In the first example, *chipfunga* was found to be more familiar than *chipfungwa*; in the second example, *jeko* was chosen above *jekwa*; and in the third example, *chiratidzo* was given preference to *chiratidzi*.

This was a huge task considering that a total of 497 of these variants and synonyms were entered in the dictionary. All of them had to be considered to decide which one would be defined and which other ones would have to be cross-referenced to it. Shona does not have standardised medical terms as all the doctors and nurses are trained and taught in English. No effort has previously been made to find suitable Shona terms.

4.4   Idioms and Euphemisms

Another challenge that presented itself to the compilers was the abundant use of idioms and euphemisms in Shona. These are used in situations where Shona

speakers refrain from making direct reference to certain terms.

When it comes to specialised fields, the history of a language is very important. The language develops or adopts terminology when a particular practice continues for a long time. English society is characterised by a long period of modern medicine practice. Often borrowing from Latin, English medical terminology has been well developed. There are certain English terms for which Shona does not have direct translations but renders them by idiomatic or euphemistic expressions. The idioms reflect the Shona people's sociocultural understanding of certain concepts. Idioms are phrasal statements, the meaning of which emerges when the words remain in a certain sequence. Some of these have always been used without ambiguity to express certain biomedical concepts in Shona. In the absence of standard medical terms, idioms had to be considered as entries in the dictionary as well. These idioms are used quite often in everyday conversations. Some of the idioms entered in the dictionary are given under (13):

(13)   **barrenness** -cheukwa netsuro (literally, to be mocked by a hare)

menstruate -enda kumwedzi (literally, to go to the moon)

sexually transmitted disease chirwere chepabonde (literally, disease of the sleeping mat)

The compilers were aware that these were not the real terms for the medical concepts they express, but at the same time it was not possible to create completely new terms. The compilers' efforts were directed towards enhancing communication between doctor and patient. In the absence of Shona terms, it was agreed that the idioms would be a temporary help. These can easily be replaced in future editions of the dictionary once more appropriate terms have been found. The compilers' mandate did not go as far as creating new terms. They therefore had to content themselves with what existed in the language. As older people would be comfortable in using these idioms, communication between patient and doctor would be clear. However, idioms are not the most appropriate communicative means in the specialised field of medical practice. They are best included in a general dictionary or a dictionary of literary and linguistic terms.

This lack of standardised terminology also surfaced in a slightly different form regarding the issue of euphemism. Euphemism is the replacement of a term seen as less refined or too direct by a more refined or less direct term. This is mostly evident when it comes to the naming of the private parts of the body and any biological functions associated with them. Mentioning them and describing their functions is considered to be highly obscene in Shona culture so that alternative descriptions for them have been created in an effort to diminish the obscenity which they are deemed to carry.

Another problem is that these euphemisms also tend to accumulate the negative connotation of obscenity with time. This results in a multiplicity of terms as new terms are created as a way of avoiding the earlier euphemisms

now also deemed obscene. This multiplicity of terms is demonstrated by the examples under (14):

(14)   **beche** (vagina) sikarudzi (creator of a tribe), mukana (passage), kunzira (the way)

mboro (penis) mhuka (animal), chirombo (a big thing), mbonausiku (a thing that sees at night), chombo (weapon), chirema (a cripple), sikarudzi (creator of a tribe)

-**svira** (copulate) -isa (put, insert), -vata (sleep with), -kwira (climb)

The first example *beche* has three euphemisms, the second *mboro* has six and the third example which is a verb *-svira* has three. What is being prioritised in the dictionary is effective communication between doctor and patient. Elderly patients would be more comfortable using the euphemisms.

The compilers had a problem with treating the entries of these euphemisms. On the one hand, the Shona community avoids the explicit terms for cultural reasons whilst, on the other hand, it is mandatory that there be effective communication between patient and doctor. It was resolved that euphemisms be entered. At the same time, it was felt that since the dictionary deals with a specialised field, there should be some initiative towards accentuating the explicit terms. As a result, all the euphemistic terms were cross-referenced to these basic terms normally considered taboo, so that the users of the dictionary can become acquainted with them. The compilers considered the dictionary as the first step towards using these terms in a direct manner. It was felt that the inclusion of the explicit Shona terms, giving them the important role of carrying the definitions, would encourage their application in medical practice.

In general purpose dictionaries, euphemism is widely employed when treating sensitive or obscene words. One of the challenges faced in this dictionary was defining private parts of the body. Euphemism has been avoided as much as possible in definitions because this is a medical dictionary using specialised language. Terms indirectly referring to parts of the body, due to society's norms of being polite, have been avoided as much as possible in the defining language. Hence, a term such as *sikarudzi* (creator of a tribe), an euphemistic reference to both the male and female genitalia, was only entered as headword but was not used in definitions. Thus, in defining sexually transmitted diseases, for example, reproductive organs were referred to by their real, not their euphemistic names. The compilers tried to be as direct as possible. The challenge was to avoid the conventions of the general purpose dictionary and keep the specialised nature of this dictionary in mind.

5.   Conclusion

Duramazwi reUrapi neUtano is a milestone in the development of Shona lexicography because it is the first scientific dictionary in one of Zimbabwe's local

languages. The aim of the dictionary is to be user-friendly to both the medical fraternity in order to enhance their efficiency and to the clients they serve. Although it is a relatively small dictionary, it will hopefully open up research of a scientific nature in the local languages and serve as a model for other African languages.

References

Dictionaries

Blackwell Scientific Editors. *Blackwell's Dictionary of Nursing.* 1994. Oxford: Blackwell.

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.

Hannan, M. 1984. *Standard Shona Dictionary*. Harare: College Press.

Mpofu, N., H. Chimhundu, E. Mangoya and E. Chabata (Eds.). 2004. *Duramazwi reUrapi neUtano.* Gweru: Mambo Press.

Strover, A.E. 1965. *A Shona–English Phrase Book*. Supplement to the *Central African Journal of Medicine* 11.

Other sources

Crystal, D. 1980. *A Dictionary of Linguistics and Phonetics.* Cambridge: Blackwell.

Fourie, D.J. 1993. Reflections on African Languages and Technical Translation Theory. *South African Journal of African Languages* 13(3): 81-85.

Kotzé, E. 1999. Translating Culture in Bilingual Dictionaries. *Lexikos* 9: 89-106.

Mtintsilana, P.N. and R. Morris. 1988. Terminography in African Languages in South Africa. *South African Journal of African Languages* 8(4): 109-113.

Trudgill, P. (Ed.). 1983. *Sociolinguistics: An Introduction to Language and Society.* London: Penguin.

Considering a Lexicographic Plan for Gabon within the Gabonese Language Landscape[*]

Hugues Steve Ndinga-Koumba-Binza, *Stellenbosch University Centre for Language and Speech Technology (SU-CLaST), Stellenbosch, Republic of South Africa (13751719@sun.ac.za / h_steve75@yahoo.fr)*

Abstract: This article raises a number of questions that should be dealt with in drawing up a lexicographic plan for Gabon. For which of the Gabonese languages should lexicographic units be established? This question entrains the issue of inventorying the Gabonese languages and their standardization as well as the issue of language planning for Gabon. What is the status of those foreign languages widely spoken in Gabon? What about French? Should Gabon keep importing its French dictionaries from France, or should the Gabonese compile their own French dictionaries, including French words and expressions exclusively used in Gabon? Finally, after trying to answer these questions, a number of suggestions are made for the establishment of a lexicographic plan for Gabon.

Keywords: GABONESE LANGUAGE LANDSCAPE (GLL), LEXICOGRAPHIC PLAN FOR GABON, LANGUAGE DIVERSITY, LANGUAGE POLICY, LANGUAGE STANDARDIZATION, GABONESE LEXICOGRAPHY, NATIONAL LANGUAGE, NATIVE LANGUAGES, FOREIGN LANGUAGES

Opsomming: Die oorweging van 'n leksikografiese plan vir Gaboen binne die Gaboenese taallandskap. Hierdie artikel bring 'n aantal vrae ter sprake waaraan aandag gegee moet word by die opstel van 'n leksikografiese plan vir Gaboen. Vir watter Gaboenese tale moet leksikografiese eenhede gestig word? Hierdie vraag bring die kwessie van die inventarisering van die Gaboenese tale en hul standaardisering ter sprake, asook die kwessie van taalbeplanning vir Gaboen. Wat is die status van daardie vreemde tale wat algemeen in Gaboen gepraat word? Wat van Frans? Moet Gaboen voortgaan om sy Franse woordeboeke uit Frankryk in te voer, of moet hy sy eie Franse woordeboeke saamstel, met insluiting van Franse woorde en uitdrukkings wat uitsluitlik in Gaboen gebruik word? Ten slotte, nadat geprobeer is om hierdie vrae te beantwoord, word 'n aantal voorstelle gemaak vir die totstandbrenging van 'n leksikografiese plan vir Gaboen.

Sleutelwoorde: GABOENESE TAALLANDSKAP (GTL), LEKSIKOGRAFIESE PLAN VIR GABOEN, TAALVERSKEIDENHEID, TAALBELEID, TAALSTANDAARDISERING, GABOENESE LEKSIKOGRAFIE, LANDSTAAL, INHEEMSE TALE, VREEMDE TALE

---

[*] An earlier version of this article was presented as a paper at the Ninth International Conference of the African Association for Lexicography hosted by the Groupe de Recherche en Langues et Cultures Orales (GRELACO) and held at the Omar Bongo University in Libreville, Gabon, 21–23 July 2004.

1.   Introduction

The aim of this article is to raise some questions Gabonese lexicographers might consider while drawing up a strategic plan for Gabonese lexicography. These questions are discussed within the Gabonese language landscape (henceforth GLL, a term coined by Ndinga-Koumba-Binza 2003). However, in order to set the framework for this article, the GLL as well as the current lexicographic plan of Gabon (henceforth LPG) are outlined before considering issues in drawing up the LPG.

2.   The Gabonese language landscape

2.1   The concept

A view or a picture of the countryside is referred to as a landscape. The term *language landscape* may be defined as an accurate picture of the languages, both foreign and native, used in a given geographical area (Ndinga-Koumba-Binza 2003). This concept applies to both multilingual countries and countries with language diversity, Gabon belonging to the latter group.

The concept of a language landscape can be wider than that of a language situation. Indeed, there are some differences between the two. The language (or linguistic) situation is viewed in terms of the native languages in a given country. Authors such as Jacquot (1978), Perrois, Jacquot and Moussavou (1983), Kwenzi-Mikala (1990), Mba-Nkoghe (1991), Blanchon (1994), Raponda-Walker (1998), Emejulu and Nzang-Bié (1999), Nyangone Assam and Mavoungou (2000), and Mavoungou (2002a), among others, have outlined the language situation in Gabon. A review of these studies shows that the linguistic situation in Gabon is described by highlighting the following four points (Mavoungou 2002a):

(a)   a mere acknowledgement of Gabon as a country with language diversity instead of a multilingual country,

(b)   a description of the particular status of French in Gabon,

(c)   an inventory and classification of the native languages, and

(d)   a brief description of the status of the native languages.

The term *linguistic situation* seems to be reserved for referring to the native languages; therefore the term *language landscape* is proposed for referring to the comprehensive linguistic countryside. When considering the language landscape, both native and foreign languages in the language diversity of the country are taken into account. Finally, the GLL reflects an on-field reality that can be broadly represented by the following figure:

Gabonese Language Landscape

```
                    Gabonese Language Landscape
                    /                        \
           Local Languages              Foreign Languages
            /         \                   /            \
      Official      Native        International    Non-international
      Language      Languages     Languages        Languages
         |            |               |                 |
      French      Language-units   e.g. English      e.g. Lingala
```
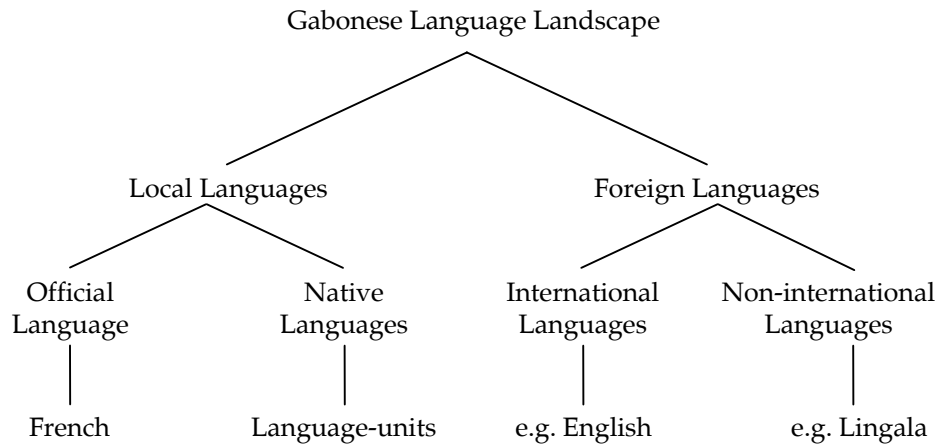
The GLL considers both local and foreign languages spoken in Gabon, since the linguistic countryside of Gabon comprises both local languages and foreign languages, as shown by the above tree diagram (taken from Ndinga-Koumba-Binza 2003 and forthcoming).

Local languages, which include native languages and the official language, refer to languages identified in accordance with a certain place through cultural inheritance (about the native languages, see Emejulu 2000 and 2003 for further details on the Gabonese heritage languages) and/or through linguistic policy (about the official language, see Mavoungou 2002b, Pambou 1998 and Ndinga-Koumba-Binza forthcoming for details on French as a Gabonese language according to the Constitution, and see Moussirou Mouyama 1984 and 1986 for details on French as a Gabonese language by reason of colonization). French has ceased being a foreign language in Gabon (Pambou 1998, Ndinga-Koumba-Binza forthcoming) since it is the only official language of the country, and the only language used in all official and public sectors (administration, business, education, media, etc.).

Foreign languages refer to languages not the mother tongues of the indigenous population of Gabon (see Addenda 2, 3 and 4), including the mother tongues of the migrant populations. Often these languages are excluded from the Gabonese language census (Blanchon 1994, Idiata 2002).

2.2   Native languages

The most recent work on making an inventory of the Gabonese native languages is that of Kwenzi-Mikala (1987, 1988, 1990 and 1998). According to his classification (see Addendum 1), he identifies 62 speech patterns (*parlers*) in Gabon that he groups into 10 language-units (*unités-langues*) on the basis of the

sociological criterion of starting a conversation with the formula "I say that" as well as the criterion of mutual intelligibility. Kwenzi-Mikala (1988: 57 and 1990: 122) states that a language-unit is a group or a set of different mutually comprehensible speech forms.

Many writers (Blanchon 1994, Emejulu and Nzang-Bié 1999, Nyangone Assam and Mavoungou 2000, Emejulu 2000, Idiata 2002 and Ndinga-Koumba-Binza forthcoming) have criticized the grouping of Gabonese languages into language-units on various aspects, mainly for being referred to as a classification and for lacking a suitable measurement for mutual intelligibility. However, in this article, Kwenzi-Mikala's grouping is viewed from the perspective of a language inventory. In fact, despite the criticism levelled at it, this inventory might sufficiently contribute to "a solution to the problem of the enumeration of all the heritage speech forms of Gabon" (Nyangone Assam and Mavoungou 2000: 254). Furthermore, it can be assumed that by using the terms *language-units* and *speech patterns* Kwenzi-Mikala tried to avoid the terms *languages* and *dialects*, which are very controversial in the Gabonese context at both political and linguistic levels. Thus, on the basis of historical dialectological criteria, it can be argued that Kwenzi-Mikala's ten language-units might be ten specific native languages found in Gabon. This is true within a historical dialectological approach seeking the standardization of these non-standardized Gabonese languages or speech forms. In other respects, many of the smaller speech forms are near extinction due to the very small number of their speakers.

2.3   Foreign languages

As far as Gabon is concerned, foreign languages are frequently spoken, since in a population of 1 379 000 inhabitants (Crystal 1999: 130), 15,2% are foreigners (according to a 2003 estimation). The fact that migrant communities speak their own languages, which in Gabon sometimes have more speakers than some native languages, should be reflected in the language census of the country.

In Ndinga-Koumba-Binza (2003 and forthcoming) the language immigration in Gabon is described as well as the settlement and spreading of the foreign languages throughout Gabon. However, apart from the sociolinguistic survey conducted between 1998 and 1999 by the Laboratoire des Sciences Humaines et de la Dynamique du Language (LASCIDYL) (Tomba Moussavou 2001), no systematic inventory of the foreign languages spoken in Gabon has ever been undertaken. Meanwhile, Tomba Moussavou (2001: 13) who is responsible for the survey of LASCIDYL, puts the figure at 49 African languages used on a daily basis in Libreville, the capital city of Gabon.

She also mentions six European languages (including French), one language from the Middle East (Arabic) and one unidentified Creole language. The latter seems to be one of the three forms of Creole found in São Tomé and Príncipe. Ndinga-Koumba-Binza (2003 and forthcoming) believes that field-

workers of LASCIDYL did not ascertain from their Creole-speaking informants the form of Creole they claimed to speak, nor did they pay attention to Asian languages, such as Japanese, Chinese and Korean, and to languages from Eastern and Southern Africa.

From both the Ethnologue records (Grimes 1996) and the work of LAS-CIDYL, and from on-field observations, Ndinga-Koumba-Binza (2003 and forthcoming) has established a wider list of foreign languages spoken in Gabon (see Addenda 2, 3 and 4).

The GLL as outlined above should be taken into account in both the language planning and lexicographic plan for Gabon.

3.   An embryonic lexicographic plan for Gabon

A lexicographic plan for Gabon (LPG) is still in its embryonic stage. By an LPG should be understood what Emejulu (2003: 205) refers to as "the strategic planning of Gabonese lexicography". Currently, Gabonese lexicography (Nyangone Assam and Mavoungou 2000; Emejulu 2001a, 2002b and 2003) is at five simultaneous phases in its development:

(a)   lexicographic training,

(b)   metalexicographical accounting for dictionaries,

(c)   circumscribing lexicography as a career,

(d)   formulating a coherent general framework for Gabonese lexicography, and

(e)   updating reference works inherited from missionaries and colonial administrators.

Although only recently established, Gabonese metalexicography experiences a fast-growing crop of literature. This is the consequence of an extensive number of Gabonese lexicographers being trained at the University of Stellenbosch. In fact, "the strategic planning of Gabonese lexicography gives the highest priority to training," as Emejulu (2003: 205) stresses.

At this stage, eight Gabonese students are undergoing doctoral training in lexicography and two have completed their doctorates in lexicography at the University of Stellenbosch (see Emejulu 2003: 205-207 and Gouws 2001 for further information on the training of future Gabonese lexicographers).

The two newly graduated doctors in lexicography have joined the staff of the Groupe de Recherche en Langues et Cultures Orales (GRELACO) at the Omar Bongo University, Libreville. In fact, in accordance with the terms of its mission, GRELACO conducts research in the field of lexicography and metalexicography headed by J.D. Emejulu and Y. Nzang-Bié, two Gabonese linguists specializing in lexicography (see Emejulu 2000 and Nzang-Bié 2002).

Thus, lexicography is becoming a career at university level as it draws scholars into basic lexicographic research. As part of Gabonese lexicography, a number of Gabonese lexicographers at present focus on the planning of dictionaries (Mavoungou 2002a, Nzang-Bié 2002, Afane Otsaga 2004) and the improvement of the quality of existing dictionaries (Nyangone Assam and Mavoungou 2000, Mihindou 2001, Mabika Mbokou 2004) within a metalexicographical framework.

Emejulu (2000, 2001a, 2002a and 2003) has proficiently defined lexicography as an opportune career in Gabon and formulated a coherent general framework for Gabonese lexicography. Mavoungou (2001a) has also contributed in this sphere by describing the inception of dictionary production in Gabon and presenting the challenges of globalization with special reference to dictionary-making in Gabon.

In general, the LPG is intended to be an economic asset in the development of Gabon (Emejulu 2000, 2003). Emejulu (2000: 65) says:

> For an immediate answer, in pure economic and commercial terms, lexicography is and should be treated as an economic and commercial variable that is governed essentially by market laws. It is, at least, job-creating in all its areas of execution and oriented to generate money. The development of such a realistic lexicography in Gabon will rely primarily on the strict adherence to basic economic and market principles both in its planning and execution.

Moreover, Emejulu (2003: 196) adds, "the remedial strategy for lexicography is to go beyond the higher institutions and reach out to broader society to unleash its full potential as an effective factor in development". Finally, the LPG is intended to develop in a general framework of the "dynamics for a viable and profit-making lexicography in a fast-changing, multilingual global melting pot", as Emejulu (2003: 197) puts it.

The organization of the Ninth International Conference of the African Association for Lexicography (AFRILEX) by GRELACO in Libreville, Gabon (July 2004), manifested a will-power for establishing lexicography in Gabon and for conceiving a specific lexicographic plan for this country.

Apart from the training of lexicographers and consistent publication-orientated research, GRELACO has equipped itself with reliable and modern technology and has signed a co-operation agreement with the Bureau of the Woordeboek van die Afrikaanse Taal (WAT) in Stellenbosch, Republic of South Africa (Emejulu 2003: 208). This agreement has allowed GRELACO researchers to receive practical computer-oriented training "in the various fields of production and marketing of dictionaries, data processing and management of lexicographic units, and strategic planning" at the Bureau of the WAT (Emejulu 2003: 205). The Bureau of the WAT has also agreed to assist GRELACO in the process of establishing lexicographic units across Gabon to cater for the need of recognized Gabonese languages (Emejulu 2003: 208). This agreement with the Bureau of the WAT is part of the GRELACO policy of co-operative lexicography (see Emejulu 2003: 207-209 for further information on this policy).

At present, it can be concluded that GRELACO has paved the way for a viable LPG, which is still to be elaborated.

No decision has yet been taken at government level with regard to dictionary-making for the Gabonese languages or the use of lexicographic techniques and models for the development of the Gabonese languages. Lexicography in Gabon is still confined to GRELACO and its research projects (most of them not sponsored by the government).

As for lexicographic practice, i.e. the practical component of lexicography (Gouws 2001: 59), two eras of dictionary production in Gabonese lexicography can be distinguished:

(a)   The earlier era, which includes reference works from missionaries and colonial administrators. Nyangone Assam and Mavoungou (2000), Mihindou (2001) and Mavoungou (2001b) have showed that the majority of lexicographic reference works available in the Gabonese languages are bilingual dictionaries, lexicons and glossaries compiled by Catholic and Protestant missionaries and colonial administrators from 1800 to 1960, when Gabon gained its independence from France.

(b)   The modern era, which starts from 2002 with the publication of *Gedandedi sa Geviya/Dictionnaire Geviya–Français* (henceforth DGF) by Van der Veen and Bodinga-bwa-Bodinga. Mavoungou (2004: 440) points out that the appearance of this dictionary opens a new era of dictionary publication in Gabon. The DGF was followed in 2004 by the *Lexique Pove–Français/Français–Pove* by Mickala Manfoumbi. This period coincides with the completion of the first theses on metalexicographical planning of dictionaries in the Gabonese languages (Mavoungou 2002a and Afane Otsaga 2004).

These two eras have this in common that dictionary production took place outside a general framework or strategic planning. None of these dictionaries was compiled on the basis of a theoretical conception, i.e. a metalexicographical design. In fact, in all countries lexicographic practice has the tendency to develop and become established in a pre-theoretical environment. This is why Gouws (2001: 59) can state: "Lexicographic practice is older than metalexicography."

The lack of a metalexicographical grounding for dictionary production in Gabon has caused that serious issues in the development of Gabonese languages have been overlooked, as will be seen in the following section. At the same time, the newly established Gabonese metalexicography has failed in effecting a concrete well-designed strategic planning, i.e. a LPG (cf. the five simultaneous phases in the development of Gabonese lexicography).

The following section deals with issues to be considered in drawing up an LPG. These include challenges in Gabonese linguistics, the status of the French language and the consideration of foreign languages. The question is how Gabonese lexicography will deal with these issues.

4.   Issues regarding the LPG

4.1   Challenges in Gabonese linguistics

Since the establishment of the Department of Language Sciences at Omar Bongo University in 1994, it has become clear that Gabonese linguistics faces a number of challenges. Though Gabonese linguistics might be said to have started with the work of Catholic and Protestant missionaries and colonial administrators, the majority of Gabonese languages are still undeveloped. According to Alberts (2004: xiv), "a language complies with the common definition of standardization if there are grammars, spellings and orthography rules, dictionaries and terminology available in that language". This definition might be suitable for a standard language which can be extended to a developed or modernized language.

Language standardization, development and/or modernization are not the only challenges for Gabonese linguistics. A study of the trends in Gabonese linguistics might show numerous linguistic issues with which the LPG will have to deal. Only a few will be mentioned, especially those that might be seen as important in the process of drawing up a language policy and/or language plan for Gabon.

4.1.1   A language policy and language plan for Gabon

The current language policy appears to be a continuation of the colonial language policy that ruled the use of French as the only official language (Kwenzi-Mikala 1990). Thus, even given the will of the government to promote the native Gabonese languages (see Nzang-Bié 2001 and Dodo-Bounguenza 2002), there is a great need for a new language policy. Ndinga-Koumba-Binza (forthcoming) argues that setting a proper language policy in place is a necessity for the development of the Gabonese languages.

The language policy should take into account the GLL and should include the whole process of language development, from fundamental research to school textbooks and dictionary compilation. Issues such as the number of languages, language standardization and the introduction of native languages into the education system should be dealt with within the framework of a language policy.[1]

The current language policy of Gabon is merely stated in one paragraph of the Constitution: "The Republic of Gabon adopts French as the offical language. Furthermore, she endeavours to protect and promote national languages" (Art. 2, par. 8). This one-paragraph statement constitutes the sole constitutional provision with regard to languages. It stipulates French to serve as "the sole medium of instruction in the national educational system, and as the language of business, in civil administration and the media" (Nyangone Assam and Mavoungou 2000: 255).

4.1.2   The native languages in Gabon

The number of native languages in Gabon has been an issue in numerous publications, and, as yet, there has not been any proposition that brings unanimity (see Kwenzi-Mikala 1987, Emejulu and Nzang-Bié 1999, Emejulu 2003, Idiata 2002 and 2003 for more information on this issue). The exact number of Gabonese native languages is still not known. According to Blanchon (1994: 1341), "due to important variation from one village to another and sometimes even between one individual to another, the exact number of languages or dialects is hard to determine".

However, the main reason for the uncertainty about the exact number of languages lies in the differentiation between languages and dialects. In fact, many dialects are regarded as different languages, and some languages, because of their limited number of speakers, are often seen as dialects of some other more extensive neighbouring languages. Ndinga-Koumba-Binza (forthcoming) points out that grouping both languages and dialects into various language-units without trying to distinguish between them (as Kwenzi-Mikala did), seems to be a way of resolving the issue.

An LPG may have to deal with this issue when answering the question for which of these languages lexicographic units should be established.

4.1.3   Language standardization

There is no standard language among the Gabonese native languages. The term *standard language* is here used in the sense of a form or a variety of any given language "which is learned and accepted as correct across a community or set of communities in which others are also used" (Matthews 1997: 352). Indeed, the 62 speech patterns distinguished by Kwenzi-Mikala are dialects and variants of various languages whose number will hardly reach ten, when the language-units of Kwenzi-Mikala are recognized.

Given the fact that "dictionaries reflect the characteristic way in which a language has become standardized in a given situation" (Gallardo 1980: 61), an LPG will also have to deal with the issue of language standardization when the question is considered for which languages lexicographic units should be established. It will be necessary to differentiate a standard variety of the language from other dialects. Therefore, an investigation will be needed to establish which process would be suitable for language standardization. In this matter, Kwenzi-Mikala's internal classification of language-units seems to be an important help. However, Afane Otsaga (2004) has recently argued that a standard translation dictionary, the first of its kind in Gabonese lexicography, might be used as an instrument for standardizing the Fang language. Earlier, Mavoungou (2002a) had suggested that the compilation of a trilingual Yilumbu–English–French dictionary might eventually contribute to the standardization of the Yilumbu language.

The standardization process of Gabonese languages will not only be "the choice and acceptance of the standard dialect by the speech community", but also "the choice and use of a unique spelling system for the reading and writing of these languages" (Afane Otsaga 2004: 218).

4.1.4   Writing system

The Gabonese native languages are still called non-written languages. However, most of these languages inherited considerable literatures from missionaries and colonial administrators. In fact, missionaries and colonial administrators established writing systems, alphabets, orthographies and spellings for the Gabonese native languages. These writing systems, alphabets, orthographies and spellings were mainly copies of their mother tongues, i.e. French, English and Portuguese.

Recently a number of proposals have been made for writing the Gabonese languages (see Idiata (2002 and 2003) and Hubert and Mavoungou (forthcoming) for further details). However, about the Gabonese orthographic alphabet adopted at a workshop in April 1999, unanimity has still not been reached, not only among linguists, but also between linguists and language practitioners and teachers. The orthographic alphabet, like many other Gabonese linguistic products, is still waiting for governmental approval and adoption for prospective national diffusion and implementation.

The role of an LPG will be to bring about, through dictionary production, the unanimity that is still lacking and to initiate the implementation of the proposed alphabet. The implementation of the proposed alphabet will provide a platform for the development of writing systems and a unique spelling system for the Gabonese languages.

4.2   The French language in Gabon

What about French in Gabon? Should Gabon keep importing its French dictionaries from France, or should the Gabonese compile their own French dictionaries, including French words and expressions exclusively used in Gabon? This implies that French should first be taken into account within an LPG. Mavoungou (2002b) has recently conceived a metalexicographical plan for a monolingual dictionary of the French language as spoken in Gabon. The first draft of this dictionary has been completed and the dictionary is ready for publication.[2]

The compilation of the dictionary entitled *Dictionnaire du Français du Gabon* (DFG) presents a view that considers French as one of the Gabonese languages. This view falls within the framework of the GLL. French, the only official language, being the language most often spoken and used, is definitely a Gabonese language, not only for being given the status of the only official language, but also for being the mother tongue of the younger generations in Gabon (Blanchon 1994, Pambou 1998, Moulougi Nguimbyt 2002).

Taking this view into account, an LPG should also have the following aims:

—   The production of all types of monolingual French dictionaries (LSP, learner's, school dictionaries, etc.) as in other French-speaking countries where the local French has been standardized (cf. Belgian French, Cameroonian French, Canadian French, Haitian French, Malian French, Senegalese French, etc.

—   The production of bilingual French dictionaries with other European languages as target languages (preferably languages taught in school, i.e. English, Spanish and German). Gabonese French can also serve as target language, other European languages being source languages of planned bilingual dictionaries.

The DFG seems to be only a repertoire of particular items in the French language as spoken in Gabon. These Gabonisms, usually designating specific extra-linguistic Gabonese realities (not in the French culture) are mostly items derived from native Gabonese languages (see Mavoungou 2002b for further details), e.g. **Ndjobi** (a secret society), **nyembwe** (sauce from palm nuts), **nkoumou** (meal from leaves of the *Gnetum africanum*).

Other items include, in addition to place names, also loan shifts, loan blends and formations based on native-language roots (this list of types of Gabonisms is not exhaustive).

However, the French language as it is spoken in Gabon does not only refer to Gabonisms (for details on the French language in Gabon, see Couvert 1982, Moulougui Nguimbyt 2002, Moussirou Mouyama 1984 and 1986, Ndinga-Koumba-Binza 2004, Ogden 1984, Pambou 1998 and Tomba Moussavou 2001). Therefore, the DFG should consider Gabonese French not only in its general core level, i.e. standard French, but also official French (acrolectal level), common French (mesolectal level) and popular French (basilectal level). All four levels are present in the French language as it is spoken in Gabon (see Mavoungou 2002b and Moulougui Nguimbyt 2002 for details).

Currently planned dictionaries (for instance Mavoungou 2002a and Afane Otsaga 2004) and a number of existing dictionaries have French as source or target language. This is the case with dictionaries such as Galley (1964), Van der Veen and Bodinga-bwa-Bodinga (2002) and Mikala Manfoumbi (2004). The *Dictionnaire Français–Mpongwé* (Raponda-Walker 1995[2]), the *Dictionnaire Français–Vili* (Marichelle 1902) and the *Dictionnaire Français–Ndumu et Ndumu–Français* (Biton 1907) could also, among others, be mentioned.

All these dictionaries view French as a foreign language. The LPG should consider French as one of the Gabonese languages.

4.3   Criteria and the nationality criterion

The LPG as currently designed focuses directly on the so-called Gabonese lan-

guages, i.e. the native Gabonese languages. However, Ndinga-Koumba-Binza (forthcoming) points out that these languages can hardly be called Gabonese languages since there is no official documentation from the state or government giving them such an identification.

In fact, as mentioned earlier, the Constitution of the Republic of Gabon only states the adoption of French as official language of the country (Art. 2, par. 8) and assigns the government to make efforts in protecting and promoting the national languages. The Constitution has to be questioned because none of the so-called national languages are mentioned in it. Moreover, there has not been any further governmental arrangements in favour of the native languages. Therefore, there is no stipulation differentiating the native languages from foreign languages. They are all just languages spoken in Gabon.

Many native languages could hardly be claimed to be Gabonese languages, since they could also be Cameroonian, Equatorial Guinean and/or Congolese languages, seeing that they are also spoken in these countries. This is true of languages such as Fang (also used in Equatorial Guinea and São Tomé and Príncipe, cf. Grimes 1996), Civili (spoken along the Atlantic Ocean coast from the southern province of Gabon to the northern province of Angola), Yipunu (spoken in Gabon, in the Republic of the Congo and in the Democratic Republic of the Congo), Yinzebi (also spoken in the Republic of the Congo), and Teke (also used in the Republic of the Congo and the Democratic Republic of the Congo).

Trends in current Gabonese lexicography show a definite focus on native languages on the basis of nationality without attention being paid to languages also used in other (surrounding) countries, though these foreign languages are spoken more in Gabon (according to the number of speakers) than some of the native languages. This is the case with Eveia compared to Yoruba, Igbo and Hausa (all three Nigerian languages).

In their recently published dictionary, Van der Veen and Bodinga-bwa-Bodinga (2003) mention in the pretexts that Eveia is a language with fewer than 50 speakers. The compilers even point out that the language is only spoken in one single village partly landlocked in the centre of the country. Such a language can obviously not contribute to a money-generating lexicography. In fact, the Eveia-speaking population has a very small impact on the Gabonese economy compared to Nigerian immigrants of whom the majority live in the capital Libreville and have an extensive influence on the national economy through their businesses (it is acknowledged that the Kalaba population has a certain control of the fishing industry at Pont Nomba and on the market of that area of Libreville). More issues are at stake than mere monetary gains in lexicography which anywhere is not known for big profits. The focus here on the monetary aspect is due to "the strategic planning for Gabonese lexicography" which intends to make lexicography an economic asset for Gabon (Emejulu 2000 and 2003).

The question to the planners of Gabonese lexicography is: Should the LPG prefer Eveia and ignore the language of the Kalaba population? If rigorous cri-

teria for selecting a language for the development of a Gabonese economically strategic lexicography have to be established, many native Gabonese languages will be disqualified due to their number of speakers, their impact on the national economy and their advance towards extinction. This is the case with Eveia.

On the contrary, language development is usually more sensitive to national than economic goals (Harris and Rampton 2003).

5.   Conclusion

This article has tried to raise a number of issues that planners of Gabonese lexicography may have to deal with. Possible directions to follow are indicated for each issue. For instance, lexicographic units should be established according to Kwenzi-Mikala's classification of Gabonese into certain language-units. This implies the issue of inventorying the Gabonese languages and the issue of their standardization as well as the issue of a language plan for Gabon.

The article also advocates the consideration of certain foreign languages widely spoken in Gabon, especially French accepted as sole official language of the country. A specific form of French that includes exclusively Gabonese words and expressions should result from Gabonese dictionary production.

Finally, the article discusses the concept of the Gabonese language landscape (GLL) which acknowledges both native and foreign languages for a better understanding and presentation of a realistic linguistic plan for Gabon.

Notes

1.   Ndinga-Koumba-Binza (forthcoming) gives more details on the Gabonese language policy and a number of suggestions for establishing a new language policy for Gabon.
2.   L. Mabika Mbokou, T. Afane Otsaga, G.-R. Mihindou and B. Nyangone Assam are among Gabonese lexicographers involved in this project of the *Dictionnaire du Français du Gabon*.

References

Dictionaries

Biton, A. 1907. *Dictionnaire français–ndumu et ndumu–français, précédé d'éléments de grammaire.* Nantes: Imprimerie Bourgeois.

Galley, Samuel. 1964. *Dictionnaire fang–français/français–fang.* Neuchatel: Henri Messeiller.

Marichelle, C. 1902. *Dictionnaire français–vili.* Loango: Imprimerie de la Mission.

Mikala Manfoumbi, Roger. 2004. *Lexique pove–français français–pove.* Libreville: Éditions Raponda-Walker.

Raponda-Walker, A. 1995². *Dictionnaire français–mpongwé.* Libreville: Éditions Raponda-Walker.

Van der Veen, Lolke J. and Sébastien Bodinga-bwa-Bodinga. 2002. *Gedandedi sa Geviya/Dictionnaire geviya–français.* Louvain/Paris/Sterling, Virginia: Peeters.

Other Literature

Afane Otsaga, Thierry. 2004. *The Standard Translation Dictionary as an Instrument in the Standardiza-tion of Fang*. Unpublished D.Litt. Dissertation. Stellenbosch: University of Stellenbosch.

Alberts, Mariëtta. 2004. A Few Words from AFRILEX. *Lexikos* 14: xiv-xv.

Blanchon, Jean A. 1994. Gabon: Language Situation. Asher, R.E. (Ed.). 1994. *The Encyclopedia of Lan-guage and Linguistics. Volume 3*. Oxford/New York: Pergamon Press.

Couvert, C. 1982. *Langue française en République gabonaise*. Paris: Haut Comité de la Langue Fran-çaise.

Crystal, David. 1999. *The Penguin Dictionary of Language*. Second Edition. Harmondsworth: Penguin Books.

Dodo-Bounguenza, E. 2002. Les langues du Gabon: le point en 2001. *Le Gabon et l'Unesco, Revue d'information* 2: 33-39.

Emejulu, James D. 2000. Lexicography, an Economic Asset in Multilingual Gabon. *Revue Gabonaise des Sciences du Langage/Gabonese Journal of Language Sciences* 1: 51-69.

Emejulu, James D. 2001a. Lexicographie multilingue et multisectorielle au Gabon: planification, stratégie et enjeux. Emejulu, J.D. (Ed.). 2001: 38-57.

Emejulu, James D. (Ed.). 2001b. *Éléments de lexicographie gabonaise. Tome 1*. New York: Jimacs-Hill-man Publishers.

Emejulu, James D. 2002a. Défis et promesses de la lexicographie intégrale dans les pays en déve-loppement. Emejulu, J.D. (Ed.). 2002: 366-381.

Emejulu, James D. (Ed.). 2002b. *Éléments de lexicographie gabonaise. Tome 2*. New York: Jimacs-Hill-man Publishers.

Emejulu, James D. 2003. Challenges and Promises of a Comprehensive Lexicography in the Devel-oping World: The Case of Gabon. Botha W.F. (Ed.). 2003. *'n Man wat beur. Huldingingsbundel vir Dirk van Schalkwyk*: 195-212. Stellenbosch: Bureau of the WAT.

Emejulu, James D. and Yolande Nzang-Bié. 1999. Linguistic Perspectives in Gabon. Unpublished paper presented at the International Meeting of the Summer Institute of Linguistics, Univer-sity of North Dakota, July 1999.

Gallardo, A. 1980. Dictionaries and the Standardization Process. Zgusta, L. (Ed.). 1980. *Theory and Method in Lexicography*: 59-69. Columbia: Hornbeam Press.

Gouws, Rufus H. 2001. Lexicographic Training: Approaches and Topics/Formation lexicographi-que: approches et thèmes. Emejulu, J.D. (Ed.). 2001: 58-94.

Grimes, Barbara (Ed.). 1996. *Ethnologue*. Thirteenth Edition. Dallas: Summer Institute of Linguis-tics, Inc.

Harris, Roxy and Ben Rampton (Eds.). 2003. *The Language, Ethnicity and Race Reader*. London/New York: Routledge.

Hubert, Jacques and Paul A. Mavoungou. (Eds.). Forthcoming. *Écriture et Standardisation des Lan-gues Gabonaises*. Submitted for publication to The Centre for Advanced Studies in African Society (CASAS) in Cape Town.

Idiata, Daniel F. 2002. *Il était une fois les langues du Gabon*. Libreville: Éditions Raponda-Walker.

Idiata, Daniel F. 2003. *Pourquoi le Gabon doit investir sur ses langues vernaculaires*. Cape Town: The Center for Advanced Studies of African Society (CASAS).

Jacquot, André. 1978. Le Gabon. Barreteau, D. (Ed.). 1978. *Inventaire des études linguistiques sur les langues d'Afrique noire d'expression française et sur Madagascar*. Paris: CILF.

Kwenzi-Mikala, Jerôme T. 1987. Contribution à l'inventaire des parlers bantu du Gabon. *Pholia* 2: 103-110.

Kwenzi-Mikala, Jerôme T. 1988. L'identification des unités-langues bantu gabonaises et leur classification interne. *Muntu* 8: 54-64.

Kwenzi-Mikala, Jerôme T. 1990. Quel avenir pour les langues gabonaises? *Revue Gabonaise des Sciences de l'Homme* 2: 121-124.

Kwenzi-Mikala, Jerôme T. 1998. Parlers du Gabon: classification du 11.12.97. Raponda-Walker, A. (Ed.). 1998. *Les langues du Gabon*. Libreville: Éditions Raponda Walker.

Mabika Mbokou, Ludwine. Forthcoming. The *Dictionnaire français–yipounou/yipounou–français* of the Église Évangélique du Sud Gabon (CMA) 1966. *Lexikos* 16.

Matthews, P.H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford/New York: Oxford University Press.

Mavoungou, Paul A. 2001a. La mondialisation et la lexicographie trilingue ou plurilingue au Gabon. J.D. Emejulu (Ed.). 2001: 160-183.

Mavoungou, Paul A. 2001b. Macro- and Microstructural Issues in Mazuna Lexicography. *Lexikos* 11: 122-138.

Mavoungou, Paul A. 2002a. *Metalexicographical Criteria for the Compilation of a Trilingual Dictionary: Yilumbu–English–French*. Unpublished D.Litt. Dissertation. Stellenbosch: University of Stellenbosch.

Mavoungou, Paul A. 2002b. Vers un dictionnaire du français du Gabon. Emejulu, J.D. (Ed.). 2002: 230-262.

Mavoungou, Paul A. 2004. Review of L.J. van der Veen and S. Bodinga-bwa-Bodinga: *Gedandi sa Geviya/Dictionnaire Geviya–Français*. *Lexikos* 14: 440-448.

Mba-Nkoghe, Jules. 1991. Le Gabon linguistique. *Notre Librairie* 105: 20-23.

Mihindou, Guy-Roger. 2001. Apports des missionnaires à la lexicographie gabonaise: dictionnaires bilingues fang–français/français–fang; français–yipounou/yipounou–français; français–mpongwé. Emejulu, J.D. (Ed.). 2001: 7-37.

Moulougui Nguimbyt, Franck V. 2002. *De la variation dialectale en français au Gabon*. Unpublished M.A. Thesis. Libreville: Omar Bongo University.

Moussirou Mouyama, Auguste. 1984. *La langue française au Gabon: Contribution sociolinguistique*. Thèse de Doctorat 3è cycle. Paris: Université René Descartes (Paris V).

Moussirou Mouyama, Auguste. 1986. L'introduction de la langue française au Gabon. *Annales de l'Université Omar Bongo du Gabon: Série Lettres, Droit, Sciences et Médecine*, 55-63.

Ndinga-Koumba-Binza, Hugues Steve. 2003. Gabonese Language Landscape (GLL). Unpublished paper presented at the 12th Biennial International Conference of the African Language Association of Southern Africa (ALASA), Stellenbosch, University of Stellenbosch, 7–11 July 2003.

Ndinga-Koumba-Binza, Hugues Steve. 2004. Le statut socio-politique du français et la promotion langues locales au Gabon. Paper presented at 18th International Conference of the Association of French Studies in Southern Africa (AFSSA), Stellenbosch, University of Stellenbosch, 8–11 September 2004.

Ndinga-Koumba-Binza, Hugues Steve. Forthcoming. Gabonese Language Landscape: Survey and Perspectives. *South African Journal of African Languages* 25(1).

Nyangone Assam, Blanche and Paul A. Mavoungou. 2000. Lexicography in Gabon: A Survey. *Lexikos* 10: 252-274.

Nzang-Bié, Yolande. 2001. Vers une éducation multilingue au Gabon: première approche. *Revue Gabonaise des Sciences du Langage/Gabonese Journal of Language Sciences* 2: 17-29.

Nzang-Bié, Yolande. 2002. Le corpus lexicographique dans les langues à tradition orale: le cas du dialecte fang-mekè. *Lexikos* 12: 111-138.

Ogden, John. 1984. French in Gabon. *Contemporary French Civilization* 8: 339-348.

Pambou, Jean-Aimé. 1998. Le français au Gabon: une langue à multiples statuts. *Iboogha* 2: 127-149.

Perrois, L., André Jacquot and Paulette Moussavou. 1983. Situation ethnolinguistique. *Géographie et cartographie du Gabon: Atlas illustré*. Paris: Edicef.

Raponda-Walker, André (Ed.). 1998. *Les langues du Gabon*. Libreville: Éditions Raponda Walker.

Tomba Moussavou, Fatima. 2001. *Les réalisations des séquences des consonnes en français de Libreville*. Unpublished M.A. Thesis. Libreville: Omar Bongo University.

Addendum 1:   Gabonese Language-units according to Kwenzi-Mikala (1987, 1988, 1990 and 1998): Inventory and Internal Classification

---

1 **MAZUNA:** Fang-Atsi, Fang-Make, Fang-Mvai, Fang-Ntumu, Fang-Nzam and Fang-Okak

2 **MYENE:** Enenga, Galwa, Mpongwe, Nkomi, Orungu and Adjumba

3 **MEKAN-MENAA:** Akele, Ungom, Lisigu, Mbangwe, Metombolo, Seki, Tumbidi, Shake, Wumpfu and Lendambomo

4 **MEKONA-MANGOTE:** Ikota, Benga, Shamayi, Mahongwe, Ndasa and Bakola

5 **MEMBE** (or **OKANDE-TSOGO**): Getsogo, Gepinzi, Kande, Gevove, Gehimbaka, Geviya, Ebongwe and Kota-Kota

6 **MERYE:** Gisira, Givarama, Givungu, Yipunu, Yilumbu, Yisangu, Ngubi, Civili, Yirimba and Yigama

7 **METYE:** Yinzebi, Yitsengi, Yiwele, Yivili, Liduma, Liwanzi and Yibongo

8 **MEMBERE:** Lembaama, Lekaningi, Lindumu, Latege and Latsitsege

9 **MAKENA:** Bekwil, Shiwa (or Makina) and Mwesa

10 **BAKA:** Baka

Addendum 2:   Foreign African Languages in Gabon

Country of Origin	Languages	Mostly spoken in
Angola	Ngola	Libreville
Benin	Adja, Fon, Nago, Watchi	Franceville, Lambaréné, Libreville, Mouila, Port-Gentil, Tchibanga
Cameroon	Aghemwun, Bafia, Bamileke, Bamun, Bassa, Batanga, Bete, Bulu, Duala, Eton, Fingwe, Hausa, Kako, Laka, Meka, Sango	Bitam, Lambaréné, Libreville, Minvoul, Oyem, Port-Gentil
Congo-Brazzaville	Kako, Kikongo, Laari, Mboshi	Franceville, Libreville, Mayumba, Ndéndé, Tchibanga
Democratic Republic of the Congo	Kisakata, Kiswahili, Lingala, Pende, Tetela, Zande	Libreville, Port-Gentil
Equatorial Guinea	Bubi	Bitam, Libreville, Oyem
Ghana	Asante, Banda, Ewe, Fingwe, Laka	Libreville, Port-Gentil
Guinea-Conakry	Halpular, Manding	Libreville, Port-Gentil
Ivory Coast	Baule, Ebrie, Nyagafulu	Libreville, Port-Gentil
Mali	Bambara, Malinke, Sarakole, Sonraï	Franceville, Lambaréné, Libreville, Mouila, Port-Gentil, Tchibanga
Nigeria	Buna, Hausa, Igbo, Laka, Yoruba	Libreville, Port-Gentil
Republic of Central Africa	Banda, Kako, Laka, Ngola, Sango	Libreville
Rwanda	Kinyarwanda	Libreville, Franceville
São Tomé and Príncipe	Creole	Libreville, Port-Gentil
Senegal	Halpular, Manding, Sossamahi, Wolof	Franceville, Lambaréné, Libreville, Mouila, Port-Gentil, Tchibanga
Togo	Fon, Kotokoli, Mina, Watchi	Libreville, Lambaréné, Port-Gentil

Addendum 3:   European Languages in Gabon

Country of Origin	Languages	Mostly spoken in
Argentina	Spanish	Libreville
Brazil	Portuguese	Libreville
Canada	English, French	Libreville
Equatorial Guinea	Spanish	Libreville, Oyem, Bitam
France	French	Libreville, Franceville, Port-Gentil
Ghana	English	Libreville, Port-Gentil
Germany	German	Libreville, Port-Gentil
Italy	Italian	Libreville
Netherlands	Dutch	Gamba, Port-Gentil
Nigeria	English	Libreville, Port-Gentil
São Tomé and Príncipe, Cape Verde, Angola	Portuguese	Libreville

Addendum 4:   Middle Eastern and Asian Languages in Gabon

Country of Origin	Languages	Mostly spoken in
China	Chinese	Libreville, Omboué
Japan	Japanese	Libreville
Lebanon	Arabic	Libreville, Port-Gentil, Franceville
Malaysia	Malaysian	Mayumba, Ndindi, Tchibanga, Lambaréné
Mauritania	Arabic	Libreville, Port-Gentil, Franceville, Lambaréné, Mouila, Oyem
Morocco	Arabic	Libreville, Port-Gentil

De la révision du Dictionnaire de Samuel Galley [*]

Pierre Ondo-Mebiame, *Département des Sciences du Langage et Groupe de Recherche en Langues et Cultures Orales (GRELACO), Université Omar Bongo, Libreville, République Gabonaise (pierreondo@yahoo.fr)*

Résumé: Le *Dictionnaire fang–français et français–fang* de Samuel Galley (1964) constitue une contribution importante dans l'histoire lexicographique du parler et, par sa nature biscopale, un intérêt pour les locuteurs du fang. Au demeurant, il présente un nombre important de lacunes qui méritent d'être corrigées. En l'occurrence, et entre autres, le lexicographe n'a pas mentionné les critères selon lesquels les items lexicaux ont été lemmatisés dans le dictionnaire. Par conséquent, on est forcé de deviner les sources du dictionnaire. Le traitement des mots composés et des homonymes doit également être amélioré. Ce sont ces améliorations que nous esquissons dans la présente proposition.

Mots clés: DICTIONNAIRE, LEXICOGRAPHIE, SOURCE, PRÉ-TEXTES, NOMENCLATURE, POSTTEXTES, SIGNE, LEMMES, BISCOPAL, MACROSTRUCTURE, MICROSTRUCTURE, ENTÊTE, PARAPHRASE, MOT COMPOSÉ, HOMONYMES

Abstract: About the Revision of Samuel Galley's Dictionary. Samuel Galley's *Dictionnaire fang–français et français–fang* (1964) forms an important contribution towards the lexicographic history of the language and, because of its biscopal nature, is of interest to speakers of Fang. However, it shows a number of important shortcomings, which should be corrected. Among others, the lexicographer does not mention the criteria according to which lexical items have been lemmatized in the dictionary. Therefore one is compelled to guess the sources of the dictionary. The treatment of compound words and homonyms needs also to be improved. These improvements are outlined in this contribution.

Keywords: DICTIONARY, LEXICOGRAPHY, SOURCE, FRONT MATTER TEXTS, CENTRAL LIST, BACK MATTER TEXTS, SIGN, LEMMATA, BISCOPAL, MACROSTRUCTURE, MICROSTRUCTURE, HEADWORDS, PARAPHRASE, COMPOUND WORD, HOMONYMS

1. Présentation

1.1 Généralités

Selon Mihindou (2001: 12) le texte de Galley intitulé *Dictionnaire fang–français et français–fang* est l'un des dictionnaires bilingues produit en fang du Gabon. Il

---

s'agit d'une œuvre produite à titre posthume, grâce au concours conjugué de l'Association Française des Amis d'Albert Schweitzer, de la veuve Galley, et du professeur Henry A.C. Samuel Galley, qui, pour mémoire, était un Mission-naire de l'Église protestante et travailla en pays fang pendant près de quarante ans.

Toujours selon Mihindou (2001: 12), le texte de Galley peut être divisé en trois parties:

(1)   les pré-textes;

(2)   la nomenclature; et

(3)   les posttextes.

Les pré-textes comprennent:

(a)   le titre de l'ouvrage, le nom de l'auteur, l'initiateur de la publication, l'éditeur et le lieu de l'édition;

(b)   la date de la fin d'impression, le nom et le lieu de l'imprimerie; et

(c)   la préface et l'avant-propos.

La nomenclature ou dictionnaire proprement dit observe l'ordre alphabétique, et comprend deux parties:

(a)   la présentation des lemmes en utilisant le fang comme idiome source; et

(b)   la présentation des lemmes en utilisant le français comme langue source.

Les posttextes enfin, sont un exposé sur la grammaire du fang rendu en sept chapitres, à savoir:

(a)   l'orthographe et la phonétique;

(b)   les classificateurs des noms;

(c)   les classes des noms;

(d)   les adjectifs;

(e)   les pronoms;

(f)   les verbes; et

(g)   les mots invariables.

En dehors de ses activités lexicographiques et évangéliques, le pasteur Samuel a également entrepris la production des ouvrages didactiques. Son outil didac-tique le plus connu est le livre *Nteni Osu Nteni Fañwe*[1] publié en 1889 sous l'é-gide de la station protestante de Talagouga.

1.2   Avantages

Le dictionnaire de Galley est prévu pour satisfaire à la fois les besoins des locuteurs du fang et du français en raison de sa nature biscopale (le terme est emprunté à Hausmann et Werner 1991: 2740) ou bi-directionnelle. Sa nomenclature présente une macrostructure dont les lemmes (ou unités du dictionnaire) sont rangés dans l'ordre alphabétique, alors que la microstructure fournit un nombre important de données, dont:

— l'indication des parties du discours;

— le numéro de la classe de nom;

— les traductions,

— la paraphrase du sens; et

— les renvois.

Comparons, pour illustrer, l'article du lemme **AKÔL**:

> **AKÔL** (h) n.4, pl. *mekôl*. Arbre dont les feuilles sont rigueuses et peuvent remplacer la toile émeri ou le papier de verre. Dans les écoles, les ardoises sont souvent remplacées par de petites planchettes en *ékukh;* les élèves écrivent dessus avec le crayon à mine de plomb et effacent bruyamment avec les feuilles d'*akôl. Akôl ntañan,* véritable papier de verre ou toile émeri.

Exemple textuel 1: Article **AKÔL** (extrait du Galley[2] 1964: 30)

Une autre caractéristique du dictionnaire est la structure d'accès. En effet, l'utilisation des marqueurs typographiques tels que la présentation des lemmes en gras et en lettres majuscules, ainsi que les **mots-guides** ou **entêtes** qui forment la structure externe d'accès rapide (Hausmann et Wiegand 1989: 329) facilitent aux utilisateurs, la quête du mot qu'ils recherchent.

Au-delà de ces avantages, notre dessein est d'inviter à une révision de l'importante proposition de Galley, en indiquant les aspects à améliorer. Nos suggestions porteront sur les plans macrostructurels et microstructurels.

2.   La macrostructure du dictionnaire de Galley

2.1   La lemmatisation des items

La macrostructure du *Dictionnaire fang–français et français–fang* contient ±13 925 articles présentés dans l'ordre alphabétique strict. C'est un recueil du parler fang tel qu'il est utilisé par la communauté de Talagouga[3].

En amont du dictionnaire, le lexicographe a pris en compte les mots utilisés dans le discours quotidien, les noms géographiques, les noms de plantes et d'animaux, les noms propres et les expressions idiomatiques en tant que

partie intégrante de ses éléments macrostructuraux. Le texte est donc une collection représentative des items lexicaux du fang en usage à Talagouga. Le lexicographe aura voulu par là, que son dictionnaire soit, selon l'expression de McArthur (1986) un "véritable réservoir de connaissances".

Au demeurant, ce dernier (en l'occurence l'éditeur) n'a pas mentionné les critères selon lesquels les items lexicaux ont été lemmatisés dans le dictionnaire. Ce genre d'informations est habituellement présenté dans le guide, à l'endroit des usagers, qui constitue les pré-textes du dictionnaire. L'on observe que la préface du dictionnaire fait simplement état des raisons pour lesquelles l'Association Française des Amis d'Albert Schweitzer a décidé de publier le travail du pasteur Genevois. La section introductive donne également un aperçu des activités de missionnaires, des lexicographes et des pédagogues de Galley. Rien n'est cependant mentionné au sujet du choix des lemmes.

L'on est donc, fort de cela, en droit de s'interroger sur:

—   Les sources primaires et secondaires du dictionnaire;

—   Les utilisateurs potentiels du dictionnaire; et

—   Les besoins et les aptitudes des utilisateurs-cibles du dictionnaire.

Comme le lecteur avisé pourrait s'en apercevoir, ces interrogations ne sont pas prises en charge par les pré-textes du dictionnaire; par conséquent, l'on est obligé de deviner les sources du dictionnaire.

2.2   Les sources primaires et secondaires du dictionnaire

Au sujet des sources primaires[4], le guide du dictionnaire nous laisse comprendre que le travail est basé sur la traduction de la Bible en fang que Galley a entreprise pendant 40 ans pour le compte de la communauté fang de Talagouga. Une mention est également faite à l'aide de l'interprète de Galley: Obam Mendôme.

Les sources secondaires[5] du dictionnaire ne sont pas explicitement discutées dans le guide aux usagers. La préface mentionne cependant l'existence d'un travail lexicographique antérieur avec des éléments macrostructuraux restreints. Il s'agit très probablement du *Lexique fãn–français* compilé par Martrou (1924).

On trouve une mention des sources secondaires en examinant minutieusement la nomenclature, et spécifiquement l'article **MVEP**. L'examen de ce lemme donne:

> **MVEP** (h) n.3, pl. *bemvep* Rongeur, appelé en anglais canerat, rat de canne à sucre ..., <u>Largeau</u> dit le même mot. ….

Exemple textuel 2: Article **MVEP** (extrait du Galley 1964: 212)

Le traitement de ce lemme-là produit donc la preuve que Galley a utilisé l'*Encyclopédie pahouine* de Largeau (1901) comme source secondaire, au cours de l'élaboration de son dictionnaire. Pour la gouverne, il est bon de savoir que l'ouvrage de Largeau porte sur les variétés fang-atsi et fang-make. On pourrait même aller plus loin et penser que Galley ne s'est pas uniquement limité à l'exploitation de l'encyclopédie de Largeau, mais a employé tous les ouvrages lexicographiques antérieurs portant sur le fang[6]. Ce soupçon est plus évident dans la section français–fang. Il semble en effet que les dictionnaires monolingues ou bilingues du français qui ont été publiés avant le texte de Galley aient constitué des sources secondaires de son ouvrage.

2.3   Les utilisateurs potentiels du dictionnaire et leurs besoins et aptitudes

La section introductive mentionne le fait que le dictionnaire ne vise pas un public sophistiqué en ce que l'auteur n'a pas proposé la transcription phonétique de ses unités de traitement. Mais il n'en demeure pas moins, à l'usage, que l'ouvrage soit également conçu pour satisfaire les besoins des étudiants, des chercheurs, et en même temps, des personnes qui souhaitent apprendre la langue.

André-Charles Henry, le préfacier de l'œuvre de Galley, déclare à ce propos que l'ouvrage de Galley n'est pas un livre pour les personnes savantes. En effet, la langue, ou mieux, le niveau de langue utilisé dans le dictionnaire est tout sauf hermétique. Les paraphrases du sens utilisées sont simples et concises.

Au demeurant, l'inclusion de données telles que le numéro de classe, la formation du mot, les marqueurs structuraux, l'étymologie, les synonymes, les homonymes et même les paraphrases du sens, présupposent une sérieuse connaissance linguistique de la part de l'utilisateur-cible, malgré l'insertion d'une mini-grammaire en posttexte. Cela fait donc penser que les lecteurs-cibles du dictionnaire pourraient d'être des chercheurs et des étudiants.

3.   La microstructure du dictionnaire de Galley

Nous allons traiter de la microstructure du dictionnaire de Galley en exposant les avantages et les inconvénients.

3.1   Avantages

(1) La microstructure du dictionnaire de Galley, et cela est de notoriété (Mihindou 2001) présente beaucoup d'informations. On relève en l'occurrence:

—   les parties du discours;

—   les numéros de classes des noms;

—   les traductions;

—   les explications de sens; et

—   les renvois.

(2) En outre, et cela n'a pas été suffisamment dit jusqu'ici, chaque signe-lemme est immédiatement et systématiquement suivi par une indication de tonalité, une entrée d'une grande importance pour l'utilisateur. Très souvent, les lexico-graphes produisant des dictionnaires dans des langues africaines ont toujours péché en ignorant cette composante de base de l'analyse phonologique de la langue.

Galley distingue trois niveaux de tonalité dans son dictionnaire:

(a)   la tonalité élevée (notée (h));

(b)   la tonalité basse (notée (b)); et

(c)   la tonalité moyenne (notée (m)).

(3) On observe qu'en traitant de la signification, l'auteur fait l'effort de faire apparaître des données ayant trait à des distinctions dans la signification. Il expose donc ce qui se rapporte aux valeurs polysémiques d'un item lexical (cf. Bureau du WAT 1999). Considérons pour cela le lemme suivant:

AKÔNA (bm) n.4, pl. *mekôna* (vbs *kon* h et *wôm* h). Amas, agglomération, quelque chose qui forme un bloc. *Akôna ndôa,* brasier out rouge. *Akôna ñkur,* amas de nuages. *Akôna mekî,* caillot de sang. *Akôna mezim,* bouillonnement d'eau dans les rapides. *Akôna kôñ, myaña,* lingot d'or, d'argent. *Akôna éki,* lingot de fer parfois donné parmi d'autres choses comme dot (syn.: *awur ékî*). *Akôna ékî* signifie aussi métal quelconque qui a été fondu. Voir *ñkôna.*

Exemple textuel 3: Article **AKÔNA** (extrait du Galley 1964: 30)

Les syntagmes *Akôna ndôa, Akôna ñkur, Akôna mekî, Akôna mezim, Akôna kôñ, myaña* et *Akôna éki* se rapportent aux expansions polysémiques de la significa-tion du lemme **AKÔNA** ('Amas, agglomération, quelque chose qui forme un bloc').

(4) Une autre caractéristique du dictionnaire se situe au niveau du traitement des signes-lemmes représentant les items lexicaux polysémiques. Parcourons le traitement du lemme **ABEÑY**.

ABEÑY (b) n.4, pl. *mebeñle*. 1. Pincettes de forgeron pour tenir le fer rouge. —2. Mirliton du chanteur de la danse *akôm*, …—3. Barre ou plaque de fer ou d'un autre métal. —4. Museau allongé du crocodile ou de la civette *mvakh*. —5. balle de fusil ou cartouche. —6. *Abeñy ndame*, chevalet à deux X en pente où l'on met en position penchée les lianes à caoutchouc coupées pour en recueillir le latex. —7. *Abeñy ntsap*, barre de savon. —8. *Abeñy ébom*, système de fer qui ferme la cangue du prisonnier. —9. *Abeñy e nžel*, grande barbe.

Exemple textuel 4: Article **ABEÑY** (extrait du Galley 1964: 16)

Dans l'exemple 4, le lexicographe fournit aux utilisateurs des données présentées dans deux sections de l'article:

(a)   le "commentaire sur la forme" (CF); et

(b)   le "commentaire sur le sens" (CS) (cf. Hausmann et Wiegand 1989: 353).

Les entrées telles que le signe-lemme (**ABEÑY**), la classe de tonalité (b), le numéro de classe (n.4) et l'indication du pluriel (*mebeñle*) qui apparaissent dans la structure gauche du noyau de l'article du dictionnaire, font partie du commentaire sur la forme ou CF; le reste des entrées dans la structure droite du noyau de l'article appartient au commentaire sur le sens CS (cf. Hausmann et Wiegand 1989: 353-354).

L'utilisateur peut considérer le lemme **ABEÑY** comme représentant un item lexical polysémique. Chaque sens est introduit par un indicateur ou marqueur structurel 1, 2 ... 9 qui fait partie de la structure interne d'accès rapide du dictionnaire (cf. Hausmann et Wiegand 1989: 354-356).

(5) Une autre caractéristique du dictionnaire porte sur le traitement des termes de spécialité et des concepts culturels. Comparez le traitement de l'article du lemme **AKUA**:

> **AKUA** (b) n.4, pl. *mekua*. Endroit où on fait fondre le minerai de fer. On fait un grand trou dans la terre, on y met le minerai et beaucoup de charbon de bois. Tous ceux qui ont des soufflets de forge indigènes les installent autour du trou. On allume le brasier, on souffle sans arrêt jusqu'à ce que la fusion s'opère. On se relaie, et les femmes apportent des vivres aux souffleurs.

Exemple textuel 5: Article **AKUA** (extrait du Galley 1964: 31)

Le traitement lexicographique commence avec une paraphrase de la signification du lemme **AKUA**. Ensuite, une explication détaillée est donnée à propos de l'extraction du minerai de fer, ses acteurs, les instruments et techniques utilisés pour la circonstance. Ces explications, qui portent principalement sur les caractéristiques extra-linguistiques, ne font évidemment pas partie de la paraphrase de la signification.

(6) Il est important d'observer comment le lexicographe traite les noms de plantes et ceux des animaux. Considérons les exemples suivants:

> **TÔM** (b) n.3, pl. *metôm*. Grand arbre parasol. Nom commercial d'après Meniaud: Tsoumbou. (Nom scientifique: Parkia Klainei Pierre). Couleur du bois: gris jaune. Arbre magnifique à voir. *Tôlba tôm* (vb *tole* h), se faire verser goutte à goutte une infusion de bois de *tôm* dans l'œil, comme épreuve pour savoir si l'adultère a été commis (ordalie).

Exemple textuel 6: Article **TÔM** (extrait du Galley 1964: 344)

> **MVON** (b) n.3, pl. *bemvon*. Gros singe à long poils noirs (fourrure), colobe (Colobus Satanas). …

Exemple textuel 7: Article **MVON** (extrait du Galley 1964: 214)

L'exemple 6 contient le nom scientifique, une description et les utilisations de l'arbre et l'exemple 7 le nom scientifique et une description du singe.

Le traitement des termes de spécialité a été fait d'une manière tout à fait satisfaisante en ce qui concerne le contexte culturel parce que l'utilisateur a accès aux données lexicographiques concernant la faune et la flore de l'habitat naturel des populations Fang. Excepté quelques exemples (qu'il n'est pas nécessaire de mentionner ici), l'utilisateur du dictionnaire a toujours accès à des entrées telles que le nom scientifique (de la plante ou de l'animal concerné), les utilisations de la plante ainsi que les sources consultées.

Par ailleurs, les feuilles, les racines, les écorces, les fruits, les graines et les ingrédients d'animaux sont énumérés avec leurs utilisations dans la médecine traditionnelle. Toutes ces données concernant le contexte culturel des lemmes traités sont importantes pour l'utilisateur qui est susceptible de trouver des plantes et des ingrédients d'animaux en vente sur les étalages des marchés.

Dans la section français–fang du texte, les éléments macrostructuraux appartiennent très probablement au vocabulaire quotidien du français. La microstructure fournit habituellement à l'utilisateur-cible les données suivantes: la traduction en fang, parfois précédée par un synonyme ou des synonymes du lemme en français, et des exemples (chaque exemple en français est immédiatement suivi de son correspondant en fang). Les articles varient en longueur. Une illustration extrême de ce point, apparaît dans le traitement de l'article du lemme **ARBRE** (presque trois colonnes):

> **ARBRE** bois, *éli* (h). Boîte en bois, *évora éli* (bh) …

Exemple textuel 8: Article **ARBRE** (extrait du Galley 1964: 419-420)

La longueur de cet article résulte de l'énumération d'environ 200 essences d'arbres de la forêt gabonaise ainsi que leurs utilisations et descriptions. En outre, le nom commercial de l'arbre apparaît avec son nom local. Le lexicographe a essayé de traiter les items lexicaux de spécialité équitablement en ce qui concerne les deux sections du dictionnaire.

3.2   Inconvénients

(1) Ce que l'on observe, c'est que l'indication de tonalité est habituellement suivie d'une paraphrase du sens du lemme dans la langue-cible (le français). Ceci amène de notre point de vue, une discrimination car, une paraphrase du sens dans la langue-source (le fang) aurait donné aux locuteurs fang plus d'aptitudes dans la maîtrise du parler. Il semble donc que le dictionnaire est plus

susceptible de satisfaire les besoins des locuteurs du français que ceux du fang. Comparons pour cela l'article du lemme **MVIA**:

> **MVIA** (h) n.2, pl. *mimvia* (vb via h). Femme acquise avec la dot de notre fille. En d'autre termes, femme achetée par le beau-père avec la dot reçu de son gendre. Et sa fille dont la dot a été prise pour cela, c'est sa *mbala* (voir ce mot).

Exemple textuel 9: Article **MVIA** (extrait du Galley 1964: 213)

La paraphrase de sens ci-dessus, n'indique réellement à l'utilisateur que ce que signifie le lemme **MVIA**, et cette information n'est nullement utile au locuteur-natif, car il en dispose déjà a priori. Ce qu'il rechercherait, c'est l'amélioration de ses aptitudes.

(2) Dans l'article d'un signe-lemme représentant un item lexical polysémique comme pour l'exemple 4 **ABEÑY**, les différentes distinctions dans la significa-tion ne devraient pas être présentées sur une base aléatoire, mais selon des critères fixes, qui devraient également être expliqués dans un des pré-textes. Or pour le lemme **ABEÑY**, on devine plutôt les principes qui gouvernent les dif-férentes distinctions.

Il semble en effet que le lexicographe a distingué entre les significations primaires et secondaires. Ou plus probablement, l'arrangement du lemme **ABEÑY** est fait selon le principe des distinctions conjonctives et disjonctives dans la signification. Le sens 1 du signe-lemme en l'occurrence semble être la valeur polysémique littérale de l'item lexical; il est donc placé en premier dans l'article. Les sens 2, 3, 4, 5, 6, 7, 8 et 9 semblent être des extensions métaphori-ques du premier sens. En outre, les sens 2, 3, 5, 6 et 8 ont des significations plus ou moins liées, parce qu'ils ont tous une relation avec l'item lexical renvoyant au *fer* ou à la *métallurgie*. Ainsi, ils auraient dû être arrangés par le lexicographe dans l'ordre mentionné ci-dessus.

Les sens 7 et 9, dans lesquels la signification figurative ne se réfère pas directement au premier sous-commentaire sur le sens, pourraient avoir été ordonnés plus loin dans la liste.

Le lexicographe aurait également pu utiliser un autre principe d'arrange-ment, à savoir: le principe de l'arrangement empirique. Dans le principe de l'ar-rangement empirique, les distinctions dans la signification sont déterminées sur la base du calcul de fréquence (cf. Bureau du WAT 1999: 17-20).

(3) L'article du lemme **ABEÑY** met en relief quelques aspects de la présentation des mots composés. *Abeñy ndame*, *Abeñy ntsap*, *Abeñy ébom* et *Abeñy e nẑel* sont des composés avec la structure: nom + (connectif) + nom. Tous ces mots com-posés sont orthographiés sans trait d'union.

À l'inverse, les lemmes **MÔR-A-MVAM**, **MVAKH-FÎN** et **AKAM-NYAR** sont présentés avec des traits d'union. Ceci est très déroutant puisque les lemmes en question sont des mots composés avec la structure, nom + (connec-

tif) + nom, comme c'est le cas avec les noms composés *Abeñy ndame*, *Abeñy ntsap*, etc. en l'exemple 4. Cette question aurait dû être réglée dans les prétextes.

Son corollaire est le fait que jusqu'à ce jour, il n'y a aucun système d'orthographe uniformisé pour le fang et même à l'échelle de l'ensemble des « langues du Gabon ». La proposition faite en 1999 sous la désignation d'*Orthographe des Langues Gabonaises* (1999) n'a pas, elle non plus malheureusement, réglé le problème de l'écriture des mots composés, ainsi que celle des possessifs, démonstratifs, et nombre d'autres catégories grammaticales.

(4) Le traitement des homonymes souffre, lui aussi, d'insuffisances. Considérons à ce sujet l'exemple ci-après:

ABI (b) (lg) n.4, pl. *mebi*. Cuisse …
ABI (h) (bf) n.4. ss pl. Mal, péché, …
ABI (h) (bf) adj. Sing. …Mauvais, …
ABI (h) (lg) n.4, pl. *mebi*. 1. Sein, …

Exemple textuel 10: Article **ABI** (h), **ABI** (h) et **ABI** (h) (extrait du
Galley 1964: 17)

Les trois dernières unités de traitement du paradigme mentionné ci-dessus, qui, selon le lexicographe, ont la même indication de tonalité (cf. (h)), ont reçu le statut de lemmes, comme homonymes. Le dictionnaire n'aide pas son utilisateur-cible à distinguer les trois membres du paradigme homonymique. L'auteur aurait pu les affecter de chiffres en exposant, ou alors, commencer par un indicateur structurel, reflétant la fréquence d'emploi comme cela est habituellement fait dans de tels cas. Par exemple, **ABI**[1]; **ABI**[2]; **ABI**[3] ou **1. ABI**; **2. ABI**; **3. ABI**. De nombreux exemples pourraient être appelés pour illustrer cette lacune systématique dans le dictionnaire.

(5) Les explications que fournit la paraphrase de signification mettent certes en exergue quelques aspects culturels de la vie des Fang et permettent à l'utilisateur de tirer beaucoup d'informations utiles de ces entrées, mais il faut constater qu'elles surchargent le texte.

L'une des manières, pour les lexicographes, de rendre compte des explications culturelles en évitant que trop de données encyclopédiques s'infiltrent dans la paraphrase de la signification consiste à concevoir des articles de synthèse; en l'occurrence, des articles présentant une vue d'ensemble d'un champ de connaissance.

(6) L'article portant sur les termes de la faune et de la flore présente une structure peu commune dans le dictionnaire, et mériterait ainsi quelques explications dans l'un des textes externes du dictionnaire.

4.   La structure d'accès du dictionnaire de Galley

Comme dans le développement qui précède, nous traitons de la structure d'accès du dictionnaire de Galley en exposant les avantages et les inconvénients.

4.1   Avantages

On distingue généralement deux types de structures d'accès:

(a)   structure d'accès externe; et

(b)   structure d'accès interne.

La structure d'accès externe guide l'usager à partir d'un point de départ externe jusqu'au signe lemme. Par exemple, le titre du dictionnaire, la table des matières et les entêtes ressortissent à la structure d'accès externe. Il est bon de faire observer que la présentation des entêtes dans le dictionnaire de Galley a grandement amélioré sa structure d'accès externe.

La structure d'accès interne, d'autre part, commence à partir du signe lemme et se poursuit à travers l'article du dictionnaire jusqu'à l'information recherchée. Pour s'assurer que l'usager aura un accès rapide et facile aux données présentées, les lexicographes utilisent généralement des marqueurs structurels typographiques et non-typographiques. Pour ce qui est des indicateurs structurels typographiques, dans le dictionnaire de Galley, les lemmes apparaissent en gras et en « petites majuscules ». Les significations des lemmes ainsi que les traductions françaises des phrases en fang sont présentées en romain. Par contre, les items lexicaux, les syntagmes et les phrases en fang sont présentés en italique. L'utilisation de tous ces indicateurs structurels typographiques facilite aux usagers l'accès aux données recherchées.

4.2   Inconvénients

(1) Les marqueurs structurels non-typographiques incluent un certain nombre de symboles, dont, et entre autres: les diamants noirs, le triangle noir à l'envers, les rectangles, souvent utilisés pour marquer les limites d'une zone de recherche spécifique (définition, exemple, catégorie grammaticale …). Le dictionnaire de Galley présente à ce sujet peu de marqueurs structurels non-typographiques, alors que ceux-ci facilitent généralement aux usagers l'accès aux données recherchées.

(2) Le texte souffre aussi au niveau du traitement des éléments de la littérature sentencieuse (cf. proverbes, maximes, adages, …). Les proverbes par exemple sont traités dans la nomenclature, et notamment dans l'article du lemme introduisant ou figurant dans le proverbe traité. Mavoungou (2002: 230-237) nous montre au contraire que les proverbes et les autres expressions idiomatiques

n'appartiennent pas, à proprement parler, à la nomenclature; pour lui, ils doivent être traités dans les textes externes du dictionnaire et de préférence dans les posttextes. Un tel traitement permet de créer une autre possibilité d'accès pour l'utilisateur, et donne au dictionnaire un caractère polyaccessible.

5.   Conclusion

À un degré important, les pré-textes du dictionnaire fournissent assez d'informations concernant les sources primaires de l'ouvrage. Cependant, l'auteur ne donne pas une indication de ses sources secondaires. Il n'y a aucun doute que le lexicographe a employé un corpus lexicographique volumineux de la langue, ce qui fait du dictionnaire une source valable de référence et un véritable réservoir de connaissances. Les éléments macrostructuraux ont été sélectionnés parmi divers champs de connaissance (la tradition orale, la religion, la sylviculture, la pharmacopée traditionnelle, entre d'autres). Malheureusement, le lexicographe ne donne pas les critères qui ont présidé à l'incorporation des items lexicaux. En outre, la macrostructure ne présente pas de caractéristiques de sophistication, c'est-à-dire, l'utilisation des procédures de « niches » et de « nids ». La microstructure contient beaucoup d'entrées utiles, mais aucune discussion métalexicographique n'est fournie dans le guide aux usagers pour rendre l'ouvrage accessible au lecteur-cible.

Notes

1.   Opuscule centenaire de 14 pages "premier livre de fang " dit le titre. C'est la mission protestante de Talagouga (près de Ndjolé) qui est à l'origine de cette publication dont la reliure est assurée par du fil à aiguille. Etrange illustration de couverture: minarets et chéchias composent un paysage sans doute exotique mais fantastiquement irréaliste (Mayer 1990: 74).
2.   Autre désignation du *Dictionnaire fang–français, français–fang* de Samuel Galley.
3.   Localité située dans les environs de Ndjolé.
4.   Encore appelées corpus lexicographique ou base du dictionnaire, qui incluent toutes les sources qui ne sont pas des dictionnaires eux-même (cf. Smit 1996: 30-31 et 111).
5.   Tous les dictionnaires consultés durant la phase de rédaction (cf. Wiegand et Kučera 1981: 100ff).
6.   Marling (1872), Lejeune (1892) et Martrou (1924).

Références bibliographiques

Bureau du WAT. 1999. *Study Guide of the Woordeboek van die Afrikaanse Taal* (WAT). Notes de cours non-publiées. Stellenbosch: Bureau du WAT.
Galley, S. 1889. *Nteni Osi Nteni Fañwe*. Talagouga: Imprimerie de la Mission.
Galley, S. 1964. *Dictionnaire fang–français et français–fang, suivi d'une grammaire fang*. Neuchâtel: Henri Messeiller.

Hausmann, F.J. et R.O. Werner. 1991. Spezifische Bauteile und Strukturen zweisprachiger Wörter-
bücher: eine Übersicht. Hausmann, F. J. et al. (Éds.). 1989–1991: 2729-2769. Berlin: Walter de
Gruyter.

Hausmann, F.J. et H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual
Dictionaries: A Survey. Hausmann, F.J. et al. (Éds.). 1989–1991: 328-360. Berlin: Walter de
Gruyter.

Hausmann, F.J., O. Reichmann, H.E. Wiegand et L. Zgusta (Éds.). 1989–1991. *Wörterbücher. Ein
internationales Handbüch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexico-
graphy/Dictionnaires. Encyclopédie internationale de lexicographie.* Handbücher zur Sprach- und
Kommunikationswissenschaft 5.1–5.3. Berlin: Walter de Gruyter.

Largeau, V. 1901. *Encyclopédie pahouine. Eléments de grammaire et dictionnaire français–pahouin.* Paris:
E. Leroux.

Lejeune, L. 1892. *Dictionnaire français–fang ou pahouin, précédé de quelques principes grammaticaux sur
cette même langue.* Paris: A. Faivre et H. Teillard.

Marling. 1872. *Dictionnaire fang–français.* New York.

Martrou, L. 1924. *Lexique fãn–français.* Procure Générale (des Pères du St Esprit) Abbeville. Paris:
Imprimerie Paillard.

Mavoungou, P.A. 2001. Macro- and Microstructural Issues in Mazuna Lexicography. *Lexikos* 11:
122-138.

Mavoungou, P.A. 2002. *Metalexicographical Criteria for the Compilation of a Trilingual Dictionary:
Yilumbu–English–French.* Thèse de doctorat non-publiée. Stellenbosch: Université de Stellen-
bosch.

Mayer, R. 1990. Histoire de l'écriture des langues du Gabon. *Revue Gabonaise des Sciences de
l'Homme* 2: 65-91.

McArthur, T. 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablets to
the Computer.* Cambridge: Cambridge University Press.

Mihindou, G.-R. 2001. Apports des missionnaires à la lexicographie gabonaise: dictionnaires bi-
lingues fang–français/français–fang; français–yipounou/yipounou–français; français/mpongwé.
Emejulu J.D. (Éd.). 2001. *Éléments de lexicographie gabonaise. Tome I:* 7-37. New York: Jimacs-
Hillman Publishers.

Smit, M. 1996. *Wiegand's Metalexicography as a Framework for a Multilingual, Multicultural, Explana-
tory Music Dictionary for South Africa.* Thèse de doctorat non-publiée. Stellenbosch: Université
de Stellenbosch.

Wiegand, H.E. et A. Kučera. 1981. Brockhaus-Wahrig: Deutsches Wörterbuch auf dem Prüfstand
der praktischen Lexikologie. I. Teil: 1. Band (A–BT); 2. Band (BU–FZ). *Kopenhagener Beiträge
zur germanistischen Linguistik* 18: 94-217.

Nuwe Woordeboek sonder Grense: A Typological and Communicative Bridge

Mariza Steyn *(mariza@mml.co.za)*, and
Liezl Gouws *(13112635@sun.ac.za)*, *Department of Afrikaans and Dutch,*
University of Stellenbosch, Stellenbosch, Republic of South Africa

Abstract: Within the South African lexicographical context, there have been several pleas for Afrikaans learner's and school dictionaries that incorporate innovative strategies and provide for specific identified target users and their particular problems. *Nuwe Woordeboek sonder Grense* (NWSG) is an Afrikaans learner's dictionary that is aimed at a specific group of users, can be used for text reception and text production and forms part of an established textbook series. Through new strategies and adaptations, this dictionary tries to form a bridge between different typological categories and to be a functional instrument for use in the classroom. In accordance with the theoretical formulation of dictionary functions, NWSG tries to further communication through initial support in and eventual assimilation of the foreign language. In this article, theoretical insights regarding learner's dictionaries as well as practical examples from NSWG are combined to illustrate the possibilities of a more user-friendly product.

Keywords: LEARNER'S DICTIONARY, USERS, TYPOLOGY, BRIDGE LANGUAGE, TEXT RECEPTION, MACROSTRUCTURE, TEXT PRODUCTION, MICRO-ARCHITECTURE, FRAME STRUCTURE

Opsomming: *Nuwe Woordeboek sonder Grense:* **'n Tipologiese en kommunikatiewe brug.** Binne die Suid-Afrikaanse leksikografiese konteks is daar reeds verskeie pleidooie gerig vir Afrikaanse aanleerder- en skoolwoordeboeke wat vernuwende strategieë inkorporeer en vir spesifieke, geïdentifiseerde teikengebruikers en hul eiesoortige probleme voorsiening maak. *Nuwe Woordeboek sonder Grense* (NWSG) is 'n Afrikaanse aanleerderwoordeboek wat op 'n spesifieke groep leerders gerig is, vir teksresepsie en teksproduksie gebruik kan word en deel van 'n gevestigde handboekreeks uitmaak. Dié woordeboek probeer om deur nuwe strategieë en aanpassings 'n brug te vorm tussen verskillende tipologiese kategorieë en 'n funksionele gebruiksinstrument in die klaskamer te wees. In ooreenstemming met die teoretiese formulering van woordeboekfunksies, probeer NWSG om kommunikasie te bevorder deur aanvanklike ondersteuning in en uiteindelike assimilering van die vreemde taal. In hierdie artikel word teoretiese insigte in verband met aanleerderwoordeboeke en praktiese voorbeelde uit NWSG gekombineer om die moontlikhede van 'n meer gebruikersvriendelike produk te illustreer.

Sleutelwoorde: AANLEERDERWOORDEBOEK, GEBRUIKERS, TIPOLOGIE, BRUGTAAL, TEKSRESEPSIE, MAKROSTRUKTUUR, TEKSPRODUKSIE, MIKRO-ARGITEKTUUR, RAAM-STRUKTUUR

1.   Introduction

Research in pedagogical lexicography has gained momentum over the past fifty years. Recently, theoretical issues such as the influence of language learning and language acquisition theories, as well as the incorporation of the mother tongue into learner's dictionaries have come under the spotlight. Dictionary practice has centred on advanced learner's dictionaries within the British lexicographical tradition. These dictionaries have been praised for their technological advances, creative presentation and structural innovations. A good deal of criticism has also been expressed about the complexity of the presentation and the level of reference skills expected from the users of these dictionaries. It is against this theoretical and practical background that planning for a new learner's dictionary began.

2.   Background

2.1   Learner's dictionaries up to now

There are two important theoretical issues within pedagogical lexicography especially relevant to this article. Firstly, present-day learner's dictionaries focus mostly on the advanced learner at the expense of learners at pre-high-school levels. Kernerman (2000: 829) predicts the following: "This will give rise to dictionary research for beginners and intermediates, and a new generation of English learners' dictionaries designed specifically for lower levels." Secondly, lexicographers experience problems with the typological classification of learner's dictionaries on account of the insertion of the learners' mother tongue. An example of the typological confusion is the switching between terms like "bilingualised" and "semi-bilingual" for hybrid learner's dictionaries.

2.2   Learner's dictionaries in South Africa

The same problems can also be identified within the South African lexicographical context. No provision has been made for beginners and intermediate learners within the existing collection of learner's dictionaries. This gave rise to a new learner's dictionary, *Nuwe Woordeboek sonder Grense,* for learners from grades four to twelve with Afrikaans as a second, third or even fourth language. This dictionary aims at assisting these learners in everyday communication and usage in the classroom. The special typological classification of the dictionary will be discussed in the following section.

3.   *Nuwe Woordeboek sonder Grense*

3.1   Users and typology

Nuwe Woordeboek sonder Grense (NWSG) is a dictionary aimed at learners who

have neither English nor Afrikaans as mother tongue. Strictly speaking, NWSG cannot be called a learner's dictionary, since it is not aimed at adult learners, but at learners in the intermediate, senior and FET phases, i.e. grades 4-12. However, it is primarily a learner's dictionary, since various characteristics of learner's dictionaries have been incorporated and the dictionary can be used in the mastering of Afrikaans. The target users are African language speakers and English speakers who fall within the beginner and sometimes the intermediate phases of language acquisition. For these learners, a learner's dictionary taking typical mistakes into account and explicitly pointing them out to users can be very helpful.

In an attempt to facilitate the learning process and to support learners, a so-called "bridge language", namely English, is used to help learners find the correct word in the target language. Translation equivalents of the lemma are added in the example sentences. Consequently the dictionary functions as a language bridge. Text reception is also facilitated and accelerated. NWSG is therefore a hybrid monolingual dictionary.

boos   adjektief ■
(bose; boser, die boosste)
❶ Wanneer ek baie kwaad is, is ek boos.
⇨ *My pa was boos [angry/cross] toe ons die venster gebreek het.*

   ⃠ sinoniem **kwaad**

❷ Iemand of iets wat boos is, is baie sleg. ⇨ *Daardie man is boos [evil/ malicious], want hy slaan sy vrou en sy kinders.* ⇨ *Dit is 'n bose [evil] wêreld waar niemand meer vir die ander een omgee nie.*

Figure 1: Translation equivalents in example sentences

Learners using monolingual dictionaries often experience the following problem, as formulated by Atkins (1985: 21): "Users of a monolingual L2 dictionary can access the material in it only by means of a foreign language headword. It might be just that word that they do not know." In order to prevent this from happening, NWSG includes an equivalent register as outer text, thus helping learners to find the lemma via an English equivalent. This outer text functions as a communicative bridge whereby learners are referred from a foreign language, English, to the object language, another foreign language, Afrikaans. According to the typological classification of Tarp (2004: 323), NWSG is a monolingual dictionary with a bilingual feature. This dictionary therefore bridges the boundary between different dictionary classes, because it is primarily monolingual with one bilingual feature, namely translation equivalents.

Lys van Engelse woorde wat in die Afrikaanse voorbeeldsinne gebruik word om die trefwoord te vertaal

*Hierdie lys bevat sommige van die Engelse woorde wat in die Afrikaanse voorbeeldsinne in die alfabetiese gedeelte van die woordeboek gegee word. In hierdie lys word die **Engelse woorde alfabeties** gegee en daarna die Afrikaanse woord of woorde waar 'n mens in die hooflys van die woordeboek die Engelse woord as vertaling in 'n voorbeeldsin kry. Dit is belangrik om te weet dat hierdie Afrikaanse woorde nie altyd 'n vertaling van die Engelse woord is nie, maar net wys op die plek in die woordeboek waar die Engelse woord as vertaling in 'n voorbeeldsin gegee word.*

a; 'n
abbreviation; afkorting
about; daaraan, daaroor, omtrent, ongeveer, so, waaraan, waaroor
above; bo, bokant, oor

above-mentioned; bogenoemde
absolutely; volkome
accept; aanvaar
accident; ongeluk
according to; volgens
accordingly; gevolglik
account; rekening
accustomed; gewoond
ache; seer, pyn
achieve; bereik
acknowledge; erken
acquaintance; kennis
across; oorkant, skuins
act; daad, handel[2], optree, speel
acting; spel[1]
action; aksie, daad, handeling, optrede
active; aktief
actor; akteur
actress; aktrise
actually; eintlik, sowaar, waarlik
activity; aktiwiteit
add; daarby, optel
address; adres
admit; toelaat
adult; grootmens
adverb; bywoord
advertise; adverteer
advertisement; advertensie
advice; raad[1]

Figure 2: Equivalent register

3.2 Why NWSG is different

There are three reasons why NWSG differs from most other learner's dictionaries.

3.2.1 Part of a textbook series

Firstly, NWSG forms part of an established Afrikaans textbook series, *Nuwe Afrikaans sonder Grense*, aimed at school learners from grades 3 to 12. The dictionary therefore agrees with the series in title and look. Otto (1989: 24) emphasizes that the ideal for a learner's dictionary would be when textbook writers and lexicographers could confer and write two books with cross references to one another. During the compilation of NWSG, there was a close cooperation between the writers of *Nuwe Afrikaans sonder Grense* and the lexicographers. While the textbook writers made recommendations for the dictionary, the lexicographers in turn gave advice on the writing of the dictionary workbooks which accompany NWSG.

3.2.2 The lemma selection

The target users play an important role in the difficult task of selecting the lemma-sign list: "What words to put in, and what to leave out," as Bakker (1999: 294) puts it. "Indeed, how does one begin to assess what words South African school children currently live with?" The fact that NWSG is linked to the textbook series definitely influenced and helped the lexicographers in the lemma selection.

A major lexicographical challenge is compiling "a truly 'tailored' dictionary for a specific target user group" (De Schryver and Prinsloo 2003: 31). According to them, a delicate balance should be struck between a general-language corpus and a custom-made corpus consisting of material culled from the immediate environment of the envisaged target user group.

The target set by the publisher was a lemma-sign list of 2 500. In compiling the corpus, a selection of lists and corpuses were used: a manual list of roughly 5 000 lemma-signs compiled by the publisher (De Schryver and Prinsloo 2003: 35), the *Pretoria Afrikaans frequency Corpus* (PAfC) (De Schryver and Prinsloo 2003: 36), the lemma list of *Junior Verklarende Woordeboek,* as well as a corpus containing words found in the textbooks. The macrostructure of the dictionary therefore includes words that learners will come across in the classroom and in typical conversations. The collection of lexical items is therefore representative of the language generally used by the target users. Once again the interaction between the selection of textbooks and the dictionary insured that NWSG is a functional and relevant instrument for its users. The detailed process of selecting the lemma-sign list is described in an article by De Schryver and Prinsloo (2003).

3.2.3 The user profile

The third advantage of the dictionary being linked to the textbook series is that the lexicographers had extensive knowledge of the target users. The primary as well as secondary characteristics of the target users, namely their needs, skills and the situations wherein the dictionary will be used, were known to the lexi-

cographers. This knowledge made the task of aiming the content and presentation at specific target users much easier. The greatest part of this article will deal with aspects of text reception and text production in NWSG.

3.3   Text reception in NWSG

When it comes to text reception, the macrostructure and the form of the lemma-item play a very important role in helping the users.

3.3.1   The presentation of irregular flexion forms

In a dictionary with a text reception function, it is very important that users can find the specific word they are looking for. Second and third language speakers of Afrikaans often have problems with irregular flexion words such as **is/was**, **hê/het**, **sal/sou**, **wil/wou** etc. In NWSG, an attempt was made to insert and present these words in a consistent way. A user who, for example, finds **sou** in a text and does not know that it is the past tense of **sal**, will find a reference at **sou** to the reference address **sal**.

sal	hulpwerkwoord ■

(sou)

❶ 'n Mens gebruik sal as jy na iets in die toekoms wil verwys. ⇨ *My pa sal **[will]** môre die gras sny.* ⇨ *Ek belowe ek sal **[shall]** volgende week by jou kom kuier.*
❷ 'n Mens gebruik sal as jy wil beklemtoon dat iemand iets moet doen. ⇨ *Jy sal **[shall]** al jou groente eet, voor jy gaan slaap.* ⇨ *As jy nie al die werk doen nie, sal **[shall]** jy in die moeilikheid beland.*
❸ 'n Mens gebruik sou as jy 'n begeerte wil uitdruk. ⇨ *Ek sou **[would have]** daardie fiets gekoop het as ek genoeg geld gehad het.* ⇨ *Ons sou **[would have]** buite gespeel het as dit nie gereën het nie.*

sou	werkwoord ■

kyk **sal**

⇨ *Hy sou **[would]** kom, maar kan ongelukkig nie meer nie.* ⇨ *Christine het gesê sy sou **[would]** my bel.*

Figure 3: Irregular flexion words

Learners also have problems finding nouns and adjectives with "strange" flexion forms. Examples of such words are **skepe** (skip), **stede** (stad), **oë** (oog), **ore** (oor), **hoë** (hoog), **nuwe** (nuut), etc. These words also get lemma status in NWSG and refer the user to the articles where they are explained.

3.3.2   The macrostructure

Certain strategies have been adopted to make the macrostructure as accessible as possible in order to facilitate and precipitate the dictionary consultation pro-

cedure. In NWSG, the main lemmas have been ordered vertically and no sub-lemmas have been added. All homonyms have been listed as separate lemmas and the lexicographical convention of marking them with superscript markers has been followed.

3.4   Text production in NWSG

NWSG can only function as a communicative instrument if attention is given to the types and presentation of data which make text production possible. One of the biggest problems concerning learners using the dictionary comes to the fore when they are not familiar with the Afrikaans words they need. Here the equivalent register gives access to words included in the macrostructure of the dictionary.

It is essential that a learner using a learner's dictionary should get confirmation that he/she has reached the correct word. To achieve this, data about parts of speech, pronunciation and meaning should be included. According to Tarp and Gouws (2004: 289), data is necessary to distinguish between the different parts of speech of a multifunctional lemma. Multifunctional lemmas in NWSG are highlighted as units to indicate to learners that one lemma has more than one function. Ordering within these grey blocks is done according to a consistent system, namely nouns, verbs, adjectives, adverbs, prepositions etc. The names of the parts of speech are not abbreviated, but given in full, because the lexicographers cannot rely on the intuition of the users.

> **blink**   werkwoord ■
> **(het geblink)**
> Iets blink wanneer dit helder skyn of
> glinster. ⇨ *Die spieëls in my badkamer*
> *blink [shine] omdat dit skoon is. ⇨ My pa*
> *se nuwe, goue horlosie blink [shines] as die*
> *son daarop skyn.*
>   adjektief ■
> **(blink; blinker, die blinkste)**
> Wanneer iets blink is, is dit glinsterend
> en helder. ⇨ *Ek het op my eerste skooldag*
> *met nuwe, blink [shiny] skoene skool toe*
> *gegaan. ⇨ Daar word gereeld na diamante*
> *verwys as blink [shiny] klippies.*

Figure 4: Highlighted multifunctional lemmas

3.4.1   Indication of stress

In NWSG, data about pronunciation is limited to the indication of stress, assisting users in the production of spoken Afrikaans.

boodskap   naamwoord ∎
(boodskappe, boodskappie)
Wanneer jy vir iemand wat nie op die
oomblik beskikbaar is nie, iets wil sê, los
of stuur jy 'n boodskap. ⇨ *"Marie, ek het*
*'n boodskap **[message]** van jou ma. Jy moet*
na skool vir haar wag." ⇨ As my vriendin
se selfoon afgeskakel is, los ek altyd 'n
*boodskap **[message]**.*

Figure 5: Indication of stress

Furthermore, typical pronunciation problems of second and third language speakers are addressed in special usage notes. These notes are especially useful during the process of producing Afrikaans. They are presented in text blocks and marked with a bee or a tortoise. Thereby the lexicographers ensure easy access to this special data category.

bom   naamwoord ∎
(bomme, bommetjie)
'n Bom is 'n voorwerp wat ontwerp
word om baie skade aan te rig. ⇨ *Die*
*polisie het gelukkig die bom **[bomb]** gevind*
voor dit ontplof het. ⇨ Tydens die Tweede
Wêreldoorlog het baie mense gesterf weens
*die bomme **[bombs]**.*

boom   naamwoord ∎
(bome, boompie)
'n Boom is 'n groot plant met 'n dik
stam, takke en blare. ⇨ *My pa het my*
gehelp om 'n boomhuis in die groot boom
***[tree]** in ons agterplaas te bou.*

kyk by **bom**

Pasop vir **bom** en **boom**.
Bom word met 'n kort o-
klank uitgespreek en het het
'n dubbele-m in die meervoud, bomme.
⇨ *Daar het gister een bom **[bomb]** in die*
stad ontplof. Vandag het daar nog twee
*bomme **[bombs]** ontplof.* Boom word met
'n lang oo-klank uitgepreek en het net
een m in die meervoud, bome. ⇨ *Daar*
*groei een boom **[tree]** in ons tuin. Die bure*
*het drie bome **[trees]** in hulle tuin.*

Figure 6: Usage note addressing pronunciation problem

3.4.2   Word formation

A learner's dictionary with a text production function should include data about word formation. NWSG, in the same manner as other Afrikaans learner's dictionaries, indicates the morphology of words in an explicit, i.e. an uncondensed way. The user does not have to deduce anything concerning word for-

mation. For example, plural and diminutive forms are given in the case of nouns. In the case of adjectives, degrees of comparison are also given in the full form and "die" is added before the superlative form to help the learner with the production of Afrikaans.

boomhuis   naamwoord ■
(boomhuise, boomhuisie)
Dit is 'n huisie wat gewoonlik met houtplanke of ander materiaal bo in 'n boom gebou word. ⇨ *Ek en my maats speel elke naweek in ons boomhuis [tree house].*

aaklig   adjektief ■
(aaklige; aakliger, die aakligste)
Wanneer iets aaklig is, is dit sleg of lelik.
⇨ *My broer lyk aaklig [terrible] met die sny aan sy voorkop.* ⇨ *My beursie is gister gesteel. Dit was 'n aaklige [horrible/awful] ervaring.* ⇨ *Sommige mense dink rook is 'n aaklige [nasty] gewoonte.*

Figure 7: Word formation

3.4.3 Definitions

Semantic commentary in the form of definitions is the most important data type most often needed by learners. The lexicographer has the task of formulating complex definitions in a restricted vocabulary. Furthermore, every word used in the definition should be explained elsewhere in the dictionary. All the definitions in the dictionary are full sentences. The learner is confronted with natural language and the language context within which the word functions.

bord   naamwoord ■
(borde, bordjie)
❶ 'n Bord is 'n plat voorwerp waaruit 'n mens eet. ⇨ *As ek die tafel dek, moet ek 'n mes en vurk langs elke bord [plate] neersit.* ⇨ *My ma sorg altyd dat daar groente op my bord [plate] is sodat ek gesond kan eet.*
❷ 'n Bord is 'n plat voorwerp waarop 'n mens iets kan skryf of verf. ⇨ *Die onderwyser skryf al die huiswerk op die bord [board].* ⇨ *Langs die hoofpad is daar 'n bord [board/sign] wat sê hoe ver dit tot by die volgende dorp is.*
❸ 'n Bord is ook 'n plat voorwerp waarop speletjies gespeel word. ⇨ *'n Mens het 'n bord [board] met swart en wit blokke nodig om skaak te speel.* ⇨ *As 'n mens "Monopoly" speel, begin jy aan die een kant van die bord [board] en beweeg tot by die ander kant.*

Figure 8: Definitions

Words presenting similar concepts, for instance colours, are defined in a similar way. The systematic presentation informs the learner that a word belongs to a specific semantic field.

3.4.4   Semantic relations

The identification and learning of words with semantic relations form an essential part of language acquisition. Learners should therefore be made aware of the relations between words. These relations should be marked explicitly. The indication of semantic relations in NWSG is done in a systematic and consistent way. Synonyms and antonyms are presented and marked by typographical markers. These word pairs are also defined and presented in a similar way.

breed   adjektief ■
(breë; breër, die breedste)
Iets wat breed is, is in wydte, dit wil sê
van een kant tot die ander kant, baie
groot. ⇨ *Daar is breë [broad] geel strepe*
op
my nuwe hemp. ⇨ *Ons moes oor die breë*
[wide] rivier swem om aan die ander kant te
kom. ⇨ *Die meisie het 'n pragtige, breë*
[broad] glimlag.
◑ teenoor **smal**

smal   adjektief ■
(smal; smaller, die smalste)
As iets smal is, is daar net 'n kort afstand
van die een kant na die ander kant.
⇨ *Daar kan net een motor op 'n slag oor*
die smal [narrow] brug beweeg. ⇨ *Die pad*
is te smal [narrow] vir voetgangers,
fietsryers en motors.
◑ teenoor **breed/wyd**

Figure 9: Semantic relations

3.4.5   Example material

A crucial element of learner's dictionaries designed for text production is the inclusion and presentation of example material. Definitions must be supported by example sentences. According to Tarp and Gouws (2004: 292), learners must be able to see how words function in a typical usage situation. In order to achieve a successful transfer of knowledge the content and presentation of example material must be user-friendly and accessible. The lexicographers of NWSG used authentic examples to explicate the usage and contexts of the lemma.

<u>aan</u>dag   naamwoord ■
(-, -)
Om aandag te gee aan iemand of iets,
beteken om te luister of daaraan te dink.
⇨ *Ek moet aandag [attention] gee wanneer*
die onderwyser iets verduidelik. ⇨ *Ek het*
baie aandag [attention] aan my projek
gegee.

Figure 10: Example material

An integrated microstructure, where example material is placed directly after the definitions, is applied in NWSG. In the case of a polysemous lemma, example material has a direct bearing on the specific polyseme.

In the case of verbs, with specific reference to particle verbs, example sentences were chosen to illustrate the different usages of the verb.

bydra   werkwoord ■
(het bygedra)
As ek iets bydra, dan skenk ek iets of gee op een of ander manier hulp. ⇨ *Ek sal enige tyd iets bydra [contribute], solank dit vir 'n goeie saak is.* ⇨ *Ek het R400 bygedra [contributed] tot die kinderhuis.* ⇨ *Die munisipaliteit dra elke jaar by [contributes] tot die opvoeding van baie kinders.*

Figure 11: Particle verbs

Once again text production is encouraged through access to the different functions of the verb.

3.4.6   Illustrations

Quite a few definitions and example sentences are supported by illustrations. An illustration enhances the direct transfer of information because of its explicitness. It is an instrument which should be used in all learner's dictionaries. In some cases, illustrations are used to explicate and illustrate the difference between polysemes of a lemma.

vlerk   naamwoord ■
(vlerk, vlerkie)
❶ 'n Voël/insek se vlerke is die dele aan weerskante van die lyf waarmee dit vlieg. ⇨ *Die voël het sy vlerk [wing] seergemaak en kan nou nie verder vlieg nie.* ⇨ *Die arend sprei sy vlerke [wings] en vlieg van die kranse af.* ❷ Vlerke van 'n vliegtuig is die dele wat die vliegtuig in die lug hou. ⇨ *As 'n mens 'n sitplek by die vlerk [wing] van 'n vliegtuig het, kan jy omtrent niks sien nie.*

Figure 12: Illustrations of "vlerk"

3.4.7   Usage notes

An important instrument to convey pragmatic information is usage notes. Special blocked usage notes are employed in NWSG, aiming to warn the user against typical mistakes. The presentation of the usage notes follows the theme of the textbooks and is marked with a bee or a tortoise.

---

aanmekaar   adjektief ■
                          en bywoord

(-; -, -)
❶ Wanneer iets aanmekaar is, is dit vas.
⇨ *Woorde wat sinne aanmekaar [join] las, noem ons voegwoorde.* ⇨ *As ek sinne aanmekaar [join] skryf, moet ek leestekens gebruik.*
❷ Dit beteken om nie op te hou met iets nie. ⇨ *My broer praat aanmekaar [continuously] en gee my nie 'n kans om iets te sê nie.*

---

aanhalingsteken   naamwoord ■
(aanhalingstekens, aanhalingstekentjie)
Wanneer 'n mens die woorde wat iemand sê, wil neerskryf, gebruik jy 'n aanhalingsteken aan die begin en einde van die woorde, byvoorbeeld "……".
⇨ *Daar is geen aanhalingstekens [quotation marks] in die indirekte rede nie.*

Nomathembu sê: **Ek wil ook 'n woordeboek hê.**

Pasop vir **aanmekaar** en **aan mekaar:**
Aan mekaar word in die volgende geval as twee woorde geskryf: Wanneer iemand letterlik iets vir iemand anders gee, gee hulle dit aan mekaar.
⇨ *Hulle gee tydens die huweliksseremonie die ringe aan mekaar* **[to one another]**.

Wanneer gebruik jy 'n aanhalingsteken?
Wanneer jy mense se direkte woorde aanhaal, byvoorbeeld:
⇨ *Bets vra: "Kan ek asseblief saamgaan fliek toe?"* ⇨ *My ma sê: "Julle moet al julle kos eet."*

Figure 13: Usage notes

The following typical problems of second and third language speakers are addressed in the usage notes:

(a)   function switching, for instance in the case of **hulp** and **help,**

(b)   meanings of words easily confused, for example **herken** and **erken**, **verveeld** and **vervelig,**

---

ver<u>vee</u>ld   adjektief ■
(verveelde; verveelder, die verveeldste)
As 'n mens verveeld is, stel jy nie in iets belang nie. ⇨ *As die onderwyser maklike werk doen, is die meeste kinders verveeld [bored]. ⇨ Die onderwyser sukkel om die verveelde [bored] leerders wakker te hou.*

kyk by **vervelig**

---

ver<u>ve</u>lig   adjektief ■
(vervelige; verveliger, die verveligste)
As iets vervelig is, is dit nie interessant nie en stel jy nie daarin belang nie.
⇨ *Die les was so vervelig [boring] dat ek amper aan die slaap geraak het. ⇨ Die lewe op ons dorpie is nooit vervelig [boring] nie. Daar gebeur elke dag iets opwindends.*

Pasop vir die verskil tussen **verveeld** en **vervelig:** iemand is **verveeld,** iets is **vervelig.** 'n Mens gebruik **verveeld** vir iemand wat nie in iets belangstel nie, byvoorbeeld: ⇨ *Die slim student is verveeld [bored], want die werk is oninteressant..* 'n Mens gebruik **vervelig** vir iets waarin jy nie belangstel nie en wat jou nie boei nie, byvoorbeeld: ⇨ *Die slim student hou nie van vervelige [boring] lesse nie.*

Figure 14: Usage note addressing confusing words

(c)   direct translations between English and Afrikaans, for example **hard** en **moeilik**,

(d)   word pairs often confused and used incorrectly, for example **ken** and **weet**, and

(e)   word order problems, especially concerning conjunctions, **want** and **sodat**.

sodat   voegwoord ■

'n Mens gebruik sodat as jy die doel van iets wil gee. ⇨ *Hy werk baie hard sodat [so that] hy nie druip nie.* ⇨ *Werk eerder in die week sodat [so that] jy nie naweke hoef te werk nie.*

Pasop vir die volgorde van die woorde na **sodat**: Verbind die volgende twee sinne met **sodat**: ⇨ *Die meisie oefen hard. Sy kan goed speel.* ⇨ *Die meisie oefen hard* sodat *[so that] sy goed* kan speel. Die werkwoorde van die tweede sin skuif na die einde van die saamgevoegde sin.

want   voegwoord ■

'n Mens gebruik want wanneer jy twee sinne aanmekaar wil skryf. Die tweede sin wat na want volg, gee gewoonlik die rede vir iets. ⇨ *Ons het huis toe gehardloop, want [because] dit het begin reën.* ⇨ *My suster is by die huis, want [because] sy is siek.*

Pasop: Die volgorde van die woorde ná **want** bly presies dieselfde as in die sin voor **want**. Verbind die volgende twee sinne met **want**: ⇨ *Ek wil nou eet. Ek is honger.* ⇨ *Ek wil nou eet, want [because] ek is honger.* Ons gebruik 'n komma (,) voor want.

⊘ sinoniem **aangesien/omdat**

Figure 15: Usage notes addressing word order problems

The lexicographers attempted to enhance and facilitate text reception and text production by including the above-mentioned data types. Another important principle that should be applied in any learner's dictionary is accessibility to

the presented data. In other words, an inaccessible presentation of data can handicap or wreck the communication process of the learner. This brings us to the micro-architecture of dictionary articles.

3.5   The micro-architecture of NWSG

The micro-architecture of dictionary articles should contribute to an accessible and user-friendly presentation. This refers to the form of the article, the coherence between the different article zones and the typographical conventions. In a learner's dictionary for school children where the users have limited dictionary skills, it is especially important to keep the structure of the articles as simple as possible. The micro-architecture of NWSG makes provision for different polysemes that are numbered and blocked under one another.

Some structural markers are used to differentiate between different data categories and search zones in order to make them more identifiable. Synonyms and antonyms are marked with different circles, example sentences are preceded by arrows (⇨) and parts of speech are shifted to the right of the article and marked with a black square (■). The search route of the learner is accelerated by the use of these access markers. Homonyms are presented in different articles and numbered. This systematic presentation informs the learner that there is no meaning correlation between the different homonyms. All the information about the structure and presentation in the dictionary is explained in the front text 'How to use your dictionary'. The learner can use this text for any inquiries about the central word list, as well as other texts included in the dictionary.

3.6   Frame structure

The frame structure is another structure lexicographers can apply in a learner's dictionary to enhance the access and improve the data distribution of the dictionary. Dictionaries are text type carriers containing different types of texts. Many dictionaries focus only on the central text and ignore the possibility of presenting data in other venues. In NWSG, good use of the frame structure has been made and a number of other texts have been added.

3.6.1   Front matter

In the front matter, user guidelines that are functional and user-friendly have been added. These guidelines include an analysis of a number of articles from the central text and show the user what types of data are presented and how they can find exactly what they are looking for. The way homonyms and polysemes are treated is also explained. The front matter further includes a table of contents.

3.6.2   Back matter

In the back matter, the user will find a number of useful lists, including the list of translation equivalents, as well as listings of the days of the week, the months of the year, the South African provinces and their capitals and a list of numerals.

4.   Conclusion

Nuwe Woordeboek sonder Grense functions as a bridge between different typological classes within the family of learner's dictionaries. The aim of this dictionary is to assist the inexperienced user during the dictionary consultation procedure. In order to achieve this goal, special attention was given to macro- and microstructural innovations. Being integrated within a textbook and workbook series, the dictionary can be used as an optimal communication instrument in the classroom.

References

Atkins, B.T. 1985. Monolingual and Bilingual Learners' Dictionaries: A Comparison. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 15-24. Oxford: Pergamon Press.

Bakker, N. 1999. Review of *The South African Oxford School Dictionary*. *Lexikos* 9: 294-295.

De Schryver, G.M. and D.J. Prinsloo. 2003. Compiling a Lemma-Sign List for a Specific Target User Group: The Junior Dictionary as a Case in Point. *Dictionaries* 24: 28-58.

Gouws, R., M. Stark and L. Gouws (Eds.). 2004. *Nuwe Woordeboek sonder Grense*. Cape Town: Maskew Miller Longman.

Kernerman, L. 2000. Pedagogic Lexicography: Some Recent Advances and Some Questions about the Future. Heid, U. et al. (Eds.). 2000. *Proceedings of the Ninth Euralex International Congress, Euralex 2000. Stuttgart, Germany, August 8–12, 2000*: 825-829.

Otto, A.N. 1989. *Kriteria vir 'n Afrikaanse aanleerderwoordeboek*. Unpublished D.Litt. Thesis. Stellenbosch: University of Stellenbosch.

Tarp, S. 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.

Tarp, S. and R.H. Gouws. 2004. Wie leer wat uit Afrikaanse (aan)leerderwoordeboeke. *Tydskrif vir Geesteswetenskappe* 44(4): 276-298.

On Defining "Imaginary" Beings and Attributes: How Do Lexicographers Cope with Culturally Determined Differences in Beliefs about Cosmology, Ontology and Epistemology?*

Piet Swanepoel, *Department of Afrikaans and Theory of Literature, University of South Africa, Pretoria, Republic of South Africa (swaneph@unisa.ac.za)*

Abstract: Members of linguistic communities often have opposing beliefs about the existence of beings denoted by lexical items or about the truth of the attributes ascribed to entities. As very little research has been forthcoming in this regard, this article focuses on how people's beliefs about existence and truth are encoded in explanatory dictionaries, and on the kind of semantics that is needed to account for these beliefs. The way in which dictionaries define issues of existence and truth against a default world view is outlined in Section 2. Section 3 indicates what happens if the default world view of lexicographic descriptions changes and how cultural biases operate in the treatment of the meaning of lexical items that denote "imaginary" beings or "imaginary" attributes. Section 4 summarizes the main findings of the article and delimits topics for further research.

Keywords: IMAGINARY BEINGS, FABULOUS CREATURES, DEFINING ATTRIBUTES, LEXICOGRAPHIC DEFINITIONS, EXISTENCE, TRUTH, SENSE, REFERENCE, DENOTATION, IDEALIZED COGNITIVE MODELS, COGNITIVE SEMANTICS, CULTURAL BIAS, CULTURAL SENSITIVITY

Opsomming: Oor die definiëring van "denkbeeldige" wesens en eienskappe: Hoe hanteer leksikograwe kultureel bepaalde verskille in beskouings oor die kosmologie, ontologie en epistemologie? Lede van taalgemeenskappe het dikwels opponerende beskouings oor die bestaan van wesens waarna leksikale items verwys of oor die waarheid van die eienskappe wat aan entiteite toegeskryf word. Omdat baie min navorsing in dié verband beskikbaar is, fokus hierdie artikel op hoe mense se beskouings oor die bestaan en waarheid ten opsigte hiervan in verklarende woordeboeke gekodeer word, en op die soort semantiek wat benodig word om hierdie opvattings te verantwoord. Die manier waarop woordeboeke kwessies van bestaan en waarheid teen 'n verstekwêreldbeskouing omskryf, word in Afdeling 2 gesket. Afdeling 3 dui aan wat gebeur indien die verstekwêreldbeskouing van leksikografiese

---

omskrywings verander en hoe kulturele vooroordele 'n uitwerking het op die behandeling van die betekenis van leksikale items wat verwys na "denkbeeldige" wesens of "denkbeeldige" eienskappe. Afdeling 4 som die hoofbevindinge van die artikel op en baken onderwerpe vir verdere navorsing af.

Sleutelwoorde: DENKBEELDIGE WESENS, FABELAGTIGE WESENS, DEFINIEEREIEN-SKAPPE, LEKSIKOGRAFIESE DEFINISIES, BESTAAN, WAARHEID, BETEKENIS, VERWYSINGS, DENOTASIE, GEÏDEALISEERDE KOGNITIEWE MODELLE, KOGNITIEWE SEMANTIEK, KUL-TURELE VOOROORDEEL, KULTURELE SENSITIWITEIT

1.   Introduction

Lexicographers faced with the task of compiling a monolingual dictionary of a language will often find that (a) such a language contains lexical items that refer to entities or beings about which the linguistic community may have opposing beliefs as to whether such entities in fact exist, and/or that (b) members of a linguistic community have opposing beliefs as to whether the ascribed attributes of some entities are in fact true.

Typical examples of (a) are lexical items denoting entities which belong to the domain of religion, the spiritual world, or the supernatural. In English, lexical items such as *God*, *god*, *heaven*, *hell*, *angel*, *spirit*, and *ghost* fall in this category; and in Zulu, lexical items such as *tokoloshe*, *mamlambo*, and *(i)mpundulu*. In the *Cambridge International Dictionary of English* (CIDE), a class of so-called "imaginary creatures" is distinguished, and the class includes such beings as dragons, unicorns, witches, wizards, fairies, goblins and ghosts, and also the grim reaper. As category they are considered to be figments of human imagination, "not real", because they are created by and only exists in the mind (cf. CIDE, p. 704). As the list indicates, the category includes, besides entities from the domain of religion, also words referring to entities in Greek mythology, fairy tales, and other narratives.

Cases of (b) are illustrated by beliefs such as sangomas possessing the power of divination, the cancer-bush having the property of healing cancer, or lucky charms bringing the wearers luck. In these cases, however, the existence of the entities is not in doubt, but members of the linguistic community may have opposing beliefs as to the truth of the attributes ascribed to these entities.

The importance of these questions of existence and truth for lexicographers lies in the fact that these beliefs often constitute core beliefs and values for a linguistic community, and that irreverence of this fact holds the danger of such a dictionary being rejected by the very community for which it is meant as linguistic tool. For example, Christian and Moslem communities believe in the existence of God, and defining *God* as an "*imaginary* being" will no doubt evoke uproar. Likewise, some Zulu speakers believe sangomas to have the power of divination, and either omitting this belief in dictionary definitions or treating it as a possibly false or misguided belief would do injustice to the meaning these

speakers ascribe to the lexical item and would be found offensive or culturally insensitive.

The question thus arises (a) whether beliefs about the existence of entities and beliefs about the truth of the attributes ascribed to some entities are encoded in dictionaries, (b) if so, what descriptive practices are followed in this regard, and (c) what kind of semantics is needed to provide an adequate account of the beliefs members of a linguistic community have about issues of existence and truth.

As will be indicated below, dictionaries in fact do encode beliefs about the existence of entities and beliefs and about the truth of the attributes of others, although very little research could be found that spells out either what lexicographic principles are followed in this regard or what kind of lexical semantics backs up lexicographic principles and descriptive practices.

The first main goal of this article is therefore to provide such an analysis.

In Section 2, current lexicographic views and practices in a number of modern dictionaries are analysed. It is indicated that ontological and epistemological beliefs are indeed incorporated into lexicographic definitions and that lexicographers have developed a number of descriptive devices to signal differences in the ontological and epistemological beliefs held by a linguistic community. It is argued, furthermore, that in the dictionaries under consideration, lexicographers operate with a default world view (i.e. cosmology) in which the physical world is the main point of reference. All other "possible" worlds are taken as some form of deviance and marked as such. Issues of existence and truth are treated on the basis of possible worlds, allowing the lexicographer — on the face of it — to avoid direct personal commitments as to such ontological and epistemological questions.

In Section 3, it is indicated, first of all, how a change in the default world view (cosmology) of a dictionary or encyclopaedia changes the way lexical items are defined and how ontological and epistemological beliefs are integrated in such a world view and in encyclopaedic entries. Defining practices in the Dutch medieval bestiary *Der naturen bloeme* ("The most beautiful from nature") are analysed and used as illustration.

Major shifts in world views often take ages, but monolingual dictionaries based on historical principles are often faced with the problem of having to reflect such changes. For example, in the Middle Ages, dragons were believed to exist in the physical world; now they are deemed to be imaginary creatures. To illustrate some of the current approaches to such historical changes in world views, an analysis is provided in Section 3.2 of the way the senses of the lexical item *dragon* is presented in the OED. The cultural bias that is operative in the treatment of *dragon* is also illustrated with regard to the way lexical borrowings can be dealt with.

A major problem confronting lexicographers in this regard is finding a semantic theory which could provide a basis for analysing people's beliefs about existence and truth and which could be used for the description of these beliefs in the definitions in an explanatory dictionary.

In the rest of this article, cognitive semantics will be used for this purpose. The three main constructs which will be analysed are dictionary definitions, lexical meanings, various kinds of Idealized Cognitive Models and the relationships between them:

dictionary definitions — lexical meanings — Idealized Cognitive Models

In line with cognitive semantics, it will be accepted that lexical meanings are mental constructs consisting of one or more beliefs (including those on existence and truth), and no strict distinction is maintained between so-called linguistic and encyclopaedic beliefs. Furthermore it is accepted that the meaning of lexical items (or its comprising beliefs) relate to, are construed against or motivated by more encompassing mental constructs, often called frames, cultural models or, more generally, Idealized Cognitive Models (ICMs) (cf. Holland and Quinn 1987, Lakoff 1987, Lakoff and Turner 1989, Palmer 1996, Shore 1996 and Taylor 1995 for a detailed discussion). World views are taken to be one class of such an ICM, which functions as justificatory contexts for certain beliefs. For example, a number of religions define a world or cosmos, and thus a reality, that includes not only the physical world but also a spiritual world with places equivalent to the Christian heaven and hell. Holy scriptures and religious dogma, as well as religious metaphysics, are the sources of these realities and they dictate how the real world is to be construed and populated, and how these worlds, their creatures and their logics or laws within such a cosmos interact.

In this regard, domains of religious experience and understanding are world views which are socially acquired and which explicates, according to Goodenough (see Keesing 1974: 77), what has to be known or believed in order to operate in a society. Such a world view furthermore acts as a model of what constitutes reality, as a model for perceiving, relating, and otherwise interpreting material phenomena; it provides standards for deciding what is or what can be, for deciding what to feel about it, what to do about it, and how to go about doing it. Such world views or ICMs thus permeate all our experiences and understanding of the realities with which we interact on a daily basis, and lexical items denote chunks of these experiences and understanding.

As noted by Lyons (1977: 209-211), there are currently no strict rules or preferred approaches of methods as to what dictionary definitions have to encode about lexical meanings as mental constructs. In most cases, dictionary definitions are constructed with pragmatic goals in mind, such as to explain the meaning(s) of lexical items; consequently their success should be judged in terms of these pragmatic goals. The focus in this article, however, will be on how the defining features chosen for lexicographic definitions relate to specific beliefs about the existence of entities and their attributes and the larger encompassing ICMs against which these beliefs are motivated.

2.   Encoding ontological and epistemological beliefs: default and deviance in some modern dictionaries

2.1   Default presuppositions and justificatory context

Dictionary definitions by default attempt to describe and explain the meaning(s) people of a linguistic community associate with a lexical item, and they do so by listing a number of such belief-attributes. For example, CIDE, p. 567, defines the primary meaning of the lexical item *frog* in terms of a belief about the category this entity belongs to and in terms of a number of beliefs about its morphology and habitat:

(1)   **frog** ANIMAL ... a small animal which has a smooth skin, lives in water and on land, has long powerful back legs with which it jumps, from place to place, has no tail, and is usually greenish-brown in colour

The first hypothesis to be ventured is that specific ontological and epistemological presuppositions are associated by dictionary users with such lexicographic descriptions of entities. The first is that the entity denoted by the lexical items described in fact exists, and more specifically that it exists as entity within the *physical world* as this is perceived by our senses (or extensions thereof such as the microscope or telescope). Lyons (1977: 209) notes in this regard that both reference and denotation with the help of language depend on and presuppose the existence of entities and their attributes to which lexical items refer. This view is termed the axiom of existence, i.e that whatever is denoted by a lexical item must exist, just as whatever is referred to by means of a lexical item must exist.

Secondly, it is postulated that a lexicographic meaning description carries a presupposition of truth with regard to the belief attributes ascribed to this entity, i.e., for example, that it is true that in our physical world a frog belongs to a cognitively salient category (the category SMALL ANIMAL) and that it is true that it has the attributes listed in the definition. To put it simply: We trust that lexicographers will provide us with true information with regard to the beliefs associated with a lexical item. However, the basis for the acceptance of such claims about existence and truth is its verifiability in terms of our sensory experiences of the empirical, physical or natural world, or, as additional justification, the theories and findings of the natural sciences.

The reality against which both the issue of the existence of a frog and the issue of the truth of its attributes are defined, is thus a reality as defined by what we consider to be the empirical, physical or natural world. Lakoff (1987: 158) uses the term *basic realism* or *experiential realism* to denote this world view and describes it as, amongst others,

—   a commitment to the existence of a real world, both external to human beings and including the reality of human experience,

— a link of some sort between human conceptual systems (the belief system alluded to above — PHS) and other aspects of reality,

— a conception of truth that is not merely based on internal coherence (of some formal system — PHS), and

— a commitment to the existence of stable knowledge of the external world.

By default, lexicographic definitions thus have a logical structure of the following form:

[Ontological Marker [[Truth Marker] Attribute 1, ..., Attribute n]]
where

the ontological marker (OM) in this case refers to the external physical or empirical world *as perceived by the language user*, although not explicitly marked or indicated in the definition itself,

[A1, ..., An] refers to whatever attributes are presented to describe or explain the conceptualization of the entity to which the lexical item being defined, refers, and

the truth marker (TM), which, although not lexically encoded in the default contexts, reflects the presupposition that the listed attributes are true as commonly believed by members of a linguistic community.

As indicated, [A1, ..., An] falls within the scope of the OM and the TM. In general, existence and truth in lexicographic definitions are thus defined relative to some domain of human experience and understanding (and thus meaning). In philosophical terms, the OM may also be interpreted as an indicator of a "possible world", thus indicating that questions of existence and truth are defined, in this default and unmarked case, as the physical world.

2.2   Deviance and its encoding

2.2.1   Encoding existence relative to some domain of experience and understanding

Deviance from this default world view and empirical reality as justificatory context is signalled in modern dictionaries by means of a wide range of explicitly encoded ontological and epistemological markers. These are labels and phrases which are presented as parts of definitions and which relativize the question of the existence of entities and the truth of attributes of entities to some possible world or domain of experience and understanding (and thus of meaning) that acts as reference point for issues of existence and truth. In the following examples, these ontological markers are in bold italics. The examples

are meant to be illustrative and are by no means exhaustive of all possible ways in which the world of reference can be marked.

(2)   **dragon** ... a large fierce *imaginary* animal, usually presented with wings, a long tail and fire coming out of its mouth (CIDE, p. 418)

(3)   **God** ... (without article) *(in Christianity and other monotheistic religions)* the creator and ruler of the universe and source of all moral authority; the supreme being (NODE, p. 785)

(4)   **heaven** ... the place where God or the gods *are supposed* to live; place of complete happiness where the souls of good people go after death (LDOCE, p. 524)

(5)   **hell** ... *in some religions*, the place where some people *are believed* to go after death to be punished forever for the bad things they have done during their lives (CIDE, p. 661)

(6)   **fairy** ... a *mythical* being *of folklore and romance* usu. having diminutive human form and magic powers and dwelling on earth in close relationship with man (Webster's Third, p. 816)

All of these definitions can be related to the propositional structure (1), where the bold italicized phrases act as ontological markers, indicating the point of reference in terms of which the existence of the entity referred to should be interpreted and the truth of the attributes assigned to them. For example, (6) can be interpreted as: "in the world/domain of myths, folklore and romance, fairies usually have diminutive human form, they have magical powers, and they dwell on earth in close relationship with man". The ontological markers in these examples refer to a wide range of human domains of experience and understanding, but they are explicitly marked to indicate that the default justificatory context is not the physical world, but specific domains of human experience and understanding.

As indicated by Tolkien (n.d), one could use the construct *Faërie* as OM in example (6), where *Faërie* then refers to the whole world created in fairy tales, which is one not only populated by fairies, but also by all kinds of creatures and places, activities, processes, etc., and where the laws of the physical world do not apply *in toto*. Likewise, religions constitute such a world and encode or present a view of the cosmos where not only the physical world exists, but also places such as a heaven and a hell which are populated by various spirits/spiritual beings. As narrative sources such as the Bible, Christian dogma and metaphysics tell us, these exist in reality and have certain properties. In dictionary definitions, an OM such as *Faërie/in fairy tales* or *in some religions* thus act, albeit often metonymical, to signal such real worlds or realities.

Although semanticists such as Lakoff (cf. Lakoff 1987) and philosophers such as Niiniluoto (cf. Niiniluoto 1999: 23-25) indicate that these domains of experience have as their source the human imagination, for some people they

are part and parcel of the reality that exists external to them, and they are made *real* by various means: by way of symbolization in images and representations which can be empirically cognized and understood, and by the fact that people act on these domains of experience once they have been created, as in religious rituals where spiritual beings are seen as living and acting and to which people pray and bring offerings. Sightings of such beings, or experiences of their presence or attributes are in fact often provided as further (empirical) justification for their existence and for the truthfulness of their attributes (cf. Palmer 1996 and Shore 1996).

Such a view of religion differs markedly from the views on religion and reality in some philosophical schools of thought. According to Niiniluoto (1999), for example, religious beliefs belong to Popper's World 3 of socially and culturally constructed entities. They are seen simply as products of the human imagination, and as such cannot be used as justificatory contexts, given empirical verifiability as criterium for both existence and truth. As Niiniluoto (1999: 233, fn. 6) notes:

> Gods do not exist, but were created by men as idealized pictures of themselves. Therefore — excluding descriptive statements about the history, psychology, and sociology of religion — genuine religious judgements about gods and their attributes are not true.

Niiniluoto (1999) argues, however, that for most people their reality is made up of various such domains of experience and understanding, and they may accept and invoke different logics and different justificatory and explanatory contexts for questions as to what exists and what is true, for example, common sense understanding and the natural sciences when it comes to the physical world, but the Bible, religious dogma and metaphysics when it comes to the domain of religion. Such "folk" models of reality therefore of necessity often contain conflicting beliefs (for example Darwinian beliefs about evolution versus creation according to Genesis), but it is an incoherence people accept and live by.

On the face of it, relativization of the questions of existence and truth to some domain of experience and understanding thus opens up the possibility for the lexicographer to "objectively" describe such ontological and epistemological commitments/beliefs of members of a language community. However, given that the empirical world is used as default justificatory and explanatory context of necessity, also encodes a commitment to and prioritization of the physical/empirical world, and thus to a specific world view which is vehemently opposed by anti-realists, anti-objectivists and subjectivists (cf. Niiniluoto 1999). The latter hold (with various nuances of belief) that there is nothing outside of human experience and understanding, i.e. epistemology is prioritized above and precedes ontology, and what is experienced as real, existing and true, first and foremost has its origin in the human mind.

With regard to the issue of existence, there are, however, also some cases

which are not so clear-cut and different ontological markers could then be used to signal different contexts of justification. In the following example, the ontological marker *mythical or actual* is used to signal that the mamlambo may in fact be some mythical creature (and exists in some mythical domain of experience) or the lexical item may be used to denote an entity that in fact exists in the physical world:

(7)   **mamlambo** ... A river snake *mythical or actual*, about which there are varied beliefs (DSAE, p. 211)

2.2.2   Encoding the relative truth of attributes

A second set of relevant examples are definitions of entities which are accepted or believed to exist in the real word, but of which there may be differences of opinion as to the truth of the attributes assigned to them. In these definitions, the ontological context is the physical world and the entities are accepted to exist in this world, but truth markers (TMs) are used as part of definitions to indicate that the issue of truth must be relativized to some point of reference. Compare the following examples:

(8)   **lucky charm** *she always wears a lucky charm* (= one *that she thinks* will cause good things to happen by chance) (CIDE, p. 846)

(9)   **sangoma** ... An African witchdoctor …, usu. a woman *often claiming* supernatural powers of divination (DSAE, p. 306)

(10)   **(i)mpundulu (bird)** ... The lightning bird ... An evil spirit *about which there are various beliefs* … It is invoked by witches, and freq. cited before judges as the instigator or cause of crime ... (DSAE, p. 231)

Again, the issue of truth is relativized to some reference point and truth markers are used to indicate which of the beliefs about the attributes of real entities should be read against or are justified in terms of the physical world (the default justificatory context) and which against some other point of reference (the deviant justificatory context). For example, definition (9) makes the statement that a sangoma is usually a woman, but the truth of the attribute of having supernatural powers of divination is relativized to the beliefs and experiences of the sangomas themselves.

One truth marker that is often used, is *believed to (be/have)*. Although the point of reference it denotes is rather vague, it does signal that the truth of what follows in the definition should be relativized to some non-empirical context. This follows mainly because such a truth marker is *not* used when the justificatory context is the physical world. This becomes evident from the oddity caused by inserting such truth markers in definitions that list attributes which are all accepted as being empirically true (by default). See the following definitions:

(11)   **frog** a small animal [*believed to have*] a smooth skin, lives in water and
on land, has long powerful back legs with which it jumps, from place to
place, has no tail, and is usually greenish-brown in colour

(12)   **frog** ANIMAL ... a small animal which has a smooth skin, lives in water
and on land, [*and is believed to have*] long powerful back legs with
which it jumps, from place to place, has no tail, and is greenish-brown in
colour

Just as uncertainties may exist with regard to the issue of existence, lexicographers may also be in doubt as to the truth of some beliefs about the attributes of entities, especially when some form of empirical justification could be brought forward to support these beliefs. In the following example, this is indicated by the truth marker *has, or is believed to have*:

(13)   **witch** ... a woman *who has, or is believed to have,* magic powers, esp.
who can cast spells on people (LDOCE, p. 1264)

3.   Changing the default world view and its lexicographic implications

World views and folk theories can and do change, as is evidenced by the differences in the dominant world views of the Middle Ages, the period of enlightenment during the Renaissance, and that of "modern man" in the 21st century. Over such extended periods, world views may change dramatically or gradually with regard to their conceptual content. Some beliefs may persist, others may undergo only slight modification. In this section, two sets of data will be considered to have a closer look at how such changing world views effect the design of definitions and the way in which issues of existence and truth are handled in such definitions. The first set of data is taken from the medieval bestiaries, and more specifically the Dutch bestiary *Der naturen bloeme* of Jacob van Maerlant. The second set of data is taken from modern dictionaries based on historical principles of which it could be accepted that they should link various periods and changing world views in the entries.

3.1   *Der naturen bloeme*: a culturally-determined "insider's" perspective

Jacob van Maerlant completed *Der naturen bloeme* (DNB) round about 1270. It is a Dutch translation and reworking of Thomas of Cantimpré's encyclopaedia *Liber de natura rerum*, which itself is a prime example of how knowledge and beliefs about nature were compiled in the Middle Ages from diverse sources — in fact, from almost everything that existed and could be attributed to respected authors, past and present. Most of this was done with an explicit didactic aim in mind, viz. to explicate nature for what it could teach us about God and His plan with/for mankind (cf. Nischik 1986: 28-36).

True to the scholastic tradition, very little was made of empirical verification when it came to issues of existence and truth — these were determined by the existing authoritarian resources/narratives. The result is a portrayal of a world with a clear demarcation of the physical and the spiritual (cf. Lakoff and Turner 1989 for a discussion of the Great Chain of Being underlying the Western Christian world view). However, this physical world was believed to be populated with a host of Plinian monster races, a host of creatures from Greek mythology, and a variety of other fabulous creatures, often accredited with equally fabulous or supernatural attributes (cf. Cohen 1992, Clair 1967 and Friedman 1981). As is exemplified in the cases discussed in Section 2, some of these fabulous attributes were also credited to creatures which in fact exist in the external world, for example the notions that the bear licks its offspring to life after birth and that the pelican feeds its offspring with its own blood.

Despite the curious mixture of fable and fact in the medieval bestiaries, the descriptions of what is now considered to be mythical animals was given with the belief that these animals in fact existed in the natural world; consequently, they were described in DNB within a conceptual framework that can be discerned for the category (EXISTING) ANIMAL underlying all the encyclopaedic entries in DNB (cf. Swanepoel 2004):

CONCEPTUAL FRAME FOR THE CATEGORY ANIMAL
1. Name (Latin, Dutch equivalents and etymology)
2. Morphological features (esp. outstanding, strange or deviant physical features)
3. Habitat (often linked to the *mappa mundi*)
4. Behaviour and personality features
5. Food/eating habits
6. Procreation
7. Natural enemies/Danger posed to man
8. Instrumental use for humans
   As source of food
   Preparation of dishes
   As source for medicinal use (linked to the doctrine of the tempers)
   How to hunt down the animal
   Body parts to be used and preparation of medicinal products
   As source for other uses (e.g. clothing)
9. Symbolic-moral values
10. Supplementary information
    Sources of information
    Reflections on factuality/truth of information
11. Illustrations

Table 1: A conceptual framework for describing animals in *Der naturen bloeme*

In DBN's description of the dragon, for example, Van Maerlant (cf. Van Maerlant 1989) lists its morphological features, its habitat, the danger it poses to

man, the elephant as its natural enemy, the medicinal properties of the magic stone it carries in its forehead, ways to catch the dragon, the potions which can be made from the magic stone, and his sources of information.

With regard to the description of animals, Van Maerlant therefore makes no explicit distinction between so-called existing and non-existing animals in DNB, and all of them are viewed, understood and described as a single category, given the belief that all of these animals, including the dragon, inhabited the external world. As such, these encyclopaedic entries present an "insider's" view of what constitutes reality within a specific culture, as backed up by a justificatory context of what the Bible and respected authors have to say about the world, its structure and inhabitants.

3.2   Dictionaries based on historical principles: an "outsider's" view

Dictionaries which base the organization of their definitions on historical principles, such as OED, WNT and Webster's Third, typically organise the definitions as these are evidenced from earliest to most recent usages of a lexical item. What is of interest in this regard, therefore, is the way in which lexical items referring to what is now considered to be imaginary beings, such as the dragon, are defined in these dictionaries. More specifically, the question is how lexicographers treat the changing world views or beliefs about the existence of these animals over the span of a specific period.

In analysing the definitions provided for *dragon* in OED, WNT and Webster's Third, the sources of citations span a period from as early as the 12th century to the 18th century, and two primary senses are discerned for *dragon*, viz. the obsolete reference to a snake or a serpent and then its meaning as it pertains to what is now considered the mythical animal. The following entries from OED are illustrative of the lexicographic approach taken in these three dictionaries. The issue of the existence of the dragon is seen from a modern perspective in as much as it is labelled a *mythical* animal (cf. sense 2), irrespective of the fact of the belief in its existence during the Middle Ages (and in some sources even as late as the 17th century; cf. Borges 1987: 154):

(14)   **dragon** ... 1. A huge serpent or snake; a python. *Obs.* in etymol. use.
2. A ***mythical*** monster, represented as a huge and terrible reptile, usually combining ophidian and crocodilian structure, with strong claws, like a beast or bird or bird of prey, and a scaly skin; it is generally represented with wings, and sometimes as breathing out fire. The heraldic dragon combines reptilian and mammalian form with the addition of wings.
It is difficult to separate senses 1 and 2 in early instances. (OED, p. 1012)

The use of the OM *mythical* therefore clearly indicates that the issue of the existence of the dragon is viewed from the outsider's perspective of the modern

world view of basic realism. The fact that the use of *dragon* in at least some of the sources used for citations in the lexicographic articles could have presupposed the existence of the animal in the real world, is not accounted for. A culturally biased perspective on the meaning and understanding of the denotation of *dragon* is therefore provided with regard to the historical facts.

Borges (1987: 154), in reflecting on this modernist perspective on the dragon within Western cultures, provides the following insightful comment:

> People believed in the reality of the Dragon. In the middle of the sixteenth century, the Dragon is recorded in Conrad Gesner's *Historia Animalium*, a work of a scientific nature.
>
> Time has notably worn away the Dragon's prestige. We believe in the lion as reality and symbol; we believe in the Minotaur as symbol but no longer as reality. The Dragon is perhaps best known but also the least fortunate of fantastic animals. It seems childish to us and usually spoils the stories in which it appears. It is worth remembering, however, that we are dealing with a modern prejudice, due perhaps to a surfeit of Dragons in fairy tales.

3.3   Cultural bias and lexical borrowings

This kind of cultural bias or prejudice is, however, not only a feature of the historical treatment of lexical meaning and the related issues of existence and truth. It could, and also does show up, for example, in the way dictionaries define lexical items borrowed from other languages. The following borrowings from Zulu into English and their definitions in the DSAE are illustrative of this phenomenon:

(15) **(i)mpundulu (bird)** ... The lightning bird ... An evil spirit about which there are various beliefs ... It is invoked by witches, *and freq. cited before judges as the instigator or cause of crime* … (DSAE, p. 231)

(16) **tokoloshe** ... An evil spirit widely believed in by both urban and rural Africans: it is invoked in witchcraft *and offered as an extenuating circumstance in criminal cases* … *Tokoloshe haunted me says the mad Hillbrow killer. I had been attacked by the tokoloshe and I was hopelessly drunk ... The first I knew of what I had done was when the police arrested me at my work on Monday.* Post 18.1.70 (DSAE, p. 372)

(17) **tokoloshe**... (in African folklore) a mischievous and lascivious hairy water spirit (NODE, p. 1949)

What is of interest in these definitions is the choice of defining attributes that was made from all the relevant beliefs people may have of the impundulu bird and the tokoloshe. In both (15) and (16), for example, it is mentioned as defining attribute that Africans frequently offer these two spirits as explanation, excuse or mitigation for some illegal deed before the law. Although merely stated as

observed fact, and without any moral comment, it reflects the lexicographer's bias as a cultural outsider in as much as it is not linked to or contextualized within the religious model of spiritual possession which is part and parcel of a number of traditional African religions (cf. Boddy 1997). In such religions, loss of personal agency through spiritual possession is seen and experienced as a reality.

This case has an analogue in the efforts of the former South African cricket captain, Hansie Cronjé, to attribute his fixing of cricket matches to the work of the devil, but both cases illustrate the basic conflict between the common sense world views of ordinary people about spiritual possession and agency and the world view of the law in which beliefs about spiritual possession and agency are linked to psychological instability — the reference to the *mad* killer in the quotation at *tokoloshe* perhaps being no coincidence.

In definition (17), however, no reference is made to the fact that spiritual possession by the tokoloshe is offered as mitigation for some misdeed, which again raises the question of how lexicographers decide on the defining attributes for their definitions. Such selections for the construction of dictionary definitions are in fact dictated by lexicographers' biases or impressions of the saliency of beliefs.

4.   Conclusion

Given the lack of research on how beliefs of existence and truth are encoded in dictionaries, this study is only exploratory in nature. The findings do, however, permit a few general conclusions.

An analysis of current descriptive practices in a number of dictionaries in Sections 2 and 3 clearly indicates that lexicographers encode the beliefs people have about the existence of the entities referred to by lexical items and about the truth of the attributes accredited to entities. This is clearly done from a specific cultural perspective, and more specifically in modern explanatory dictionaries, from the perspective of reality construed as existing as an independent external real or empirical world and people's common sense understanding of this world. The latter acts as default reference point with regard to issues of existence and truth; all deviances from this point of reference are therefore explicitly marked with the help of ontological and epistemological/truth markers.

The second major point to be made is that one can in principle not construct an explanatory dictionary and do so in an objective way, i.e. without adhering to a specific world view. The decision on what the default world view should be for any dictionary therefore becomes critical in compiling an explanatory dictionary, and this has to be made by a careful analysis of the cultural model(s) of the linguistic communities whose language will be the object of

description. Such models may or may not coincide with those of the lexicographer, but obviously those of the linguistic community should have priority. As indicated in the case of *Der naturen bloeme,* homogeneity with regard to the dominant world view simplifies the lexicographer's task as deviances become unnecessary to mark; heterogeneity, on the other hand, forces one to make an explicit choice for a default world view and deviances have to be marked.

Thirdly, lexicographers should be extremely sensitive to cultural biases in defining lexical items. Careful consideration should be taken as to whose beliefs are encoded in dictionary definitions, especially also in linguistic communities which are culturally very diverse — speakers of (varieties of) English in South Africa, being a prime example.

The analysis and discussion of the encoding of truth and existence in dictionary definitions started off with a rather minimalist set of assumptions about dictionary definitions, lexical meanings as mental constructs, ICMs as broader cognitive constructs against which lexical meanings are defined and motivated and the links that hold between them. It served this exploratory analysis well, but obviously the theoretical framework needs further refinement, especially with regard to the kinds of human domains of experience, understanding and meaning (or ICMs) taken by default or by deviance to act as justificatory contexts and the variety of ways they are encoded in dictionary definitions.

Furthermore, the way in which dictionaries relativize aspects of existence and truth to specific justificatory ICMs, is most probably only a small manifestation of a much wider phenomenon in dictionaries that needs to be researched and explicated. As Niiniluoto (1999: 227-228) indicates, relativism abounds. Concepts of existence and truth are only one of a host of cognitive categories which are relativized to such ICMs:

Cognitive categories taken to be relative in some respect
  Ontological concepts: Objects, facts, world, reality
  Semantic concepts: Truth, reference, meaning
  Epistemological concepts : Perception, belief, justification, knowledge
  Methodological concepts: Inference, rationality, progress
  Moral categories: Customs, values, ethics, law, politics, religion

Cognitive categories to which other categories are taken to be relative
  Persons
  Groups
  Cultures
  Environments
  Languages
  Conceptual frameworks
  Theories
  Paradigms

Points of view
Forms of life
Gender
Social class
Social practices
Social interests
Values

Obviously, then, this article has only superficially touched on a much larger and complex phenomenon.

Bibliography

Dictionaries

CIDE: Proctor, P. (Ed.). 1995. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press.

DSAE: Branford, J. 1987. *A Dictionary of South African English*. (Third edition. Revised and enlarged.) Cape Town: Oxford University Press.

HAT: Schoonees, P.C., C.J. Swanepoel, S.J. du Toit and C.M. Booysen. 1971. *Verklarende handwoordeboek van die Afrikaanse taal*. Klerksdorp: Voortrekkerpers.

LDOCE: Proctor, P. (Ed.). *Longman Dictionary of Contemporary English*. Essex: Longman.

NODE: Pearsall, J. (Ed.). 1998. *The New Oxford Dictionary of English*. Oxford: Oxford University Press.

OED: Simpson, J.A. and E.S.C. Weiner. 1989. *The Oxford English Dictionary*. (Second edition). Oxford: Clarendon Press.

Webster's Third: Gove, P.B. (Ed.). 1961. *Webster's Third New International Dictionary of the English Language*. London: G. Bell and Sons.

WNT: Knuttel, J.A.N. 1916. *Woordenboek der Nederlandsche Taal*. 's-Gravenhage/Leiden: Martinus Nijhoff/A.W. Sijthoff.

Other literature

Boddy, J. 1997. Spiritual Possession. Middleton, J. (Ed.). 1997. *Encyclopedia of Africa South of the Sahara. Volume 3*: 156-160. New York: Charles Scribner's Sons

Borges, J.L. 1967. *The Book of Imaginary Beings*. New York: Random House.

Clair, C. 1967. *Unnatural History. An Illustrated Bestiary*. New York/Toronto: Abelard-Schumann.

Cohen, J.J. 1992. The Use of Monsters in the Middle Ages. *Journal of the Spanish Society of Medieval English Language and Literature* 2: 47-69.

Friedman, J.B. 1981. *The Monstrous Races in Medieval Art and Thought*. Cambridge, Mass.: Harvard University Press.

Holland, D. and N. Quinn (Eds.). 1987. *Cultural Models in Language and Thought*. Cambridge: Cambridge University Press.

Keesing, R.M. 1974. Theories of Culture. *Annual Review of Anthropology* 3: 73-97.

Lakoff, G. 1987. *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.

Lakoff, G. and M. Turner. 1989. *More than Cool Reason. A Field Guide to Poetic Metaphor*. Chicago/London: The University of Chicago Press.

Lyons, J. 1977. *Semantics. Volume 1*. Cambridge: Cambridge University Press.

Niiniluoto, I. 1999. *Critical Scientific Realism*. Oxford: Oxford University Press.

Nischik, T-M. 1986. *Das volksprachliche Naturbuch im späten Mittelalter; Sachkunde und Dinginterpretation bei Jacob van Maerlant und Konrad von Megenberg*. Tübingen: Max Niemeyer Verlag.

Palmer, G.B. 1996. *Toward a Theory of Cultural Linguistics*. Austin: University of Texas Press.

Shore, B. 1996. *Culture in Mind. Cognition, Culture, and the Problem of Meaning*. New York/Oxford: Oxford University Press.

Swanepoel, P. 2004. Wonderbaarlike volke en monsterdiere in *Der naturen bloeme* van Jacob van Maerlant en die kognitiewe verankering van denkbeeldige kategorieë in die Middeleeuse lewens- en wêreldbeskouing**.** *Tydskrif vir Nederlands en Afrikaans* 11(2): 162-183.

Taylor, J.R. 1995. *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.

Tolkien, J.R.R. n.d. *On Fairy Stories* [Online]. Available at: <http://larsen-family.us/~1066/onfairystories.html [viewed 15 April 2005].

Van Maerlant, J. 1989. *Het boek der natuur.* Compilation and translation by Peter Burger. Amsterdam: Em. Querido's Uitgeverij.

Über die Datenakzessivität in Printwörterbüchern. Einblicke in neuere Entwicklungen einer Theorie der Wörterbuchform[*]

Herbert Ernst Wiegand, *Germanistisches Seminar, Universität Heidelberg, Heidelberg, Bundesrepublik Deutschland*
(herbert.ernst.wiegand@gs.uni-heidelberg.de)

Zusammenfassung: In diesem Beitrag werden die Begriffe eingeführt, die man benötigt, um das Datenakzessivitätsprofil von Printwörterbüchern genau beschreiben zu können. Es wird zwischen externer und interner Datenakzessivität unterschieden; erstere ist obligatorisch, letztere ist fakultativ. Die externe Datenakzessivität wird durch äußere, die interne Datenakzessivität durch innere Zugriffsstrukturen innerhalb von akzessiven Wörterbucheinträgen hergestellt. Es werden unterschiedliche Typen von äußeren Zugriffsstrukturen als lineare Ordnungsstrukturen, wie z.B. alphabetische äußere Zugriffsstrukturen, numerische mediostrukturelle Zugriffsstrukturen, Registerzugriffsstrukturen u.a., beschrieben und ihr Funktionieren erklärt. Weiterhin werden äußere Zugriffsstrukturen von Zugriffspfaden abgegrenzt, die durch die Ausführung externer Zugriffshandlungen der Benutzer etabliert werden. Der Beitrag gibt insgesamt fünf Einblicke, aber keine zusammenhängende Übersicht über die vielfachen Ausprägungen der Eigenschaften der Wörterbuchform, die die Datenakzessivität sicherstellen.

Stichwörter: AKZESSIVER WÖRTERBUCHEINTRAG, ALPHABET, ALPHABETISCHE ÄUßERE ZUGRIFFSSTRUKTUR, ÄUßERER ZUGRIFFSPFAD, EXTERNE DATENAKZESSIVITÄT, EXTERNE ZUGRIFFSHANDLUNG, INNERE ZUGRIFFSSTRUKTUR, INTERNE DATENAKZESSIVITÄT, MEDIOSTRUKTURELLES LEITELEMENT, MEDIOSTRUKTURELLE ZUGRIFFSSTRUKTUR, MONOAKZESSIVES WÖRTERBUCH, REGISTER, REGISTERZUGRIFFSSTRUKTUR, POLYAKZESSIVES WÖRTERBUCH, SCHNELLZUGRIFFSSTRUKTUR, ZUGRIFFSTEXTELEMENT

Abstract: **On the Data Accessibility in Printed Dictionaries. Insights into Recent Developments of a Theory of the Form of Dictionaries.** In this contribution, concepts are introduced that are needed for a precise description of the data accessibility profile of printed dictionaries. A distinction is made between external and internal data accessibility, with the first being obligatory and the second being optional. The external data accessibility is presented by means of external access structures and the internal data accessibility by means of internal access structures in accessible dictionary entries. Distinctive types of external access structures as linear ordering structures, e.g. alphabetical external access structures, numerical mediostructural

---

access structures and register access structures are described and their functions explained. External access structures are furthermore distinguished from access routes, established by the user by means of the execution of external access procedures. The contribution gives a total of five insights but no coherent overview of the multiple manifestations of the characteristics of the form of dictionaries established by the data accessibility.

1.   Vorbemerkung

Printwörterbücher gehören zur wissensvermittelnden Literatur. Dies ist nicht so zu verstehen, dass Wissen in den Wörterbüchern zu finden ist, sondern vielmehr so, dass anhand der lexikographischen Daten Wissen von den Benutzern erarbeitet werden kann. Wissen ist ein kognitives Phänomen von Personen, das u.a. anhand von semiotisch gebundenem Wissen und damit vor allem anhand von geäußerten Sprachzeichen erhältlich ist.[1] Innerhalb der wissensvermittelnden Literatur haben die Wörterbücher eine Sonderstellung: Sie sind vor allem — wenn auch nicht nur — zum Nachschlagen konzipiert; entsprechend gilt einerseits: Ein gedrucktes Buch, in dem man nicht nachschlagen kann, wird nicht als Printwörterbuch betrachtet. Es ist aber andererseits nicht jedes Buch, das zur wissensvermittelnden Literatur gehört und in dem man beispielsweise anhand eines Namenregisters und/oder eines Sachregisters und/oder eines Wortregisters Daten nachschlagen kann, ein Wörterbuch. Die *externe Datenakzessivität* als eine Eigenschaft der Wörterbuchform, oder — anders ausgedrückt — die *Nachschlagbarkeit* und damit die *Zugriffsbereitschaft* der lexikographischen Daten für einen Zugriff, der von außen auf ein Wörterbuch erfolgt, ist für die Frage, ob ein gedrucktes Werk als Wörterbuch gelten kann, nur eine objektiv notwendige und keine objektiv hinreichende Bedingung. Der Frage, welche Eigenschaften der Wörterbuchform und/oder des Wörterbuchgegenstandes (*sensu* Wiegand 1998: 302) als objektiv hinreichende Bedingung zu gelten haben, werde ich hier nicht nachgehen, weil dann ein ganz anderes Thema zu behandeln wäre.

Gerade habe ich festgestellt, dass Wörterbücher *vor allem* zum Nachschlagen konzipiert sind. Dies möchte ich so verstanden wissen: Es gibt einerseits Wörterbücher, die ausschließlich zum Nachschlagen in dem Sinne konzipiert sind, dass ein Zugriff auf die Daten von außerhalb des Wörterbuchs kommend erfolgen kann; andererseits gibt es Wörterbücher, die nicht ausschließlich für

das Nachschlagen von außen und damit für den direkten externen Zugriff auf Daten anhand von Leitelementen konzipiert sind, sondern auch zum wörterbuchinternen Recherchieren mittels eines Systems von Verweisungen zum Zwecke der Vermittlung von Wissenszusammenhängen, die durch die wörterbuchinterne Datendistribution auf zugriffsverschiedene Positionen nicht erkennbar sind. In den erstgenannten Wörterbüchern ist die Organisation der Datenakzessivität weniger komplex als in den Letztgenannten.

Ich gehe daher in diesem einführenden Beitrag wie folgt exemplarisch vor: Ich führe zunächst die für das Thema benötigten grundlegenden lexikographietheoretischen Termini am Beispiel von Wörterbüchern ein, die ausschließlich zum Nachschlagen konzipiert sind. Dann gehe ich zu solchen Wörterbüchern über, bei denen die Datenakzessivität aufgrund der Existenz von Verweisungen und integrierten Umtexten komplexer ist.[2] Insgesamt gebe ich fünf Einblicke, die aufeinander abgestimmt sind.

2.   Akzessive Wörterbucheinträge, externe versus interne Datenakzessivität

Für die Datenakzessivität von gedruckten Nachschlagewerken jeder Art, also auch für die Datenakzessivität von nichtlexikographischen Nachschlagewerken, gilt das Folgende: Man muss zwischen der *externen Datenakzessivität* und der *internen Datenakzessivität* unterscheiden. Die externe Datenakzessivität wird durch die Einrichtung von *äußeren Zugriffsstrukturen* erreicht. Äußere Zugriffsstrukturen sind diejenigen Komponenten der Wörterbuchform, die sicherstellen, dass ein Wörterbuch *akzessive Wörterbucheinträge* und damit *akzessive Daten* aufweist.[3] Akzessive Wörterbucheinträge, die im Englischen *basic reference units* heißen, sind die Grundbausteine jedes Nachschlagewerkes; die Wörterbuchartikel sind die wichtigsten. Ein Ausschnitt aus einer Typologie von akzessiven Wörterbucheinträgen findet sich in Abb. 1. Um bereits hier einem möglichen Irrtum vorzubeugen, sei festgestellt: Nicht alle äußeren Zugriffsstrukturen sind solche, auf die man von außerhalb eines Wörterbuchs zugreifen kann. Das Kriterium dafür, dass eine Zugriffsstruktur eine *äußere* Zugriffsstruktur ist, besteht mithin nicht darin, dass man von außerhalb des Wörterbuchs kommend zugreifen kann.

Die interne Datenakzessivität wird durch die Einrichtung von *inneren Zugriffsstrukturen* erreicht. Innere Zugriffsstrukturen dienen dem kundigen Benutzer (*sensu* Wiegand 1998: 506) zur Orientierung innerhalb akzessiver Wörterbucheinträge, insonderheit innerhalb von kondensierten Wörterbuchartikeln. Äußere Zugriffsstrukturen sind für gedruckte Nachschlagewerke obligatorisch; innere Zugriffsstrukturen sind dagegen fakultativ. Äußere und nichtalphabetische innere Zugriffsstrukturen unterscheiden sich — wie wir noch genauer im 4. Abschnitt sehen werden — in ihren Zugriffseigenschaften grundsätzlich.
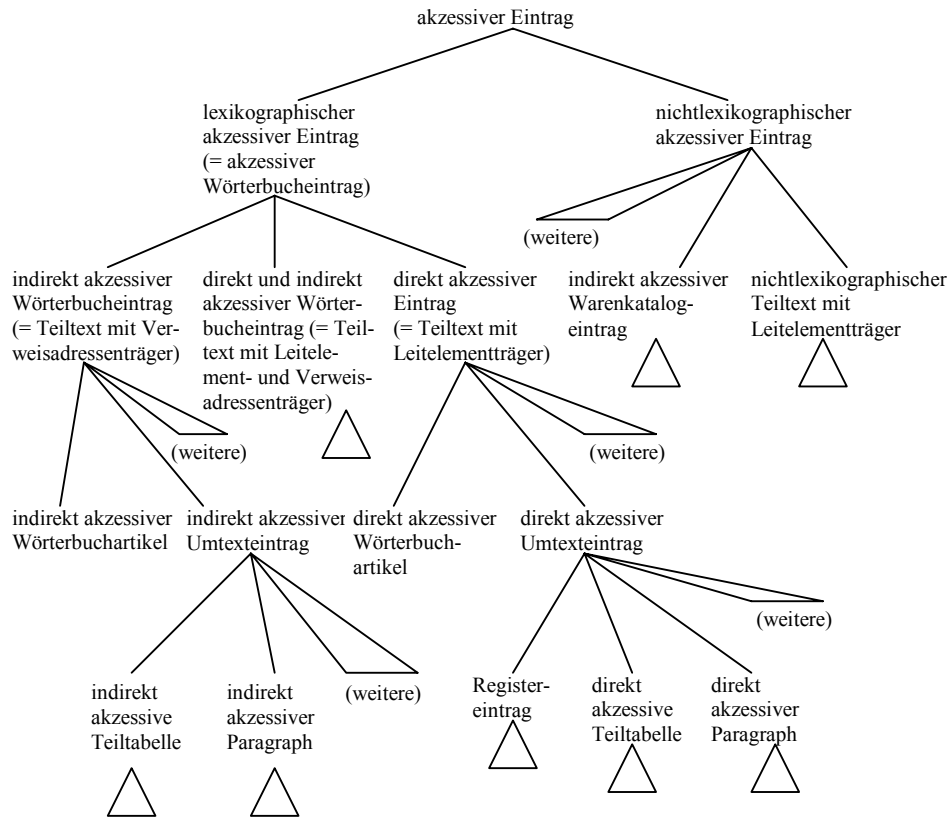
akzessiver Eintrag

lexikographischer
akzessiver Eintrag
(= akzessiver
Wörterbucheintrag)

nichtlexikographischer
akzessiver Eintrag

(weitere)

indirekt akzessiver
Wörterbucheintrag
(= Teiltext mit Ver-
weisadressenträger)

direkt und indirekt
akzessiver Wörter-
bucheintrag (= Teil-
text mit Leitele-
ment- und Verweis-
adressenträger)

direkt akzessiver
Eintrag
(= Teiltext mit
Leitelementträger)

indirekt akzessiver
Warenkatalog-
eintrag

nichtlexikographischer
Teiltext mit
Leitelementträger

(weitere)

indirekt akzessiver
Wörterbuchartikel

indirekt akzessiver
Umtexteintrag

direkt akzessiver
Wörterbuch-
artikel

direkt akzessiver
Umtexteintrag

(weitere)

(weitere)

indirekt
akzessive
Teiltabelle

indirekt
akzessiver
Paragraph

(weitere)

Register-
eintrag

direkt
akzessive
Teiltabelle

direkt
akzessiver
Paragraph

Abb. 1: Ausschnitt aus einer Typologie akzessiver Einträge; vgl. Wiegand 2003: 196.

3.   Externe Datenakzessivität

Wir betrachten zunächst die externe Datenakzessivität für den einfachsten Fall, und zwar bei gedruckten Sprachwörterbüchern, deren Wörterbuchgegenstands-bereich (*sensu* Wiegand 1998: 303) solche Sprachen sind, die nach einem Alpha-bet verschriftet sind. Als Beispiel für einen Wörterbuchgegenstandsbereich nehmen wir die deutsche Standardsprache der Gegenwart. Zu deren Schrift-system gehört das lateinische Alphabet in einer bestimmten Ausprägung: Es weist 26 Buchstaben auf. Unter dem Gesichtspunkt der Datenakzessivität ist dann der einfachste Fall ein glattalphabetischer monoakzessiver Einbänder der deutschen Standardsprache, der ausschließlich zum Nachschlagen von außen konzipiert ist, so dass es in ihm auch keine Verweisungen gibt.[4]

3.1   Erster Einblick: Alphabet und alphabetische äußere Zugriffsstrukturen

Will man verstehen, was eine *alphabetische äußere Zugriffsstruktur* ist, muss man als Verstehensvoraussetzung einige Eigenschaften des Alphabets kennen. Hier ist zunächst auf einen durchaus störenden unterschiedlichen Gebrauch der Wörter *alphabetisch* und *Alphabet* in verschiedenen wissenschaftlichen Disziplinen kurz einzugehen. In der Theorie der formalen Sprachen wird mit *Alphabet* eine nichtleere Menge von Zeichen bezeichnet, die nicht geordnet sein muss (vgl. Wiegand 1989: 374). Dies bedeutet: Die wohlgeformten Ketten über diesem Alphabet können nicht alphabetfundiert geordnet werden. Dieser Gebrauch von *Alphabet* ist für die Lexikographie ungeeignet. In der Schriftlinguistik ist der Gebrauch unterschiedlich. Die meisten Autoren verstehen unter einem Alphabet eine endliche Menge von Schriftzeichen, die eine historisch entstandene kanonische Abfolge aufweisen, so dass ein Alphabet eine endliche geordnete Menge ist. Diese Auffassung ist für die Lexikographie und Wörterbuchforschung dann brauchbar, wenn man sie präzisiert. Relativ zu diesem schriftlinguistischen Gebrauch heißt *alphabetisch* dann, dass eine bestimmte endliche Menge von sprachlichen Ausdrücken gemäß einem Alphabet geordnet ist. Wir werden noch genauer sehen, wie ungenau eine solche Redeweise ist. Als Hinweis zu dieser Feststellung an dieser Stelle nur dies: Man versuche einmal die deutschen Wörter *Märchen*, *Öl*, *müde* und *heiß* nach dem lat. Alphabet zu ordnen. Andere Autoren sprechen immer dann von der *Alphabetreihe*, wenn es nicht um den Verschriftungsaspekt, sondern um den Ordnungsaspekt eines Alphabets geht.[5] Obwohl gerade dieser Terminus auf einen für die Lexikographie wesentlichen Aspekt jedes Alphabets aufmerksam macht, den wir noch genauer kennen lernen werden, werde ich beim Terminus *Alphabet* bleiben. In der lexikographischen Werkstattsprache wird *Alphabet* jargonhaft verwendet; werkstattsprachlich kann man z.B. formulieren: „Einige Ausdrücke sind falsch ins Alphabet eingeordnet." Gegen solche handlungsbegleitenden und/oder handlungskommentierenden werkstattsprachlichen Redewesen ist nichts einzuwenden, wenn sie bei den Beteiligten verständlich sind. In Umtexten von Wörterbüchern ist dieser Gebrauch von *Alphabet* jedoch fehl am Platz, auch wenn er sich in den Metatexten deutscher Wörterbücher öfters findet; statt vieler gebe ich nur ein Beispiel. In Mackensen (1977: V) heißt es: „Die geographischen Benennungen wurden in das Alphabet ebenso eingearbeitet wie die [...] Abbildungen." In allen lexikographietheoretischen Zusammenhängen ist dieser oder jeder andere werkstattsprachliche Gebrauch von *Alphabet* selbstverständlich unangemessen.

Im Folgenden gehe ich kurz auf Alphabete am Beispiel des lateinischen Alphabets des neuhochdeutschen Schriftsystems ein. Jedes Alphabet lässt sich unter zwei Aspekten betrachten, nämlich dem Verschriftungs- und dem Ordnungsaspekt. Anders ausgedrückt heißt das: Jedes Alphabet ist bifunktional: Es dient der Verschriftung einer gesprochenen Sprache und hat damit einen Bezug zu Lauten und Phonemen. Weiterhin ist es ein Mittel zur Herstellung von Ord-

nung(en). Welcher der beiden Aspekte historisch der ältere ist, ist unter den „Schriftgelehrten" strittig und interessiert hier nicht. Klar ist allerdings Folgendes: Die Universalität des Alphabets besteht darin, dass es als Mittel zur Herstellung von Ordnungen gerade dann am besten funktioniert, wenn man von jedem Laut-Buchstaben-Bezug vollständig abstrahiert. Für den Ordnungsaspekt des lat. Alphabets ist vielmehr nur von Relevanz, dass es sich um eine *Ordnungsstruktur* im mathematischen Sinne handelt, welche die Axiome einer totalen (oder: linearen) Ordnung erfüllt, so dass Isomorphien zur Zahlenreihe definiert werden können. Dies ist — sehr vereinfacht und ausdrücklich für Leser ohne mathematische Vorkenntnisse dargestellt — folgendermaßen zu verstehen:

Betrachten wir einmal ein lat. Alphabet in seiner üblichen Schreibweise im Vergleich mit den ersten 26 Zahlen der Zahlenreihe:

$$1 < 2 < 3 < 4 < 5 < 6 < 7$$
a   b   c   d   e   f   g

$$8 < 9 < 10 < 11 < 12 < 13 < 14$$
h   i   j   k   l   m   n

$$15 < 16 < 17 < 18 < 19 < 20 < 21$$
o   p   q   r   s   t   u

$$22 < 23 < 24 < 25 < 26$$
v   w   x   y   z

Dann kann man z.B. das Folgende wahrnehmen, nämlich direkt sehen (!!):

a geht b voraus; dem entspricht 1 < 2 (zu lesen wie *1 ist kleiner als 2*),

a geht c voraus; dem entspricht 1 < 3,

a geht d voraus; dem entspricht 1 < 4,

a geht e voraus; dem entspricht 1 < 5 usw. bis

a geht z voraus, was 1 < 26 entspricht.

Oder man sieht z.B.: b geht c voraus, c geht d voraus usw. sowie die dazugehörigen Entsprechungen in der Zahlenreihe.

Das Alphabet in seiner kanonischen Schreibweise ist daher nichts weiter als eine geniale verkürzte, auf die Anschauung rekurrierende Präsentation einer Menge von 26 Buchstaben, auf der eine zweistellige Ordnungsrelation vom Typ der Vorgänger-Nachfolger-Relation (oder: der Präzedenzrelation) definiert ist, zu welcher der Relationsterm *x geht y voraus* oder der als synonym geltende Relationsterm *x vor y* gehört. Eine solche Relation lässt sich (in extensionaler Schreibweise) auch als Menge M von geordneten Paaren hinschreiben; sehr stark abgekürzt wäre die Schreibweise wie folgt:

M = {‹a, b›, ‹a, c›, ‹a, d›, ‹a, e›, ‹a, f›, ..., ‹y, z›}.

Bei einer vollständigen Präsentation in dieser Mengenschreibweise müssten in der Schweifklammer genau n · (n-1) : 2 geordnete Paare stehen, mit n = 26, also (26 · 25) : 2 = 325 geordnete Paare. Für alle 325 geordnete Paare der Relation geht der Relationsterm *x geht y voraus* immer dann in einen wahren Satz über, wenn für „x" die erste und für „y" die zweite Komponente eines geordneten Paares eingesetzt wird. Für das Paar (a, b) ergibt sich dann: a geht b voraus. Oder für das Paar (b, z) ergibt sich: b geht z voraus.

Eine auf einer Menge M definierte zweistellige Ordnungsrelation R_P, die eine Teilmenge des kartesischen Produkts M x M ist (also $R_P \subseteq$ M x M) vom Typ der Präzedenzrelation hat folgende Eigenschaften, die als Ordnungsaxiome fungieren: R_P ist (1) transitiv, (2) asymmetrisch und (3) konnex.

Für die Präzedenzrelation, die auf der Menge der 26 Buchstaben des lat. Alphabets definiert ist, kann dies informell folgendermaßen erläutert werden (mit B_x, B_y und B_z als Variablen für Buchstaben des lat. Alphabets):

(a)   *Transitivität*: Für drei verschiedene Buchstaben des Alphabets B_x, B_y, B_z gilt stets: Wenn B_x B_y vorausgeht und wenn B_x B_z vorausgeht, dann geht auch B_x B_z voraus.
Beispiel: Wenn a b vorausgeht und wenn b z vorausgeht, dann geht auch a z voraus. Dem entspricht: Wenn 1 kleiner ist als 2 und wenn 2 kleiner ist als 26, dann ist auch 1 kleiner als 26.

(b)   *Asymmetrie*: Für zwei verschiedene Buchstaben des Alphabets B_x, B_y gilt stets: Wenn B_x B_y vorausgeht, dann geht B_y nicht B_x voraus.
Beispiel: Wenn b d vorausgeht, dann geht d nicht b voraus. Dem entspricht: Wenn 2 kleiner ist als 4, dann ist 4 nicht kleiner als 2.

(c)   *Konnexität*: Für die zwei geordneten Paare von 2 Buchstaben des Alphabets B_x, B_y gilt stets: Entweder ist ‹B_x, B_y› ein Element der alphabetischen Relation oder ‹B_y, B_x›. Oder anders ausgedrückt: Zwei beliebige Buchstaben B_x, B_y des Alphabets sind hinsichtlich der alphabetischen Relation vergleichbar.
Beispiel: Nimmt man die Buchstaben a und q und bildet die beiden möglichen Paare ‹a, q› und ‹q, a›, dann ist nur ‹a, q› ein Element der Relation, da gilt: a geht q voraus. Dem entspricht: Nimmt man die geordneten Paare ‹1, 16› und ‹16, 1›, dann ist nur das erstgenannte Paar ein Element der *Kleiner-als*-Relation.

Es ist klar, dass die soeben grob erläuterte Isomorphie eines Ausschnitts der Zahlenreihe mit dem lat. Alphabet die Grundlage dafür ist, dass man auf der Basis des Alphabets lineare Ordnungen herstellen kann. Das bedeutet aber nicht, dass Personen, wenn sie eine Menge von sprachlichen Ausdrücken alphabetisch sortieren, bewusst auf diese Isomorphie zurückgreifen. Günther (1996: 1568f) schreibt über das alphabetische Sortieren:

> Die Technik alphabetischen Sortierens und des Auffindens von Information in alphabetisch sortierten Listen ist den meisten literaten Menschen so vertraut, daß es gar nicht so einfach ist, sich klarzumachen, was dieser Technik und Tätigkeit zugrundeliegt; wissenschaftliche Literatur ist dazu kaum zu finden. Worin besteht eine vollständige alphabetische Anordnung, was macht man beim vollständigen, absoluten oder mechanischen alphabetischen sortieren (Wiegand 1989)? Die Elemente einer Liste, z.B. die Lemmata eines Wörterbuchs, werden beim absoluten Alphabetisieren analysiert als den jeweiligen Buchstabenketten zugeordnete Zahlenwerte. Dabei wird jeder Buchstabenposition des Wortes diejenige Ziffer zugeordnet, die der Stelle des Buchstabens in der Alphabetreihe entspricht.

Zwar ist es richtig, dass die Grundlagen des Alphabetisierens mathematisch expliziert werden können. Beim normalen alphabetischen Sortieren und beim externen Zugreifen auf alphabetische äußere Zugriffsstrukturen, bei dem ja das alphabetische Sortieren nachvollzogen wird, also insgesamt in allen praktischen Handlungszusammenhängen, in denen alphabetisiert wird, spielen die möglichen Zuordnungen von Buchstaben zu Zahlenwerten überhaupt keine Rolle. Günther verwechselt die extrasituative bewusstseinsfremde Perspektive eines Mathematikers bzw. eines Wissenschaftlers mit der intrasituativen bewusstseinseigenen Perspektive derjenigen, die alphabetisch sortieren und alphabetfundiert extern zugreifen möchten. Vielmehr gilt: In beiden Handlungssituationen greifen die Agenten lediglich in intuitiver Weise auf die erläuterten Eigenschaften der Ordnungsstruktur zurück. Entsprechend lässt sich das Alphabetisieren am einfachen Beispiel wie folgt erläutern: Die drei Wortformen *Mona*, *Monist* und *Monster* seien alphabetisch zu sortieren. Dann werden diese drei Wortformen Buchstabenposition für Buchstabenposition verglichen: Das Ergebnis ist: Die drei Wortformen sind in den ersten drei Buchstabenpositionen gleich (*Mon*). Also bestimmen die Buchstaben in der vierten Buchstabenposition ihre alphabetische Reihenfolge, und zwar gilt u.a.: Im Alphabet geht a i voraus; also kommt *Mona* vor *Monist*. Weiterhin gilt: Im Alphabet geht i s voraus; entsprechend kommt *Monist* vor *Monster*, so dass die lineare Ordnungsstruktur, die als Ergebnis des alphabetischen Sortierens zustande kommt, folgende Form aufweist: *Mona < Monist < Monster* (hier mit „<" für *geht voraus*). Beim alphabetischen Sortieren ebenso wie beim alphabetfundierten externen Zugreifen auf alphabetische Zugriffsstrukturen benötigt man also keine Zahlen, sondern die Kenntnis des Alphabets und damit die intuitive Kenntnis der Ordnungsaxiome, die für das Alphabet als lineare Ordnungsstruktur gelten, sowie eine gewisse Übung, so dass man die allgemeine Wörterbuchbenutzungspraxis (*sensu* Wiegand 1998: 370) beherrscht.

Wir kennen jetzt diejenigen ordnungstheoretischen Eigenschaften eines Alphabets, deren Kenntnis die Voraussetzung dafür ist, dass man genau versteht, was die zentralen Eigenschaften einer *alphabetischen äußeren Zugriffsstruktur* in ordnungstheoretischer Hinsicht sind. Für eine alphabetische äußere Zugriffsstruktur gelten die gleichen Ordnungsaxiome wie für das Alphabet, rela-

tiv zu dem sie erzeugt wurde. Entsprechend ist eine alphabetische äußere Zugriffsstruktur eine Menge, sie heiße M_z, von *sprachlichen äußeren Zugriffstextelementen*, auf der eine zweistellige alphabetische (a) Ordnungsrelation $R_a \subseteq M_z$ x M_z definiert ist, die transitiv, asymmetrisch und konnex ist. Eine alphabetische äußere Zugriffsstruktur ist mithin eine lineare Ordnungsstruktur. Dadurch ist es möglich, dass ein Benutzer, der die nötigen Kenntnisse hat, auf jedes sprachliche äußere Zugriffstextelement direkt extern zugreifen kann. Direkt extern auf eine alphabetische äußere Zugriffsstruktur zugreifen zu können, heißt dann — informell und vereinfacht ausgedrückt — anhand der Buchstabenfolge eines sprachlichen Ausdrucks, die *Leitelement* heißt (vgl. Abb. 3), z.B. anhand von |Mona|, auf der Basis der Kenntnis einer *wörterbuchexternen* linearen Ordnungsstruktur, nämlich der Kenntnis des Alphabets, ein alphabetisch eingeordnetes äußeres Zugriffstextelement, z.B. das Lemma **Mona**, in einer *wörterbuchinternen* linearen Ordnungsstruktur, welche die gleichen Ordnungsaxiome erfüllt wie das Alphabet, Ausdrücke gezielt suchen und finden zu können bzw. sicher entscheiden zu können, dass ein äußeres Zugriffstextelement, das die gleiche Buchstabenfolge aufweist wie das Leitelement, kein Element der sprachlichen äußeren Zugriffsstruktur ist und daher nicht gefunden werden kann, weil eine Lemmalücke eines bestimmten Typs vorliegt.
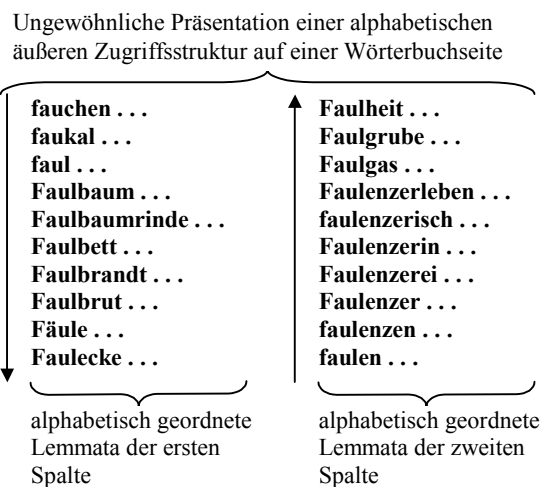
Ungewöhnliche Präsentation einer alphabetischen
äußeren Zugriffsstruktur auf einer Wörterbuchseite

fauchen . . .	**Faulheit** . . .
faukal . . .	**Faulgrube** . . .
faul . . .	**Faulgas** . . .
Faulbaum . . .	**Faulenzerleben** . . .
Faulbaumrinde . . .	**faulenzerisch** . . .
Faulbett . . .	**Faulenzerin** . . .
Faulbrandt . . .	**Faulenzerei** . . .
Faulbrut . . .	**Faulenzer** . . .
Fäule . . .	**faulenzen** . . .
Faulecke . . .	**faulen** . . .

alphabetisch geordnete   alphabetisch geordnete
Lemmata der ersten   Lemmata der zweiten
Spalte   Spalte

Abb. 2:   Veranschaulichung zu einer ungewöhnlichen Präsentation einer alphabetischen äußeren Zugriffsstruktur im zweidimensionalen Druckraum

Bisher haben wir uns mit dem Ordnungsaspekt des Alphabets und der alphabetischen äußeren Zugriffsstrukturen befasst. Zugriffsstrukturen jeder Art müssen jedoch im zweidimensionalen Printraum auch präsentiert werden. Diese Präsentation richtet sich in der Regel nach Konventionen des jeweiligen Schriftsystems (vgl. zum Folgenden Wiegand 1989: 381). Für das Deutsche

heißt dies, dass für ein normales Schriftstück oder Druckwerk gilt: Es wird auf (oder: in) Zeilen von links nach rechts geschrieben und auch so gelesen. Werden mehr als eine Zeile benötigt, dann wird nacheinander auf jeder Zeile von links nach rechts weiter geschrieben, und auf diese Weise wird dann auch von oben nach unten geschrieben. Auf der nächsten Seite geht es dann genau so weiter. Die Präsentation von alphabetischen Zugriffsstrukturen richtet sich nach diesen Konventionen. Das erkennt man sozusagen „auf einen Blick", wenn man die Abb. 2 genauer betrachtet. In dieser ist in der ersten Spalte diese Konvention eingehalten. Die alphabetisch sortierte Lemmareihe läuft von oben nach unten. In der zweiten Spalte läuft sie dagegen von unten nach oben. Niemand wird behaupten wollen, die Lemmata seien deswegen insgesamt nicht alphabetisch sortiert, und selbstverständlich liegt auch eine alphabetische äußere Zugriffsstruktur vor, diese ist lediglich im zweidimensionalen Printraum auf eine Weise präsentiert, die nach dem deutschen Schriftsystem nicht vorgesehen ist.

Das Beispiel in Abb. 2 lehrt: (1) Die Ordnungseigenschaften einer alphabetischen äußeren Zugriffsstruktur sind von ihrer zweidimensionalen Präsentation unabhängig. (2) Da es im Rahmen der Schriftsystemkonventionen eine Reihe von Möglichkeiten gibt, die äußeren Zugriffstextelemente zu präsentieren, lassen sich alphabetische äußere Zugriffsstrukturen auch nach der Art ihrer Präsentation und damit nach den Lageeigenschaften der äußeren Zugriffstextelemente im zweidimensionalen Druckraum unterscheiden. Dies kann hier nur an einem Beispiel anhand der Abb. 3 erläutert werden.

In der Abb. 3 wird u.a. veranschaulicht, dass auf der (nur ausschnittsweise dargestellten) Menge aller äußeren Zugriffstextelemente, die im vorliegenden Fall Lemmata sind, nicht nur eine zweistellige Ordnungsrelation vom Typ der alphabetischen Ordnungsrelation definiert ist (auf deren Elemente der Relationsterm *x geht y alphabetisch voraus* zutrifft), sondern auch eine zweistellige texttopologische (oder: textarchitektonische) Ordnungsrelation, nämlich eine vom Typ der *oberhalb*-Relation (auf deren Elemente der Relationsterm *x ist (senkrecht) oberhalb von y* zutrifft). Auch Relationen dieses Typs sind totale Ordnungsrelationen, die transitiv, asymmetrisch und konnex sind. Es gilt jetzt — am Beispiel zweier Lemmata aus Abb. 3 erläutert — nicht nur „**Mona** geht alphabetisch **Monist** voraus", sondern auch „**Mona** ist senkrecht oberhalb (kurz: oberhalb) von **Monist**". Ist M_L die Menge aller Lemmata (L) eines Wörterbuchs, dann gilt: Für alle Paare, die mit Lemmata L_x, L_y aus M_L gebildet werden können, entweder: Auf das Paar ‹L_x, L_y› trifft sowohl der Relationsterm *L_x geht alphabetisch L_y voraus* zu, als auch der Relationsterm *L_x ist oberhalb von L_y*; oder es gilt: Auf das Paar ‹L_y, L_x› trifft sowohl der Relationsterm *L_y geht alphabetisch L_x voraus* zu, als auch der Relationsterm *L_y ist oberhalb von L_x*. Oder anders ausgedrückt: Zwei beliebige Lemmata L_x, L_y der Menge M_L sind immer sowohl hinsichtlich der alphabetischen Relation als auch hinsichtlich der *oberhalb*-Relation vergleichbar. Nehmen wir als Beispiel: L_x = **Mona** und L_y = **Monist**, dann ist offenbar das geordnete Paar ‹**Mona**, **Monist**› und nicht das geordnete Paar ‹**Monist**, **Mona**› ein Element beider zweistelligen Ordnungsrelationen.

Zugriffsbereich der Zugriffsstruktur
(hier: das Wörterverzeichnis)

Ausschnitt aus der
glattalphabetischen
makrostrukturellen
äußeren Zugriffs-
struktur

A
⋮
B
⋮
M

Mona

Monist

4 glattalpha-
betisch
geordnete
Zugriffstext-
elemente, und
zwar
Leitelement-
träger, die
Lemmata sind

4 akzessive Wörter-
bucheinträge, die
Wörterbuchartikel
sind, auf die das
Lemma direkten
externen Zugriff
eröffnet

Monster

Montage

EXTERNE
ZUGRIFFS-
HANDLUNG

N
⋮
Z

⌐Montage⌐ = Leitele-
ment (L)

Montage = deutsches
Wort

Verschriftung

Lat. Alphabet im deutschen Schriftsystem
(wörterbuchexterne lineare Ordnungsstruktur)

a, b, c, d, e, f, g,
h, i, j, k, l, m, n,
o, p, q, r, s, t, u, v,
w, x, y, z

Abb. 3:   Veranschaulichung zur externen Datenakzessivität, die mittels einer
glattalphabetischen makrostrukturellen äußeren Zugriffsstruktur her-
gestellt ist. Der Doppelpfeil bedeutet soviel wie *x ist oberhalb von y*,
mit „x" und „y" als Variablen für äußere Zugriffstextelemente

Wir können nun also feststellen: Ist auf der Menge aller Lemmata eines Wörter-
buchs sowohl eine zweistellige Ordnungsrelation vom Typ der alphabetischen

als auch eine vom Typ der *oberhalb*-Relation definiert, dann liegt eine *glattalpha-betische makrostrukturelle äußere Zugriffsstruktur* vor.

Ist auf der Menge aller Lemmata keine *oberhalb*-Relation definiert, dann liegt eine *nicht glattalphabetische makrostrukturelle äußere Zugriffsstruktur* vor. Zugriffsstrukturen dieses Typs, die nicht zu einem seiner speziellen Untertypen gehören — wie z.B. die nischenalphabetischen makrostrukturellen Zugriffsstrukturen — kommen kaum vor, weil sie extrem benutzerunfreundlich sind. Bei den alphabetischen *makrostrukturellen* äußeren Zugriffsstrukturen lässt sich die für die Typen von Makrostrukturen entwickelte Terminologie (vgl. Wiegand 1989; 1998a; 2002; Bergenholtz, Tarp und Wiegand 1999: 1811ff) weitgehend übernehmen. Dies sei wenigstens an einem Beispiel demonstriert. In Abb. 4 findet sich ein Ausschnitt aus dem Wörterverzeichnis des Duden-²GW.

> **Wạll|holz,** das; -es, Wallhölzer [zu mund-artl. walen = (sich wälzen] (schweiz.): *Nudelholz.*
> **Wạl|lis,** das; -: Schweizer Kanton; ¹**Wạl|li-ser,** der; -s, -: Ew.; ²**Wạl|li|ser** ⟨indekl. Adj.⟩; **Wạl|li|ser Al|pen** ⟨nur mit Art.; Pl.⟩: südlich des Walliser Rhonetals gelegener Teil der Schweizer Alpen; **Wạl|li-se|rin,** die; -, -nen: w. Form zu ↑ ¹Walliser; **wạl|li|se|risch** ⟨Adj.⟩.
> **Wal|lo|ne,** der; -n, -n: Nachkomme romanisierter Kelten in Belgien u. Nordfrankreich; **Wal|lo|ni|en:** *Gebiet, in dem Wallonisch gesprochen wird;* **Wal|lo|nin,** die; --nen: w. Form zu ↑ Wallone; **wal|lo|nisch** ⟨Adj.⟩; **Wal|lo|nisch,** das; -[s] u. ⟨nur mit best. Art.:⟩ **Wal|lo|ni|sche,** das; -n: *von den Wallonen gesprochenes Französisch.*
> **Wall|street** ['wɔ:lstri:t], die; -: Geschäftsstraße in New York (Bankzentrum); Geld- u. Kapitalmarkt, Finanzzentrum der USA.

Abb. 4:   Ausschnitt aus dem Wörterverzeichnis des Duden-²GW

Der Ausschnitt in Abb. 4 beginnt mit einem Einzelartikel zum Lemmazeichen *Wallholz.* Es folgen zwei gruppierte Artikelnischen. Die erste Nische beginnt mit dem Nischeneingangslemma **Wal lis**; die zugehörige unten und binnenerweiterte Lemmazeichengestaltangabe ist **Wạl|lis**. Die zweite Nische beginnt mit dem Nischeneingangslemma **Wal lo ne**; die zugehörige unten und binnenerweiterte Lemmazeichengestaltangabe ist **Wal|lo|ne**. In der ersten gruppierten Artikelnische stehen fünf, in der zweiten stehen ebenfalls fünf gruppierte Nischenlemmata. Auf einer Teilmenge der Menge aller Lemmata des Duden-²GW, nämlich der Menge derjenigen Lemmata, die keine gruppierten Nischenlemmata sind, ist eine *oberhalb*-Relation definiert; entsprechend liegt eine *komprimiert-nischenalphabetische makrostrukturelle Zugriffsstruktur* vor. Eine Veranschaulichung zu diesem Zugriffsstrukturtyp findet sich in der Abb. 5.
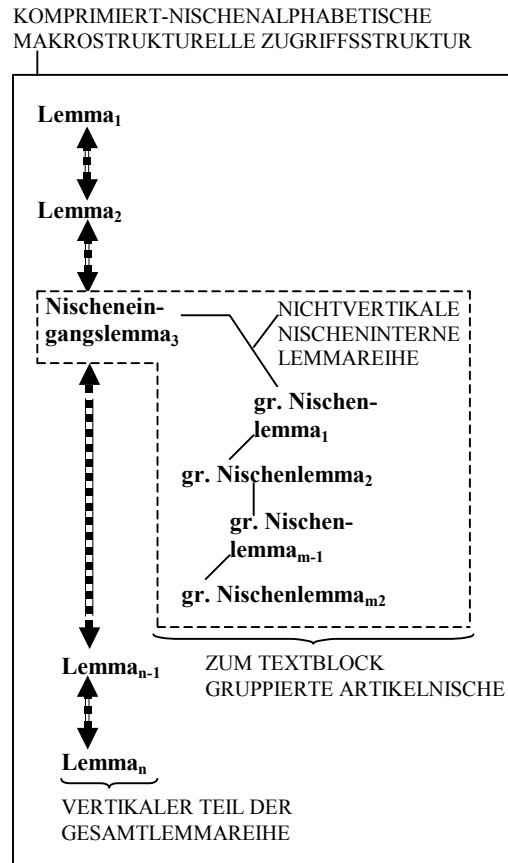
KOMPRIMIERT-NISCHENALPHABETISCHE
MAKROSTRUKTURELLE ZUGRIFFSSTRUKTUR

Lemma₁

Lemma₂

**Nischenein-
gangslemma₃**

NICHTVERTIKALE
NISCHENINTERNE
LEMMAREIHE

**gr. Nischen-
lemma₁**

gr. Nischenlemma₂

**gr. Nischen-
lemma_{m-1}**

gr. Nischenlemma_{m2}

ZUM TEXTBLOCK
GRUPPIERTE ARTIKELNISCHE

Lemma_{n-1}

Lemma_n

VERTIKALER TEIL DER
GESAMTLEMMAREIHE

Abb. 5:   Veranschaulichung zur komprimiert-nischenalphabetischen makro-
strukturellen Zugriffsstruktur

Es ist oben schon darauf hingewiesen worden, dass das lat. Alphabet nicht aus-
reicht, um alle in deutschen Texten vorkommende Wörter so alphabetisch zu
sortieren, dass eine lineare Ordnungsstruktur gegeben ist. Denn in den Leit-
elementen, die zu deutschen Wörtern gehören, können neben den 26 alphabet-
internen Buchstaben weitere alphabetexterne, aber schriftsysteminterne Buch-
staben auftreten, nämlich die deutschen Zusatzbuchstaben ä, ö, ü und ß sowie
darüberhinaus schriftsystemexterne Buchstaben. Weiterhin können numerische
Elemente auftreten, das Leerzeichen muss — je nach gewählter Alphabetisie-
rungsmethode — in verschiedener Weise berücksichtigt werden u.v.a.m. Das
lat. Alphabet zusammen mit der Menge aller Zusatzregelungen ergibt dann
das je *wörterbuchspezifische Zugriffsalphabet*, das in einem der Metatexte erklärt
werden muss (vgl. Wiegand 1989: 376ff).

Entsprechend lassen sich dann der *uneingeschränkte Zugriff* von dem *eingeschränkten Zugriff* auf alphabetische Zugriffsstrukturen unterscheiden. Ein uneingeschränkter Zugriff im Zuge der Ausführung einer externen Zugriffshandlung (*sensu* Wiegand 1998: 393ff) ist möglich, wenn das Leitelement nur aus alphabetinternen Buchstaben besteht. Besteht dagegen das Leitelement nicht nur aus alphabetinternen Buchstaben, wie dies z.B. bei folgenden Leitelementen der Fall ist: |öde|, |küssen|, |Bär|, |weiß| oder soll auf Bezeichnungen, wie z.B. *α-Strahlung* oder *8-Azaguanin* zugegriffen werden, bei denen nicht ohne weiteres klar ist, was zum Leitelement gehört, dann ist nur ein eingeschränkter Zugriff möglich. Die Einschränkung besteht darin, dass die Kenntnis des lat. Alphabets nicht ausreicht, so dass der Benutzer nicht extern zugreifen kann, ohne über ein Zusatzwissen zum wörterbuchspezifischen Zugriffsalphabet zu verfügen, das anhand der Metatexte des Wörterbuchs erworben wurde.

Damit sei der erste Einblick, der uns zu dem im Schriftkulturraum der Alphabetschriften bekanntesten Typ der äußeren Zugriffsstrukturen führte, nämlich zum Typ der alphabetischen Zugriffsstrukturen, abgebrochen.

3.2   Zweiter Einblick: Weitere Typen von äußeren Zugriffsstrukturen

Im Folgenden betrachten wir weitere Typen von äußeren Zugriffsstrukturen, und zwar solche, die nicht — wie die glattalphabetischen makrostrukturellen und die komprimiert-nischenalphabetischen makrostrukturellen Zugriffsstrukturen — zu einem Untertyp des Typs der alphabetischen makrostrukturellen äußeren Zugriffsstrukturen gehören. Dazu gehen wir von dem einfachsten Fall, nämlich den monoakzessiven Einbändern ohne Verweisungen, zu einem komplexeren Fall über: Wir berücksichtigen jetzt Wörterbücher, die integrierte Umtexte aufweisen und in denen Verweisungen anzutreffen sind.

Um zu verstehen, wie Zugriffsstrukturen funktionieren, die nicht zu den alphabetischen makrostrukturellen äußeren Zugriffsstrukturen gehören und weiterhin für das Verständnis der Bifunktionalität von alphabetischen makrostrukturellen äußeren Zugriffsstrukturen in solchen Wörterbüchern, in denen lemmatisch orientierte Verweisungen vorgenommen wurden, muss zunächst in vereinfachter Form erläutert werden, was unter einer externen Zugriffshandlung und unter einem externen Zugriff zu verstehen ist. Selbst wenn eine solche Erläuterung hier nur sehr grob ausfallen kann, lässt sie sich nicht ohne einen Rückgriff auf handlungstheoretische Grundeinsichten geben (vgl. Wiegand 1998: 270-292). Zunächst sollte das Folgende klar sein: Eine Handlung h_1, die zu einem bestimmten Handlungstyp HT_1 gehört, wird stets *dadurch* ausgeführt, *dass* mindestens eine Handlung h_2 ausgeführt wird, die zu einem bestimmten anderen Handlungstyp HT_2 gehört. Hierzu ein Beispiel aus der Grundschule: Eine Handlung h_3 vom Typ EINEN BRUCH DURCH EINEN ANDEREN BRUCH DIVIDIEREN (= HT_3) wird *dadurch* ausgeführt, *dass* eine Handlung h_4 vom Typ EINEN BRUCH MIT DEM KEHRWERT DES ANDE-

REN BRUCHES MULTIPLIZIEREN (= HT$_4$) ausgeführt wird. Man schreibt hierfür auf der generischen Ebene: {HT$_3$} → {HT$_4$}; hier repräsentiert der Rechtspfeil „→" die *dadurch-dass*-Relation. Handlungstypen (und damit die zugehörigen individuellen Handlungen h$_i$), die unmittelbar rechts vom Rechtspfeil genannt werden (im Beispiel also HT$_4$) liegen im *Ausführungsbereich* des Handlungstyps, der unmittelbar vor dem Rechtspfeil genannt wird (im Beispiel HT$_3$). Die *dadurch-dass*-Relation ist transitiv; entsprechend können mehrere Rechtspfeile hintereinander auftreten, so dass ein unmittelbarer und mindestens ein mittelbarer Ausführungsbereich gegeben sein können. Für den vorliegenden Zusammenhang ist es nun besonders wichtig zu berücksichtigen, dass externe Zugriffshandlungen im unmittelbaren Ausführungsbereich von usuellen Benutzungshandlungen liegen können, die zu zwei unterschiedlichen Typen von Benutzungshandlungen gehören. Dagegen ist hier nicht wichtig, wie externe Zugriffshandlungen genau ausgeführt werden; d.h.: Die Handlungstypen (und die zugehörigen Handlungen), die im Ausführungsbereich von externen Zugriffshandlungen liegen, sind hier nicht von Interesse. Im Folgenden nehmen wir als Beispielwörterbuch Wahrig (1989). Dessen vollständig kondensierte Wörterbuchartikel (*sensu* Wiegand 2003: 207f) weisen innere Zugriffsstrukturen auf (zu diesen vgl. 4). Eine Nachschlagehandlung anhand von Wörterbüchern, die innere Zugriffsstrukturen aufweisen, kann dann vereinfacht wie folgt charakterisiert werden: Eine Nachschlagehandlung, die zum Handlungstyp IN EINEM WÖRTERBUCH ETWAS NACHSCHLAGEN (= HT-6) gehört, ist eine Benutzungshandlung, die dadurch ausgeführt wird, dass nach der erfolgreichen Ausführung einer Handlung vom Typ AUF EIN WÖRTERBUCH EXTERN ZUGREIFEN (= HT-10) eine Handlung vom Typ AUF DIE DATEN EINES AKZESSIVEN WÖRTERBUCHEINTRAGES INTERN ZUGREIFEN (= HT-7) ausgeführt wird ({HT-6} → {erst: HT-10, dann: HT-7}).

Hier steht also der Typ der externen Zugriffshandlung (HT-10) im unmittelbaren Ausführungsbereich des Typs der Nachschlagehandlung (HT-6). Wer also z.B. (vgl. Abb. 3) nicht weiß, was *Montage* bedeutet und deswegen im Wörterbuch nachschlagen möchte, der muss dies *dadurch* tun, *dass* er — ausgestattet mit dem Leitelement |Montage| — zuerst eine externe Zugriffshandlung erfolgreich ausführt, so dass er das Lemma **Montage** gefunden hat. Als externer Zugriff gilt der erfolgreiche Abschluss der Ausführung einer externen Zugriffshandlung, der gerade darin besteht, dass ein Benutzer das äußere Zugriffstextelement gefunden hat. Erst danach kann er eine interne Zugriffshandlung ausführen, die ihm zur Bedeutungsparaphrasenangabe im Wörterbuchartikel zu *Montage* führt.

Anders verhält es sich dagegen, wenn ein Benutzer im gleichen Wörterbuch im Wörterbuchartikel zu *Zaun* ein verweisvermittelndes Textsegment findet, z.B. die lemmatisch orientierte vollständige Verweisangabe „→ Montage", die aus der Verweisbeziehungsangabe „→" und der Angabe der lemmatischen Verweisaußenadresse „Montage" besteht (vgl. Abb. 6).
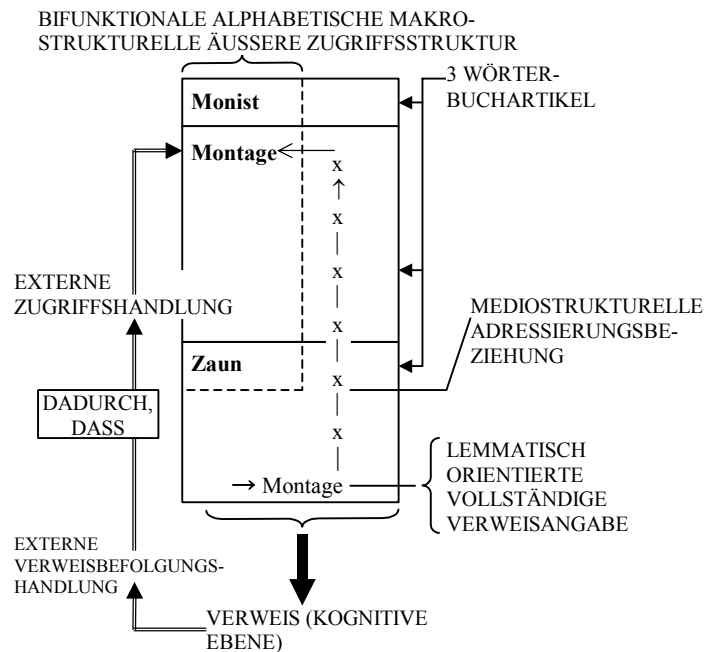
BIFUNKTIONALE ALPHABETISCHE MAKRO-
STRUKTURELLE ÄUSSERE ZUGRIFFSSTRUKTUR

3 WÖRTER-
BUCHARTIKEL

Monist

Montage

EXTERNE
ZUGRIFFSHANDLUNG

MEDIOSTRUKTURELLE
ADRESSIERUNGSBE-
ZIEHUNG

Zaun

DADURCH,
DASS

LEMMATISCH
ORIENTIERTE
VOLLSTÄNDIGE
VERWEISANGABE

→ Montage

EXTERNE
VERWEISBEFOLGUNGS-
HANDLUNG

VERWEIS (KOGNITIVE
EBENE)

Abb. 6:   Veranschaulichung zur Bifunktionalität alphabetischer makrostruktu-
reller äußerer Zugriffsstrukturen; A → B bedeutet soviel wie *anhand
von A wird B* erschlossen.

Dieser Benutzer kann dann anhand der Verweisangabe „→ Montage" einen
Verweis erschließen; dieser besteht in seinem Wissen, dass er, um das Verweis-
ziel zu erreichen, eine externe Verweisbefolgungshandlung ausführen muss,
und zwar mit Hilfe der im Artikel zu *Zaun* genannten Verweisadresse | Mon-
tage|, die damit als mediostrukturelles Leitelement fungiert. Die Ausführung
der externen Verweisbefolgungshandlung kann nur *dadurch* geschehen, *dass*
der Benutzer eine externe Zugriffshandlung ausführt, die ihn zum Lemma
Montage führt. Das Lemma ist damit nicht nur ein *Leitelementträger*, sondern
auch ein *Verweisadressenträger*. Damit sollte deutlich sein, dass es auch externe
Zugriffshandlungen gibt, die im unmittelbaren Ausführungsbereich von Ver-
weisbefolgungshandlungen liegen. Während bei externen Zugriffshandlungen,
die im Ausführungsbereich von Nachschlagehandlungen liegen, das zur Hand-
lungsausführung benötigte Leitelement *wörterbuchextern* erschlossen wird, wird
bei externen Zugriffshandlungen, die im Ausführungsbereich von Verweis-
befolgungshandlungen liegen, das zur Handlungsausführung benötigte Leitele-
ment, nämlich die jeweilige Verweisadresse, anhand von verweisvermittelnden
Textsegmenten und damit *wörterbuchintern* erschlossen. Aus diesem Grunde
heißen solche Leitelemente *mediostrukturelle Leitelemente.* In einem Wörterbuch,

in dem lemmatische Verweisaußenadressen, wie z.B. |Montage| mit verweis-vermittelnden Textsegmenten, wie z.B. der vollständigen Verweisangabe „→ Montage", genannt werden, fungiert die alphabetische makrostrukturelle äuß-ere Zugriffsstruktur immer auch als *mediostrukturelle äußere Zugriffsstruktur*, also als Fundortdomäne, in der die Verweisaußenadressen zu finden sind. Sie ist damit bifunktional. Dies ist in modernen Printwörterbüchern der Normal-fall, wurde aber bisher in der Wörterbuchforschung kaum explizit dargestellt.

Im Folgenden soll erläutert werden, was eine *mediostrukturelle äußere Zugriffsstruktur* ist. Dies geschieht anhand von Abb. 7.

Abb. 7: Veranschaulichung zu mediostrukturellen äußeren Zugriffsstrukturen

Die Abb. 7 zeigt ein Beispiel aus dem Wahrig (1989). Links ist ein Ausschnitt aus dem Wörterverzeichnis zu sehen, rechts ein Ausschnitt aus einem zweispaltigen tabellarischen Umtext, der sich im Vorspann des Wörterbuchs findet. Im Artikel zum Lemmazeichen *nachleben* findet sich nach der verdichteten Wortartangabe „V.", der Abkürzung für *Verb*, eine numerische reduzierte Verweisangabe, nämlich „600"; reduziert ist sie deswegen, weil sie keine Verweisbeziehungsangabe als Teilangabe aufweist. Die reduzierte Verweisangabe ist umtextorientiert: Der Adressenfundort der genannten Verweisadresse │600│ ist die numerische äußere Umtextzugriffsstruktur, von der der Ausschnitt ‹**600, 601, 602, 605, 607, 608**› in Abb. 7 zu sehen ist; sie bildet die erste Spalte des tabellarischen Umtextes, die damit eine Zugriffsspalte ist. Es besteht mithin eine mediostrukturelle Adressierungsbeziehung von der reduzierten Verweisangabe „600" im Artikel zu *nachleben* zum Verweisadressenträger „**600**" in der numerischen äußeren Umtextzugriffsstruktur. Auf letztere kann nur im Zuge der Ausführung von externen Zugriffshandlungen zugegriffen werden, die im unmittelbaren Ausführungsbereich von externen Verweisbefolgungshandlungen liegen. Zwar kann man in den tabellarischen Umtexten des Wahrig (1989) „herumlesen" und sich z.B. darüber beim Lesen informieren, wieviele und welche Typen von Satzmusterangaben im Wahrig (1989) vorgesehen sind. Bei einer solchen Wörterbuchbenutzung werden aber keine externen Zugriffshandlungen anhand eines nichtmediostrukturellen Leitelements ausgeführt, das wörterbuchextern gegeben ist. Für die numerische äußere Umtextzugriffsstruktur, von der in Abb. 7 ein Ausschnitt wiedergegeben ist, gilt Folgendes: Weil auf sie ausschließlich mit Hilfe eines mediostrukturellen Leitelements, das wörterbuchintern gegeben ist, im Zuge der Ausführung einer externen Zugriffshandlung zugegriffen werden kann, die im Ausführungsbereich einer externen Verweisbefolgungshandlung liegt, gehört sie zu einem bestimmten Typ der *mediostrukturellen äußeren Zugriffsstrukturen.*

Unter dem Aspekt der externen Datenakzessivität muss der Zugriffsbereich der äußeren Umtextzugriffsstruktur (also die gesamte zweite Spalte des zweispaltigen tabellarischen Umtextes; vgl. Abb. 7) als eine Menge von Verweiszielbereichen betrachtet werden. Die einzelnen Umtexteinträge in Abb. 7, also z.B.

600   S + Vb + DatO
           er ähnelt seinem Vater

fungieren nur als Verweiszielbereiche, d.h.: Auf sie kann — wie wir ja bereits gesehen haben — nur im Zuge der Ausführung von Verweisbefolgungshandlungen extern zugegriffen werden. Aus diesem Grund gehört die numerische mediostrukturelle äußere Umtextzugriffsstruktur zu dem Untertyp der mediostrukturellen Zugriffsstrukturen, die *Verweiszielbereichs-Zugriffsstrukturen* heißen. Der zweite wichtige Untertyp der mediostrukturellen Zugriffsstrukturen sind die *adressenvermittelnden Zugriffsstrukturen*; auf diese wird unter 3.3 eingegangen.

Bereits unter 1. wurde darauf hingewiesen, dass nicht alle äußeren Zugriffsstrukturen solche sind, auf die man von außerhalb eines Wörterbuches kommend zugreifen kann. Wir haben jetzt die begrifflichen Mittel zur Verfügung, um dies genauer auszuführen. Anhand der Abb. 3 hatten wir gesehen, dass man auf eine alphabetische makrostrukturelle äußere Zugriffsstruktur „von außerhalb des Wörterbuchs kommend" (also mit Hilfe eines nichtmediostrukturellen Leitelements) extern zugreifen kann. Anhand der Abb. 6 hatten wir gesehen, dass man auf eine alphabetische makrostrukturelle äußere Zugriffsstruktur auch „von innerhalb des Wörterbuchs kommend" (also mit Hilfe eines mediostrukturellen Leitelements) extern zugreifen kann. Schließlich haben wir anhand der Abb. 7 gesehen, dass man auf eine numerische äußere Umtextzugriffsstruktur nur „von innerhalb des Wörterbuchs kommend" (also nur mit Hilfe eines mediostrukturellen Leitelements) extern zugreifen kann. Damit ist zunächst deutlich, dass äußere Zugriffsstrukturen (von welchem Untertyp auch immer) nicht dadurch definiert werden können, dass man sagt, der externe Zugriff erfolge „von außerhalb des Wörterbuchs". Vielmehr ist das entscheidende Definitionskriterium dafür, dass eine Zugriffsstruktur eine *äußere* Zugriffsstruktur ist, der Sachverhalt, dass auf eine äußere Zugriffsstruktur (beliebig oft) extern zugegriffen werden kann, und zwar sowohl im Zuge von Nachschlagehandlungen („von außerhalb des Wörterbuchs kommend") als auch im Zuge von Verweisbefolgungshandlungen („von innerhalb des Wörterbuchs kommend").

Verfügt man über den Begriff des äußeren (oder: externen) Zugriffspfads, lässt sich auch wie folgt argumentieren: Eine äußere Zugriffsstruktur ist eine Zugriffsstruktur, zu der von Benutzern beliebig viele *äußere Zugriffspfade* etabliert werden können. Ein äußerer Zugriffspfad ist der je individuell etablierte, durch die Reihe der äußeren Zugriffstextelemente (z.B. die Lemmareihe) führende Weg, den ein Benutzer-in-actu gegangen ist, wenn er eine externe Zugriffshandlung erfolgreich ausgeführt hat, so dass diese mit dem externen Zugriff auf ein äußeres Zugriffstextelement (als Element der äußeren Zugriffsstruktur) endet. Jeder einfache (i.S.v. nicht zusammengesetzte) äußere Zugriffspfad hat einen Pfadanfang und ein Pfadende. Als Pfadanfang gilt der Ort, an dem das Leitelement erfasst wird. Als Pfadende gilt das zugehörige äußere Zugriffstextelement. Alle Pfadenden sind damit wörterbuchintern. Der Pfadanfang kann dagegen wörterbuchextern oder wörterbuchintern situiert sein. Ersteres ist genau dann der Fall, wenn die pfadetablierende externe Zugriffshandlung im Ausführungsbereich einer Nachschlagehandlung liegt; letzteres ist dagegen immer dann der Fall, wenn die pfadetablierende externe Zugriffshandlung im Ausführungsbereich einer Verweisbefolgungshandlung liegt. Entsprechend lassen sich zwei Typen von einfachen äußeren Zugriffspfaden unterscheiden, und zwar Folgende:

(a)   Der Typ des *adkurrenten äußeren Zugriffspfades*. Dies ist ein Zugriffspfadtyp, zu dem alle einfachen äußeren Zugriffspfade gehören, deren Pfadanfang wörterbuchextern und deren Pfadende wörterbuchintern situiert ist.

(b)   Der Typ des *binnenkurrenten äußeren Zugriffspfades*. Dies ist ein Zugriffs-
pfadtyp, zu dem alle einfachen äußeren Zugriffspfade gehören, deren
Pfadanfang und Pfadende wörterbuchintern situiert sind. Die zuletzt
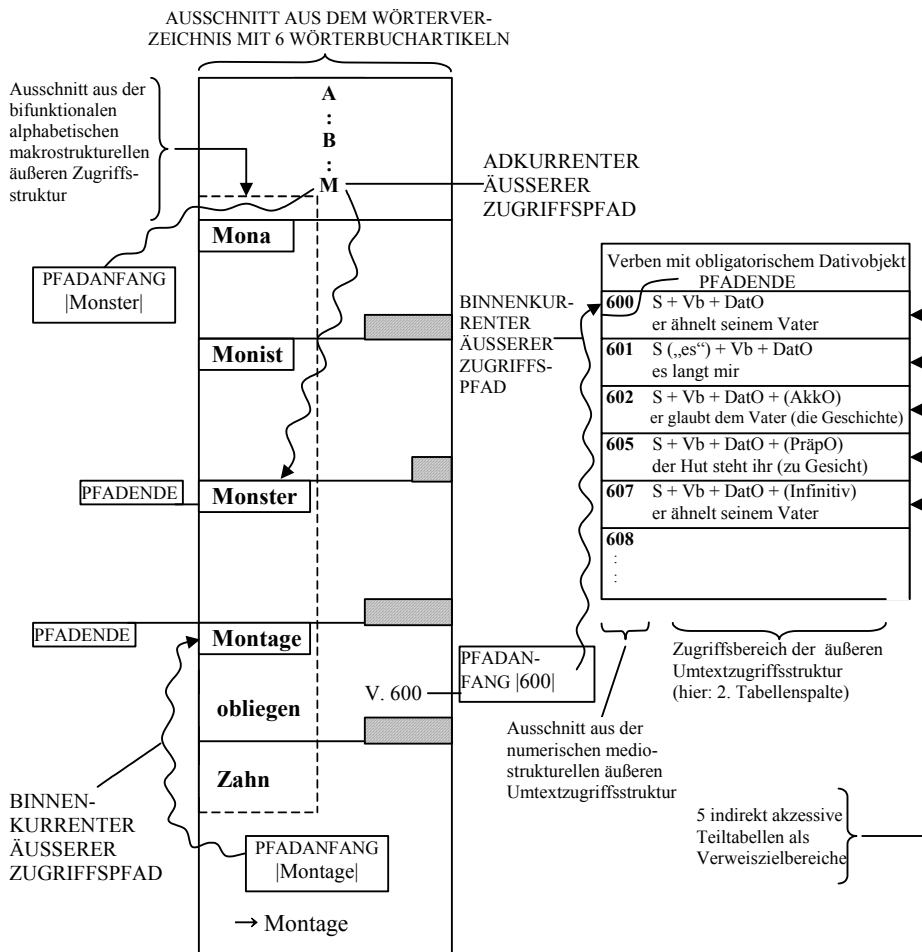eingeführten Begrifflichkeiten sind in Abb. 8 in einen Zusammenhang
gebracht.

Abb. 8:   Veranschaulichung zu äußeren Zugriffspfaden bei äußeren Zugriffs-
strukturen

Hier sei noch ergänzt, dass Zugriffspfade auch *Suchpfade* genannt werden (vgl.
Hausmann und Wiegand 1989: 337f).[6] Unabhängig davon, ob von *Zugriffspfa-
den* oder von *Suchpfaden* gesprochen wird, gilt: Sie dürfen nicht mit Zugriffs-
strukturen verwechselt werden. Leider ist dies in der Literatur nach Hausmann
und Wiegand (1989) und Wiegand (1989), wo zum ersten Mal Zugriffsstruk-

turen erörtert und von Suchpfaden deutlich unterschieden wurden, öfters der Fall. So beispielsweise in Bergenholtz und Tarp (1995: 16 u. 219ff) oder im Anschluss an diese bei Steyn (2004: 276), wo es heißt: „The access structure can be regarded as the search route the dictionary user follows during a dictionary consultation procedure."

Diese Feststellung ist nicht vereinbar mit der Folgenden: „In most general monolingual dictionaries, the central list is the only outer access structure; therefore the macrostructure of these dictionaries coincides with the outer access structure" (Steyn 2004: 280). Hier nun ist die äußere Zugriffsstruktur nicht mehr der Suchpfad (search route) sondern eine textuelle Struktur, die mit der Makrostruktur zusammenfällt, wobei noch die zusätzliche Verwechslung auftritt, dass das Wörterverzeichnis (central list) eine Zugriffsstruktur sein soll! Es sind mithin folgende Klarstellungen dringend erforderlich:

(a)   Such- oder Zugriffspfade werden durch die Ausführung von Such- und Zugriffshandlungen etabliert und sind damit einmalige konkrete Phänomene. Äußere Zugriffspfade, die beim gleichen äußeren Zugriffstextelement enden, sind Elemente der gleichen Klasse von äußeren Zugriffspfaden.

(b)   Äußere Zugriffsstrukturen sind dagegen nicht durch Handlungen konstituierte Phänomene, sondern immer gleiche textuelle Ordnungsstrukturen mit unterschiedlicher Funktion, bei deren Elementen äußere Zugriffspfade enden.

(c)   Ein Wörterverzeichnis ist keine Zugriffsstruktur. Vielmehr weist ein Wörterverzeichnis in einem alphabetischen Wörterbuch eine alphabetische makrostrukturelle äußere Zugriffsstruktur auf. Dies heißt — anders ausgedrückt: Das Wörterverzeichnis ist der direkte Zugriffsbereich der alphabetischen makrostrukturellen äußeren Zugriffsstruktur und ist eine Menge von akzessiven Einträgen, die in diesem Fall Wörterbuchartikel sind.

Auch Registerzugriffsstrukturen sind äußere Zugriffsstrukturen. Diesen gilt unser nächster Einblick.

3.3   Dritter Einblick: Einige Typen von Registerzugriffsstrukturen

Derzeit existiert m.W. weder eine ausgearbeitete Theorie zu Registern, noch gibt es ein ausgearbeitetes typologisches System zu lexikographischen Registern (oder: Wörterbuchregistern).[7] Entsprechend ist auch die Terminologie für die recht zahlreichen Typen von Wörterbuchregistern — wenigstens im Deutschen — derzeit noch unterentwickelt. Im Folgenden werden daher lediglich anhand von einigen Registern, die zu verschiedenen Registertypen gehören,

grundlegende Eigenschaften von Registerzugriffsstrukturen erläutert, was allerdings nur möglich ist, wenn man auch einige wichtige Eigenschaften von Wörterbuchregistern berücksichtigt. Am Ende dieses Einblickes werden wir dann über die Grundbausteine einer Registertheorie verfügen.

Register können im Vorspann oder im Nachspann eines Printwörterbuches stehen; meistens stehen sie im Nachspann, bei Lieferungswerken und Mehrbändern ausschließlich dort. Jedes Register weist mindestens eine äußere Registerzugriffsstruktur auf, die auch *registerinterne Zugriffsstruktur* heißt (vgl. z.B. Wiegand 1989: 393f). Jedes Register ist eine geordnete Menge von *Registereinträgen*. Jeder Registereintrag ist ein akzessiver Wörterbucheintrag, der zu den direkt akzessiven Wörterbucheinträgen gehört (vgl. Abb. 1); er besteht aus einem *Registereingang*, der ein Element (der Trägermenge) der äußeren Registerzugriffsstruktur ist, sowie aus mindestens einer *Registerangabe*, die an den Registereingang adressiert ist. Ein Registereintrag heißt *einfacher Registereintrag*, wenn er genau eine solche Registerangabe aufweist, die direkt auf den Registereingang folgt; er heißt dagegen *erweiterter Registereintrag*, wenn er mehr als eine Registerangabe aufweist, die auf den Registereingang folgen. Hat er n Registerangaben, liegt ein n-fach erweiterter Registereintrag vor. Der Registereintrag „ramponieren ital 1788 V" aus dem Alphabetischen Register des DFW ist z.B. ein *dreifach erweiterter Registereintrag*, weil drei Registerangaben auf „ramponieren", den Registereingang, folgen. Alle drei Registerangaben sind an den Registereingang adressiert. Wird mit dem Registereingang ein sprachlicher Ausdruck genannt, ist der Registereingang stets auch eine *Zeichengestaltangabe*; diese zählt zu den Registerangaben genau so wie die Lemmazeichengestaltangabe zu den Artikelangaben gehört. Register, deren Registereinträge nur aus Registereingängen bestehen, die Zeichengestaltangaben sind, heißen *reduzierte Register* und weisen *reduzierte Registereinträge* auf. Die Menge aller derjenigen Registerangaben eines Registers, die auf einen Registereingang folgen, bildet den *direkten Zugriffsbereich* einer Registerzugriffsstruktur.

Wir betrachten zunächst Register, die Komponenten der Mediostruktur eines Wörterbuchs sind. Solche Register liegen vor genau dann, wenn alle Registerangaben, die auf Registereingänge folgen, nur Fundortdaten für die Ausführung von Verweisbefolgungshandlungen liefern, so dass alle diese Registerangaben als reduzierte Verweisangaben zu gelten haben, anhand derer der Benutzer Verweisadressen erschließen kann, die ihm als mediostrukturelle Leitelemente dienen. Solche Register weisen mediostrukturelle äußere Zugriffsstrukturen auf, die zu dem Untertyp der *adressenvermittelnden Registerzugriffsstruktur* gehören.

Eine adressenvermittelnde Registerzugriffsstruktur findet sich in Prosinger (1984). Da dieses Wörterbuch der alternativen Lexikographie auch weitere interessante Aspekte der Wörterbuchform aufweist, dient es im Folgenden als Beispielwörterbuch. Im vorliegenden Zusammenhang wichtige Aspekte seiner Wörterbuchform sind in Abb. 9 wiedergegeben.
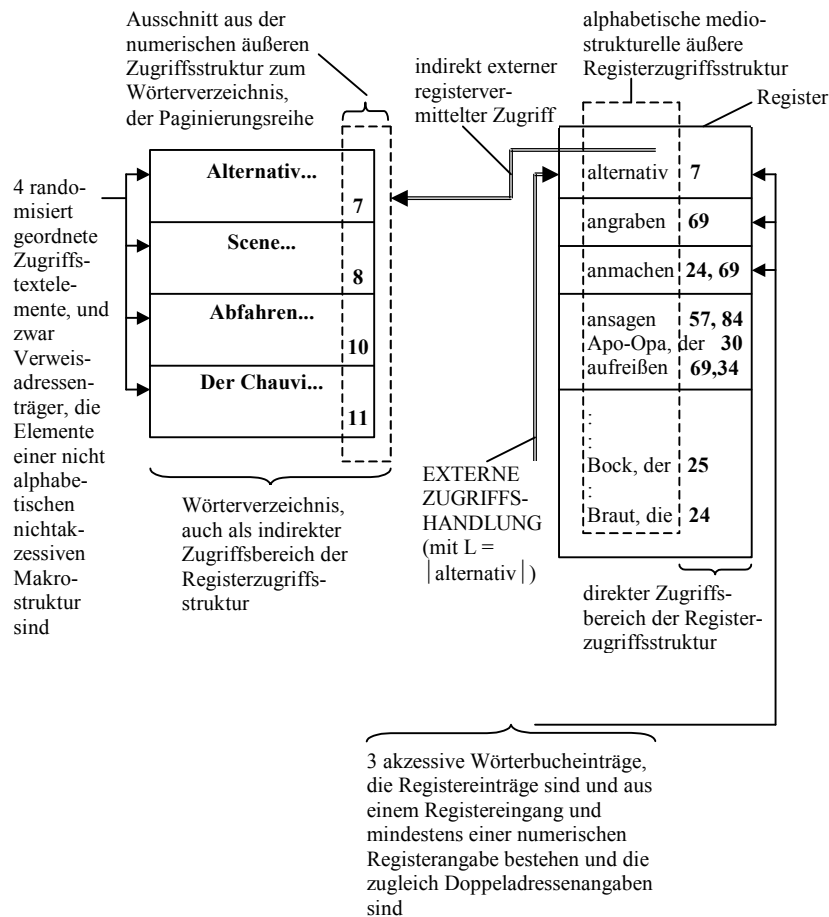
Abb. 9: Veranschaulichung zu einer alphabetischen mediostrukturellen äuße-
ren Registerzugriffsstruktur

In Prosinger (1984) weist die Makrostruktur weder eine der alphabetischen
noch eine der systematischen Anordnungsformen auf. Vielmehr sind die äuße-
ren Zugriffstextelemente zufällig geordnet. Das bedeutet: Aus der Menge aller
möglichen Reihenfolgen, die zu einer Menge von sprachlichen Ausdrücken
gebildet werden können, ist eine Reihenfolge zufällig ausgewählt. Auf eine
solche Makrostruktur kann nicht direkt extern zugegriffen werden. Zwar liefert
z.B. das Wort *Scene* dem Benutzer die Buchstabenfolge |Scene|. Aber diese
funktioniert nicht als Leitelement für den direkten externen Zugriff auf die
Makrostruktur, da deren Elemente keine Leitelementträger und damit keine
Lemmata sind, sondern lediglich Wörterbuchartikeltitel (kurz: Artikeltitel). Der
Makrostruktur von Prosinger (1984) ist mithin keine makrostrukturelle Zu-
griffsstruktur inhärent. Betrachtet man diese Makrostruktur für sich alleine, liegt

eine *nichtakzessive Makrostruktur* vom Typ der randomisierten Makrostruktur vor (vgl. auch Wiegand (2002: 419 u. 2003: 189ff). Prosinger (1984) wäre kein Wörterbuch ohne das Register im Nachspann, das in der Abb. 9 rechts ausschnittsweise abgebildet ist. Dieses Register weist eine alphabetische mediostrukturelle äußere Registerzugriffsstruktur auf, so dass man direkt extern auf dieses Register zugreifen kann. Entsprechend sind alle Registereingänge Leitelementträger. Im direkten Zugriffsbereich der Registerzugriffsstruktur finden sich nur Seitenangaben und damit nur Fundortangaben für denjenigen sprachlichen Ausdruck, der mit dem Registereingang genannt wird. Aus diesem Grund ist die Registerzugriffsstruktur eine adressenvermittelnde Zugriffsstruktur. Denn mit den numerischen Registerangaben, die auf den Registereingang folgen, werden numerische Verweisadressen genannt, die in der Paginierungsreihe von Prosinger (1984) liegen. Entsprechend ist die eleganteste Interpretation eines Registereintrages wie „alternativ **7**" die Folgende: „alternativ **7**" ist eine reduzierte Registerverweisangabe (also eine Verweisangabe, die als Teilangabe keine Verweisbeziehungsangabe, wie z.B. → aufweist). Mit ihr wird eine Verweisdoppeladresse genannt, die aus einer numerischen Verweishauptadresse, nämlich der Seitenzahl 7 besteht, sowie aus einer Verweisunteradresse, nämlich dem erwähnten Artikeltitel alternativ. Ein Benutzer, der mit einer Frage nach den Fundortdaten von *alternativ* im Zuge der Ausführung einer externen Zugriffshandlung den Registereintrag „alternativ **7**" erreicht hat, muss anhand dieses Eintrages einen Verweis erschließen und daraufhin eine externe Verweisbefolgungshandlung ausführen, die ihn zunächst zur Seitenzahl 7 und damit zur numerischen Hauptadresse |7| führt, und danach auf der Wörterbuchseite 7 zum Artikelteil **alternativ**. Dieser fungiert damit als Träger der Verweisunteradresse |alternativ|. Alle Elemente der Makrostruktur in Prosinger (1984) sind damit Verweisadressenträger. Betrachtet man also die Makrostruktur in Prosinger (1984) nicht isoliert, sondern berücksichtigt die adressenvermittelnde Registerzugriffsstruktur, dann ist der Makrostruktur eine indirekt akzessive makrostrukturelle äußere Zugriffsstruktur inhärent.

An dieser Stelle sei zum präzisen Verständnis einer externen Zugriffshandlung, die entweder im unmittelbaren Ausführungsbereich einer Nachschlagehandlung oder im unmittelbaren Ausführungsbereich einer Verweisbefolgungshandlung liegt, Folgendes festgestellt: Der Zugriff, mit dem die Ausführung der jeweiligen Benutzungshandlung beendet ist, ist immer ein *direkter Zugriff*. Das Attribut *indirekt* in entsprechenden Mehrworttermini, die z.B. die verschiedenen Untertypen von akzessiven Wörterbucheinträgen bezeichnen, ist daher wie folgt zu verstehen: Der Benutzer kommt nur über eine andere zugriffsverschiedene (wörterbuchinterne) Position, in der er Fundortdaten (oder: neue Zugriffsdaten) als mediostrukturelle Leitelemente findet und damit in diesem Sinne „indirekt" zum Benutzerziel, obwohl er auch mit der externen Zugriffshandlung, die im unmittelbaren Ausführungsbereich einer Verweisbefolgungshandlung liegt, „direkt" zugreift! Man kann die *indirekte externe Datenakzessivität*, um die es hier geht, auch wie folgt erläutern: Akzessive Wörter-

bucheinträge und damit die Daten, die in ihnen textuell präsentiert werden, sind indirekt extern akzessiv genau dann, wenn die äußeren Zugriffstextelemente stets das Pfadende eines *zusammengesetzten äußeren Zugriffspfades* bilden. Der zusammengesetzte äußere Zugriffspfad besteht aus einem adkurrenten äußeren Zugriffspfad, an den sich unmittelbar ein binnenkurrenter äußerer Zugriffspfad anschließt.

Sucht beispielsweise ein Benutzer anhand des Leitelementes | alternativ | im Register von Prosinger (1984) den Registereingang „alternativ" und hat eine entsprechende externe Zugriffshandlung erfolgreich ausgeführt, dann wurde ein adkurrenter äußerer Zugriffspfad etabliert, dessen Pfadanfang das Leitelement | alternativ | und dessen Pfadende der Registereingang „alternativ" ist. Schließt sich dann direkt eine Verweisbefolgungshandlung an, die den Benutzer zum Wörterbuchartikel zu *alternativ* führt, wird ein binnenkurrenter äußerer Zugriffspfad atabliert, dessen Pfadende ein äußeres Zugriffstextelement, nämlich das Lemma **alternativ** ist. Betrachtet man die gesamte Wörterbuchbenutzung, dann wurde bis zur Erreichung des Benutzerziels ein zusammengesetzter äußerer Zugriffspfad etabliert. Der Weg zum Benutzerziel führt über das Register, so dass ein *registervermittelter Zugriff* erfolgt. In diesem Sinne sind die Wörterbuchartikel in Prosinger (1984) sowie die artikelinternen Daten nur *indirekt extern akzessiv.* Auch die Daten im tabellarischen Umtext des Wahrig (1989) sind nur indirekt extern akzessiv. Auf sie kann extern nur *artikelvermittelt* zugegriffen werden.

Im Folgenden betrachten wir anhand von Abb. 10 ein bifunktionales Äquivalentregister.

Bilingualisierte Fachwörterbücher, wie z.B. die geplanten WSK-Bände (vgl. zu diesen Schierholz und Wiegand 2005) oder polylingualisierte Fachwörterbücher, wie z.B. das Wörterbuch zur Lexikographie und Wörterbuchforschung (vgl. Wiegand 2003a sowie die Demoversion unter: www.fabulex.de), weisen häufig Äquivalentregister auf. In diesen werden z.B. — wie in der Abb. 10 — die englischen Äquivalente mit den Registereingängen genannt, die in den Fachwörterbuchartikeln ebenfalls genannt und an die Lemmazeichengestaltangabe adressiert sind. Mit den Registerangaben von Äquivalentregistern werden die fachsprachlichen Lemmazeichen erwähnt. Möchte ein Benutzer z.B. wissen, wie das deutsche Äquivalent zu *article constituent* lautet, findet er die Antwort bereits im Äquivalentregister selbst: Das Äquivalent heißt *Artikelkonstituente.* Die Menge aller Registerangaben, die auf Registereingänge folgen, bilden daher den *direkten Zugriffsbereich* des Äquivalentregister. Eine Funktion des Äquivalentregisters besteht mithin darin, Benutzerfragen zu beantworten, die zum Fragentyp gehören: WIE HEISST DAS DEUTSCHE ÄQUIVALENT ZU EINEM ENGLISCHEN FACHAUSDRUCK X? Eine weitere Funktion des Äquivalentregisters besteht darin, dem Benutzer die Fundortdaten, also die lemmatischen Verweisaußenadressen als mediostrukturelle Leitelemente für die Ausführung von externen Verweisbefolgungshandlungen zu liefern. Demgemäß sind die Registerangaben zugleich reduzierte Registerverweisangaben, anhand

derer artikelorientierte Verweise erschlossen werden können. Entsprechend ist das Wörterverzeichnis der *indirekte Zugriffsbereich* der alphabetischen Register-zugriffsstruktur, das mithin auch eine mediostrukturelle Funktion aufweist. Für Fachwörterbücher, die ein Äquivalentregister aufweisen, gilt, dass ihre Wörterbuchartikel zu den direkt und indirekt akzessiven Wörterbuchartikeln gehören. Man kann direkt extern auf sie zugreifen und registervermittelt. Anders ausgedrückt heißt das: Alle Lemmata können sowohl als Pfadende eines adkurrenten äußeren Zugriffspfades als auch als Pfadende eines zusammengesetzten äußeren Zugriffspfades auftreten.
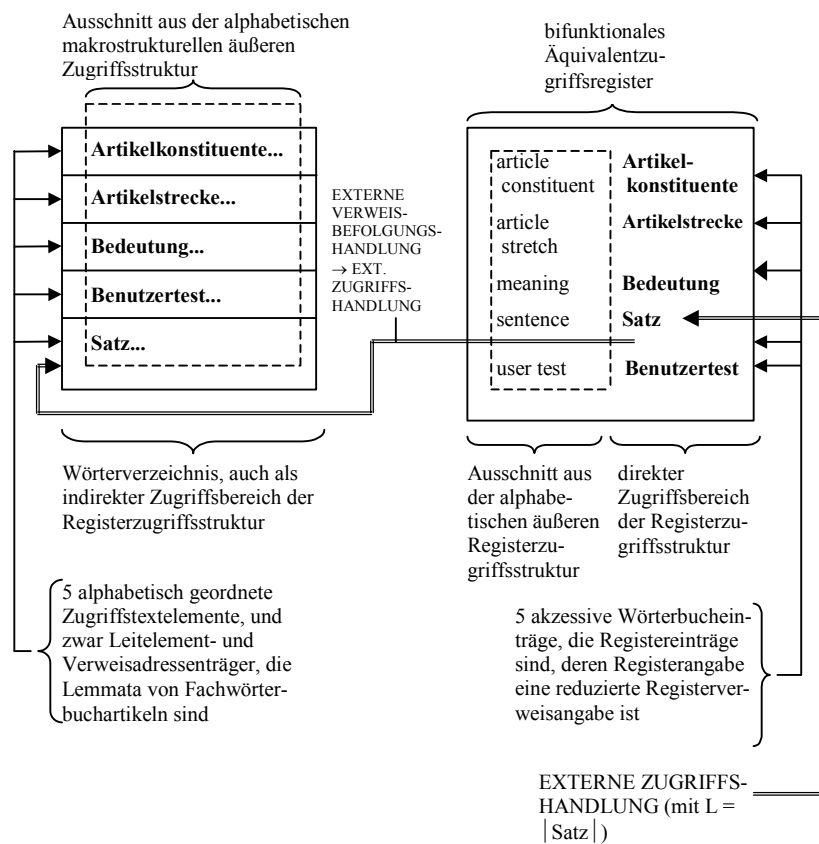
Abb. 10: Veranschaulichung zu einem bifunktionalen Äquivalentregister

Mit Blick auf die externe Datenakzessivität lässt sich auch eine Wörterbuch-typologie angeben. Wörterbücher mit Äquivalentregistern sind z.B. polyakzessiv, was bedeutet: Zu den makrostrukturellen äußeren Zugriffstextelementen gibt es stets zwei Zugriffspfade, die zu verschiedenen Typen gehören. Den polyakzessiven stehen die monoakzessiven Wörterbücher gegenüber, bei denen dies nicht der Fall ist. Sowohl der Typ des poly- als auch der des monoakzes-

siven Wörterbuchs weisen mehrere Untertypen auf, worauf hier jedoch nicht näher eingegangen werden kann.

Im Folgenden betrachten wir anhand der Abb. 11 und 12 das Chronologische Register aus dem 7. Bd. des Deutschen Fremdwörterbuchs (Kirkness 1988).
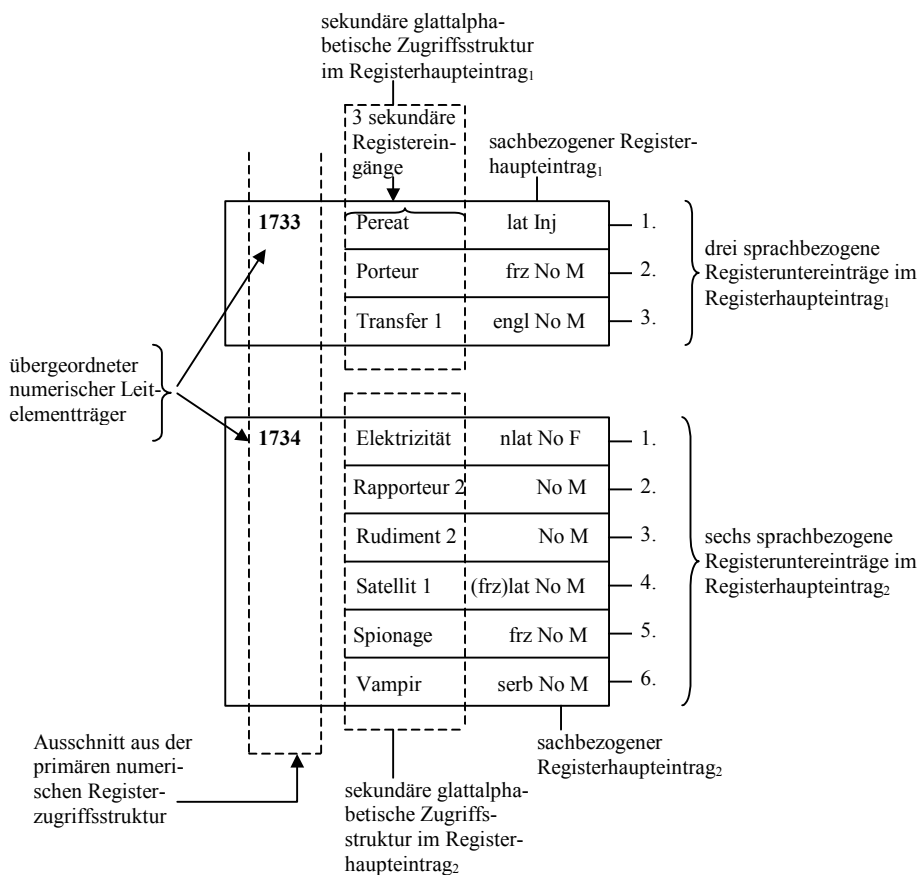

Abb. 11: Erste Veranschaulichung zum Chronologischen Register aus Kirkness (1988); vgl. auch Wiegand (2003: 186)

Das Chronologische Register in Kirkness (1988) gehört zu den Wörterbuchregistern, in denen zwei Zugriffsstrukturen gegeben sind. Zum Status der Zugriffsstrukturen sind zwei lexikographietheoretisch gut begründbare Positionen möglich. Nach der ersten Position liegen nur äußere Zugriffsstrukturen vor. Nach der zweiten Position ist die numerische Zugriffsstruktur eine äußere Zugriffsstruktur, und die glattalphabetischen Zugriffsstrukturen sind innere Zugriffsstrukturen von Registereinträgen. Ich stelle zunächst die erste theoretische Position anhand der Abb. 11 dar.
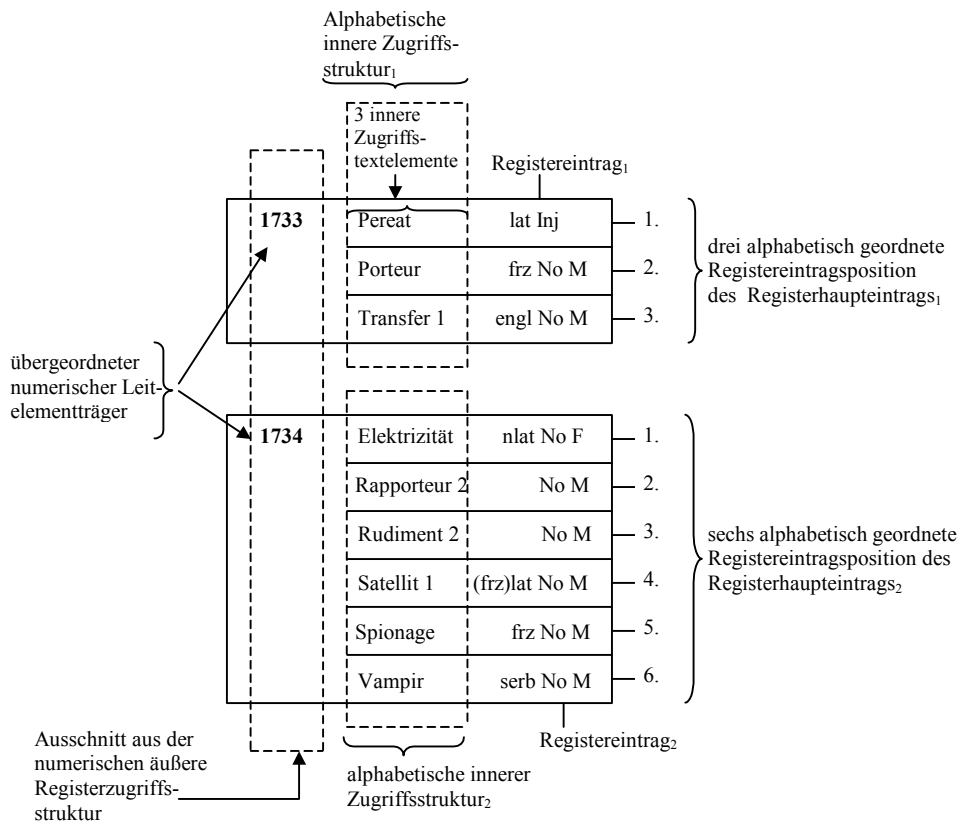
Abb. 12: Zweite Veranschaulichung zum Chronologischen Register aus Kirkness (1988)

Die *primäre* (oder: vorgeordnete) Zugriffsstruktur ist eine *numerische äußere Registerzugriffsstruktur*. Die *sekundären* (oder: nachgeordneten) Zugriffsstrukturen sind kleinere *glattalphabetische äußere Registerzugriffsstrukturen*. Die Leitelementträger der primären Registerzugriffsstruktur sind Jahreszahlen. Auf der Menge aller berücksichtigten Jahreszahlen ist eine zweistellige Ordnungsrelation definiert, auf deren Elemente (also geordnete Paare, deren erste Komponente z.B. 1733 und deren zweite Komponente 1734 ist) der Relationsterm *x geht y zeitlich voraus* zutrifft (mit „x" und „y" als Variablen für Jahreszahlen). Die primäre Registerzugriffsstruktur ist damit eine lineare Ordnungsstruktur, auf die Benutzer z.B. mit Fragen vom Typ WELCHE WÖRTER WURDEN IM JAHR 1734 ENTLEHNT? direkt extern zugreifen können.

Zu jedem numerischen Leitelementträger als *primärem Registereingang* (oder: Registerhaupteingang) gehört ein *Registerhaupteintrag*. Jeder Registerhaupteintrag weist eine eigene *sekundäre* (oder: nachgeordnete) glattalphabetische äußere

Zugriffsstruktur auf, deren Zugriffstextelemente die *sekundären Registereingänge* sind. Entsprechend ist z.B. ‹Pereat, Porteur, Transfer› die sekundäre glatt-alphabetische Zugriffsstruktur, die zum Registerhaupteintrag gehört (vgl. Abb. 11). Ein Registerhaupteintrag besteht aus mehreren *Registeruntereinträgen*, die mit sekundären Registereingängen beginnen. „Pereat lat Inj" (vgl. Abb. 11) ist dann ein *Registeruntereintrag*. Ein Benutzer, der den numerischen Register-haupteingang (z.B. 1733) gefunden hat, kann daraufhin extern z.B. auf den sekundären Registereingang „Pereat" zugreifen.

Anhand der Abb. 12 wird im Folgenden die zweite theoretische Position erläutert. Nur die numerische Zugriffsstruktur wird als äußere Zugriffsstruktur angesehen. Es wird nicht zwischen Registerhaupt- und Registeruntereinträgen unterschieden; vielmehr weist ein Registereintrag n *Registereintragsposi-tionen* auf (mit n ≥ 1). Der Registereintrag$_2$ weist beispielsweise sechs Register-eintragspositionen auf (vgl. Abb. 12). Die Registereintragspositionen sind nach ihrer ersten Angabe, der Fremdwortangabe, alphabetisch geordnet; dadurch ist eine *alphabetische innere Zugriffsstruktur* gegeben. Da auf der Menge aller Fremd-wortangaben aller Registereinträge eine *oberhalb*-Relation definiert ist, sind alle inneren Zugriffsstrukturen aller Registereinträge des Chronologischen Regis-ters in Kirkness (1988) vertikalarchitektonisch ausgebaut und damit glattalpha-betische innere Zugriffsstrukturen.

Welche der beiden erläuterten theoretischen Positionen als die angemesse-nere zu gelten hat, kann nur in einem größeren theoretischen Rahmen entschie-den werden, der in diesem Beitrag nicht zur Verfügung steht.

Obwohl es zahlreiche weitere Typen von Wörterbuchregistern gibt, kann der dritte Einblick hier abgebrochen werden: Die wichtigsten lexikographie-theoretischen Begriffe, die benötigt werden, um Wörterbuchregister hinsicht-lich ihrer Form und Funktion angemessen und exakt analysieren und anlegen zu können, wurden eingeführt. Für die Darstellung zahlreicher theoretischer Details sowie für die Entwicklung eines umfassenden Systems von Register-typologien wird wesentlich mehr Raum benötigt.

4.   Interne Datenakzessivität

Wir wissen bereits aus dem 2. Abschnitt das Folgende: Die interne Datenak-zessivität wird durch innere Zugriffsstrukturen hergestellt. Innere Zugriffs-strukturen findet man nur innerhalb von akzessiven Wörterbucheinträgen. Während für ein Printwörterbuch die externe Datenakzessivität eine *conditio sine qua non* ist, so dass mithin mindestens eine äußere Zugriffsstruktur für ein Printwörterbuch obligatorisch ist, ist die interne Datenakzessivität und damit auch das Vorhandensein von inneren Zugriffsstrukturen fakultativ. Beispiels-weise weisen Wörterbücher, die nur nichtkondensierte Wörterbuchartikel (*sensu* Wiegand 2003: 207ff) aufweisen, die nur aus einem Lemma und einem Lemmakotext bestehen, niemals innere Zugriffsstrukturen auf. In solchen Arti-keln kann man nur lesen, aber nicht — nachdem man das Lemma gefunden hat

— durch die Ausführung von internen Zugriffshandlungen auf eine bestimmte textuelle Artikelposition gezielt zugreifen.

4.1   Vierter Einblick: Einige Typen von inneren Zugriffsstrukturen

Die wichtigste Unterscheidung bei den inneren Zugriffsstrukturen kann wie folgt angegeben werden: Es gibt innere Zugriffsstrukturen, die genau so wie äußere Zugriffsstrukturen funktionieren, und es gibt solche, bei denen das nicht der Fall ist. Die Erstgenannten sind die *alphabetischen inneren Zugriffsstrukturen*. Ein Beispiel sind die inneren Zugriffsstrukturen in den Registereinträgen des Chronologischen Registers in Kirkness (1988), wenn man die zweite Interpretation wählt (vgl. 3.3).

In zahlreichen Fachwörterbüchern finden sich am Ende des Artikels alphabetisch geordnete Literaturangaben, die insgesamt eine textuelle Artikelposition bilden, so dass diese Literaturposition dann eine *positionenspezifische alphabetische innere Zugriffsstruktur* aufweist.

Den alphabetischen stehen die *nichtalphabetischen inneren Zugriffsstrukturen* gegenüber. Sie unterscheiden sich grundsätzlich von allen äußeren Zugriffsstrukturen, da es keine allgemein bekannte, wörterbuchexterne Ordnungsstruktur gibt, nach deren Ordnungsgesetzlichkeiten innere Zugriffsstrukturen angelegt sind, so dass ein Benutzer, der diese Ordnungsgesetzlichkeiten kennt, auf innere Zugriffsstrukturen zugreifen könnte. Vielmehr muss die Kenntnis innerer Zugriffsstrukturen, vor allem die von partiell und vollständig kondensierten Wörterbuchartikeln, auf deren Betrachtung wir uns im Folgenden beschränken, durch die Lektüre entsprechender Ausschnitte der jeweils zuständigen Metatexte, z.B. durch die Lektüre der Benutzungshinweise, erst erworben werden. Dies bedeutet, dass nur der kundige Benutzer, also der, der die Metatexte gelesen hat und über entsprechende Kenntnisse bei der Wörterbuchbenutzung verfügt, interne Zugriffshandlungen ausführen kann.

Innere Zugriffsstrukturen werden in den Benutzungshinweisen dem Benutzer dadurch erklärt, dass ihm mehr oder weniger genau der so genannte Artikelaufbau oder das so genannte Grundschema für den Artikelaufbau erklärt werden. Untersucht man — und eine solche Untersuchung ist relativ schwierig und aufwendig —, was eigentlich in Wörterbüchern genau erklärt wird, wenn der „Artikelaufbau" erklärt wird, dann ist das Ergebnis bei verschiedenen Wörterbüchern des gleichen Typs im Detail recht verschieden. Im Prinzip ist es aber gleichartig: Vor dem Hintergrund der lexikographietheoretisch bisher bekannten Strukturtypen von kondensierten Wörterbuchartikeln, nämlich dem Typ der Artikelkonstituentenstrukturen, dem Typ der Mikrostrukturen und dem Typ der Suchbereichsstrukturen wird unter dem Namen „Artikelaufbau" (oder unter ähnlichen Namen) stets ein Mischtyp aus den drei genannten Typen erklärt. In den Erklärungen wird z.B. auf ausgewählte nichttypographische Strukturanzeigertypen (wie z.B. Kommata auf der Ebene der

Wörterbuchform oder auf senkrechte Striche) und damit auf Elemente der Artikelkonstitutentenstruktur Bezug genommen. Zugleich wird auf bestimmte Typen von elementaren und nichtelementaren Angaben und auf Typen von charakteristischen Angabegruppen und damit sowohl auf Elemente der Artikelkonstituentenstruktur als auch auf solche der Mikrostruktur Bezug genommen. Weiterhin wird auf typographische Mikrostrukturanzeiger (wie z.B. kursiv und fett) hingewiesen und damit auf Eigenschaften von mikrostrukturellen Textkonstituenten; und schließlich werden auch ausgewählte Gestalteigenschaften der Artikeltexte berücksichtigt und damit Elemente von Suchbereichsstrukturen.

Von der in den Benutzungshinweisen erläuterten „Mischstruktur" behält der einzelne kundige Benutzer — wie ich anhand von Wörterbuchbenutzungsprotokollen zeigen konnte — meistens nur einen Strukturausschnitt. Bei verschiedenen Benutzern sind die kognitiv verfügbaren Strukturausschnitte verschieden. Dieses je individuelle Wissen von der in den Benutzungshinweisen erklärten Mischstruktur, mit der zugleich eine oder meistens mehrere innere Zugriffsstrukturen relativ zu bestimmten Artikeltypen beschrieben werden, fungiert als Orientierungs- und Zugriffshilfe für die Ausführung von inneren Zugriffshandlungen.

Während also die inneren Zugriffsstrukturtypen in den Metatexten als „Mischstrukturen" erklärt werden und in den einzelnen Wörterbuchartikeln der verschiedenen Typen als textuelle Mischstruktur tatsächlich auftreten und wahrgenommen werden können, fungiert als Zugriffshilfe für interne Zugriffshandlungen ein je individuelles Strukturwissen von den gegebenen Strukturen, das bei der Handlungsausführung korrigiert und ergänzt werden kann, so dass durch häufige Benutzung des gleichen Wörterbuchs aus dem kundigen allmählich ein erfahrener Benutzer (*sensu* Wiegand 1998: 508) wird.

Damit dürfte der grundsätzliche Unterschied zwischen alphabetischen äußeren Zugriffsstrukturen und nichtalphabetischen inneren Zugriffsstrukturen unter dem Aspekt der Datenakzessivität klar sein: Alphabetische äußere Zugriffsstrukturen garantieren die externe Datenakzessivität jedem potentiellen Benutzer, der das Zugriffsalphabet beherrscht und gelernt hat, das alphabetische Sortieren nachzuvollziehen. Alle potentiellen Benutzer dieser Benutzerklasse sind sozusagen „vor den Gesetzen des Alphabets" gleich. Jeder kann auf jedes äußere Zugriffstextelement punktgenau zugreifen. Der eine kann es schneller, der andere langsamer. Und keiner muss die alphabetische äußere Zugriffsstruktur in dem Sinne „im Kopf" haben, dass er die genaue Reihenfolge aller alphabetisch geordneten äußeren Zugriffstextelemente kennt.

Nichtalphabetische innere Zugriffsstrukturen muss ein Benutzer — wenn er interne Zugriffshandlungen ausführen möchte, die zu einem punktgenauen Zugriff führen — dagegen so „im Kopf" haben, wie sie in den Metatexten beschrieben sind, und das bedeutet: Er muss eine bestimmte Reihenfolge von unterschiedlichen Textsegmenttypen „im Kopf" haben und diese gegebenenfalls für unterschiedliche Artikeltypen variieren können.

4.2   Fünfter Einblick: Bemerkungen zu inneren Schnellzugriffsstrukturen

Um die Eigenart innerer Schnellzugriffsstrukturen leichter erkennen zu können, ist es nützlich, kurz einen Blick auf die äußeren Schnellzugriffsstrukturen zu werfen, und zwar anhand der beiden wohl bekanntesten Typen. Manche alphabetische Wörterbücher weisen zur Erleichterung des direkten externen Zugreifens ein Daumenregister auf. Dieses ist der Träger einer *alphabetischen äußeren Schnellzugriffsstruktur* ‹A, B, C, ..., Z›. Die äußeren Zugriffsstrukturanzeiger A, B, C, ..., Z, die im Wörterverzeichnis vor den Artikelstrecken stehen, werden auf den Ausschnitten des Registers erwähnt, so dass der Anfang einer Artikelstrecke leichter gefunden werden kann. Ein Daumenregister kann wesentlich dazu beitragen, dass die äußeren Zugriffszeiten kürzer sind.

Das zweite Beispiel ist eine *alphabetische lebende Lemmakolumne*. Meistens ist sie so gestaltet, dass auf zwei aufgeschlagenen Wörterbuchseiten auf der linken Seite über dem Wörterverzeichnis das erste Lemma dieser Seite genannt wird und auf der rechten Seite das letzte Lemma dieser Seite. Der Benutzer kann dann schneller entscheiden, ob das gesuchte Lemma auf den beiden aufgeschlagenen Seiten zu finden ist oder nicht.

Um die Vorteile, die eine alphabetische äußere Zugriffsstruktur bietet, nutzen zu können, muss der Benutzer nur das Alphabet kennen. Dies ist bei *nichtalphabetischen inneren Schnellzugriffsstrukturen* anders. Diese sind Teilstrukturen innerer Zugriffsstrukturen. Zu ihrer Trägermenge gehören nur besonders markante strukturanzeigende Angaben und nichttypographische Strukturanzeiger. Diese, ihre Reihenfolge sowie die wesentlichen Angabetypen (zu der die Angaben gehören, die auf sie folgen) muss der Benutzer kennen. Dies sei an einem einfachen Beispiel erläutert.

> **Riemen,** der; -s, - **1.** *Streifen aus Leder, festem Gewebe, Gummi, synthetischem Material*: ein langer, schmaler R.; einen R. verkürzen, zurechtschneiden, um etw. schnallen; die Hose wird mit einem R. (*Gürtel*) gehalten — **2.** *endloses Band aus Leder, Gummi, festem Gewebe, mit dessen Hilfe Drehmomente von einer Welle auf eine andere weitergeleitet werden* — **3.** Rudern *zur Fortbewegung eines Ruderbootes* (*paarweise*) *verwendete, an einem Ende zu einem Blatt verbreiterte hölzerne Stange*: die R. einziehen, ergreifen; sich in die R. legen (*kräftig rudern*)
> + umg. den R. enger schnallen (*sich einschränken*); salopp sich am R. reißen (*sich anstrengen, sich zusammennehmen*)

Abb. 13: Wörterbuchartikel wa_1 aus HWDG

Die innere Schnellzugriffsstruktur von wa_1 kann wie folgt angegeben werden: ‹**Rie**men < **1**. < **2**. < **3**. < **+**›.

Ein kundiger Benutzer, der mit einer Frage „Welche Phraseme mit *Riemen* sind gebucht?" eine externe Zugriffshandlung erfolgreich ausgeführt und damit das Lemma **Riemen** erreicht hat, kann sofort auf die Phrasemidentifizierungsangabe intern zugreifen, die durch „+" realisiert wird, weil er weiß, dass die Phrasemangaben am Schluss des Artikels nach „+" zu finden sind. Er muss also nicht den ganzen Artikel durchsuchen. Dadurch verkürzt sich die innere Zugriffszeit deutlich. Die innere Schnellzugriffsstruktur in wa_1 gehört zu den nicht glatten inneren Schnellzugriffsstrukturen; bei glatten inneren Schnellzugriffsstrukturen ist auf den Elementen der Trägermenge neben einer Präzedenzrelation noch eine *oberhalb*-Relation definiert. Dies ist in wa_1 nicht der Fall.

5.   Schlussbemerkung

In diesem einführenden Beitrag wurde versucht, die wichtigsten Eigenschaften der Wörterbuchform, die mit der Herstellung der Datenakzessivität in Printwörterbüchern zusammenhängen, herauszuarbeiten. Die Vielfalt der Ausprägungen dieser Eigenschaften sowie die Fülle ihrer unterschiedlichen Kombinationsmöglichkeiten konnten dabei nur angedeutet werden. Die zentralen Begrifflichkeiten, um das Datenakzessivitätsprofil von Printwörterbüchern genau beschreiben zu können, wurden jedoch bereit gestellt.

Anmerkungen

1.   Ein für die Wörterbuchforschung angemessener Wissensbegriff sowie eine passende Terminologie werden in Wiegand (2002) und (2002a) vorgeschlagen.
2.   Integrierte Umtexte sind solche Umtexte, in denen Daten zu finden sind, die dazu beitragen, dass ein Wörterbuch seinen genuinen Zweck erfüllt; meistens sind integrierte Umtexte mit dem Wörterverzeichnis mediostrukturell vernetzt.
3.   Akzessive Daten finden sich nur in akzessiven Wörterbucheinträgen; oder anders ausgedrückt: Akzessive Daten sind nur solche Daten, auf die anhand mindestens einer äußeren Zugriffsstruktur extern zugegriffen werden kann.
4.   Solche Wörterbücher finden sich z.B. bei den frühen deutschen Rechtschreibwörterbüchern. Alphabetische Mehrbänder weisen eine vorgelagerte alphabetische Zugriffsstruktur auf. Diese findet sich auf dem Buchrücken und hat bei einem Zweibänder z.B. die Form A–K, L–Z.
5.   Vgl. z.B. Günther 1996.
6.   Für spezifischere Zusammenhänge lassen sich Suchpfade von Zugriffspfaden unterscheiden. Zugriffspfade sind dann solche Suchpfade, die von erfolgreichen Zugriffshandlungen etabliert werden, also solchen, die als Ergebnis einen Zugriff aufweisen.
7.   Kirkness (1989) ist lediglich eine nützliche Beschreibung einiger Funktionen von Registern und eine Sammlung von Beispielen, wo sich Register mit bestimmten Strukturen und Funktionen finden lassen.

Literatur

Bergenholtz, H. und S. Tarp (Hrsg.). 1995. *Manual of Specialised Lexicography.* Amsterdam/Philadelphia: John Benjamins.

Bergenholtz, Henning, Sven Tarp und Herbert Ernst Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Hoffmann, L., H. Kalverkämper und H.E. Wiegand (Hrsg.). 1998–1999. *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research*: 1762-1832. Berlin/New York: Walter de Gruyter.

DFW = Kirkness, Alan (Hrsg.). 1988. *Deutsches Fremdwörterbuch.* Begonnen von Hans Schulz, fortgeführt von Otto Basler, weitergeführt im Institut für deutsche Sprache. 7. Bd. Berlin/New York: Walter de Gruyter.

Duden-²GW = Duden. 1993–1995. *Das große Wörterbuch der deutschen Sprache in acht Bänden.* 2. völlig neu bearb. u. stark erw. Aufl. Hrsg. u. bearb. v. Wissenschaftlichen Rat u. den Mitarbeitern der Dudenredaktion unter der Leitung v. Günther Drosdowski. Mannheim: Dudenverlag.

Günther, Hartmut. 1996. Schrift als Zahlen- und Ordnungssystem — alphabetisches Sortieren. Günther, H. und O. Ludwig (Hrsg.). 1994–1996. *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung/Writing and Its Use. An Interdisciplinary Handbook of International Research*: 1568-1583. Berlin/New York: Walter de Gruyter.

Hausmann, Franz Josef und Herbert Ernst Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J., O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.). 1989–1991: 329-360.

Hausmann, F.J., O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopaedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie.* Berlin/New York: Walter de Gruyter.

HWDG = Kempcke, Günter et al. (Hrsg.). 1984. *Handwörterbuch der deutschen Gegenwartssprache.* In zwei Bänden. Berlin: Akademie-Verlag.

Kirkness, Alan (Hrsg.). 1988. *Deutsches Fremdwörterbuch.* Begonnen v. Hans Schulz, fortgeführt v. Otto Basler, weitergeführt im Institut für deutsche Sprache. 7. Bd. Berlin/New York: Walter de Gruyter.

Kirkness, Alan. 1989. Wörterbuchregister. Hausmann, F.J., O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.). 1989–1991: 767-771.

Mackensen, Lutz. 1977. *Deutsches Wörterbuch. Rechtschreibung. Grammatik. Stil. Worterklärung. Fremdwörterbuch. Geschichte des deutschen Wortschatzes.* 9. völlig neubearb. u. stark erw. Aufl. München: Südwest Verlag.

Prosinger, Wolfgang. 1984. *Das rabenstarke Lexikon der Scene-Sprache. Der große Durchblick für alle Freaks, Spontis, Schlaffis, Softies, Flipper und Hänger sowie deren Verwandte und sonstige Fuzzis.* Illustriert von Peter Gaymann. Frankfurt am Main: Eichborn Verlag.

Schierholz, Stefan und Herbert Ernst Wiegand. 2005. Die Wörterbücher zur Sprach- und Kommunikationswissenschaft. Eine neue Konzeption der linguistischen Fachlexikographie und ihre computergestützte Praxis. *Lexicographica* 20: 164-264.

Steyn, Mariza. 2004. The Access Structure in Learner's Dictionaries. *Lexikos* 14: 275-298.

Wahrig, Gerhard (Hrsg.). 1989. *dtv-Wörterbuch der deutschen Sprache.* 10. Aufl. München: Deutscher Taschenbuchverlag.

Wiegand, Herbert Ernst. 1989. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: Alphabetische Anordnungsformen und ihre Probleme. Hausmann, F.J., O. Reichmann, H.E. Wiegand und L. Zgusta (Hrsg.). 1989–1991: 246-280.

Wiegand, Herbert Ernst. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. 1. Teilband. Mit 159 Illustrationen im Text. Berlin/New York: Walter de Gruyter.

Wiegand, Herbert Ernst. 1998a. Altes und Neues zur Makrostruktur in alphabetischen Printwörterbüchern. Wiegand, H.E. (Hrsg.). *Wörterbücher in der Diskussion III. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 348-372. Lexicographica. Series Maior 84. Tübingen: Max Niemeyer.

Wiegand, Herbert Ernst. 2002. Zur Makrostruktur und zu den äußeren Zugriffsstrukturen im *de Gruyter Wörterbuch Deutsch als Fremdsprache*. Wiegand, Herbert Ernst (Hrsg.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des de Gruyter Wörterbuch Deutsch als Fremdsprache:* 413-442. Lexicographica. Series Maior 110. Tübingen: Max Niemeyer.

Wiegand, Herbert Ernst. 2002a. Altes und Neues zur Mediostruktur in Printwörterbüchern. *Lexicographica* 18: 168-252.

Wiegand, Herbert Ernst. 2003. Überlegungen zur Typologie von Wörterbuchartikeln in Printwörterbüchern. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica* 19: 169-313.

Wiegand, Herbert Ernst. 2003a. Wörterbuch zur Lexikographie und Wörterbuchforschung/Dictionary of Lexicography and Dictionary Research. Städtler, Thomas (Hrsg.). 2003. *Wissenschaftliche Lexikographie im deutschsprachigen Raum. Im Auftrag der Heidelberg Akademie der Wissenschaften*: 417-437. Heidelberg: Winter Verlag.

Capturing Cultural Glossaries: Case-study II [*]

Matete Madiba, *Department: Teaching and Learning Development, Tshwane University of Technology, Soshanguve, Republic of South Africa (madibanrm@tut.ac.za);*
Lorna Mphahlele, *Department: Applied Languages, Tshwane University of Technology, Soshanguve, Republic of South Africa (mphahleleml@tut.ac.za);*
and
Matlakala Kganyago, *Nkoshilo High School, Mankweng, Republic of South Africa (tetema@mweb.co.za)*

Abstract: This case-study attempts to present a brief glossary of Northern Sotho medical terms. It is a follow-up of a similar case-study (Madiba, Mphahlele and Kganyago 2003), which was an attempt to capture and present Northern Sotho cooking terms. Case-study I consists of the names for utensils, ingredients and the processes involved in the preparation of cultural dishes. With both these case-studies, the intention has been to use the opportunities availing themselves for the preservation and valorisation of Northern Sotho, including the extension of its corpora to support national dictionary-making processes. The case-study methodology has been very useful for the purposes of this project and the context within which it was undertaken. It aims to provide a model for the collection and presentation of authentic Northern Sotho terminology which otherwise would hardly have been accessible.

Keywords: CULTURAL GLOSSARY, INDIGENOUS KNOWLEDGE SYSTEMS, CORPUS, OUTCOMES-BASED EDUCATION, ASSESSMENT CRITERIA, TRADITIONAL MEDICAL TERMS, COMMUNICATIVE APPROACH, STRUCTURAL APPROACH, METONYMY, HANDS-ON ACTIVITIES, TASK-BASED LEARNING

Opsomming: Die totstandbrenging van kulturele woordversamelings: Gevallestudie II. Hierdie gevallestudie probeer om 'n beperkte woordelys van Noord-Sotho-mediese terme aan te bied. Dit is 'n voortsetting van 'n soortgelyke gevallestudie (Madiba, Mphahlele en Kganyago 2003) wat 'n poging was om Noord-Sothokookterme te versamel en aan te bied. Gevallestudie I bestaan uit die name van gereedskap, bestanddele, en die prosesse betrokke by die voorbereiding van kulturele geregte. Met altwee hierdie gevallestudies was die bedoeling om die geleenthede te gebruik wat hulleself aanbied vir die bewaring en bestendiging van Noord-Sotho, insluitende die uitbreiding van sy korpora om die nasionale woordeboeksamestellingsprosesse te ondersteun. Die gevallestudiemetodologie was baie nuttig vir die doeleindes van hierdie projek en die konteks waarin dit onderneem is. Dit beoog om 'n model te verskaf vir die versameling en aanbieding van Noord-Sothoterminologie wat andersins skaars toeganklik sou gewees het.

---

[*] This article is a revised version of a 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.

Sleutelwoorde: KULTURELE WOORDVERSAMELING, INHEEMSE KENNISSTELSELS, KORPUS, UITKOMSGEBASEERDE ONDERWYS, WAARDEBEPALINGSKRITERIA, TRADISIO-NELE MEDIESE TERME, KOMMUNIKATIEWE BENADERING, STRUKTURELE BENADERING, METONIMIE, PRAKTIESE WERKSAAMHEDE, TAAKGEBASEERDE LEERWYSE

Background

This project is a continuation of a previous one initiated in 2002 (see Madiba, Mphahlele and Kganyago 2003). It aims at capturing cultural glossaries within the authentic context of a school setting in a rural area in the Limpopo Prov-ince. Of particular interest is that potential projects of this nature have to record cultural words that would otherwise be lost. This study, case-study II, is de-voted to medical terms, gleaned from the preparation and administering of medicinal remedies in Northern Sotho culture. This is not meant to present a comprehensive glossary but to demonstrate the results of a school project, with the hope of stimulating similar ventures with a more comprehensive focus.

As suggested in the previous presentation, working on projects like these will show how these glossaries can help realise and implement innovative methodologies and concepts such as 'simultaneous feedback' (De Schryver and Prinsloo 2000), and 'hybrid dictionaries' to support major lexicographical work. As with the previous project, it must be noted that the glossary is a 'secondary' and not a 'primary' product of the project, because the project had a different aim. The main target is the teaching and learning of Northern Sotho as a first language within the Outcomes Based Education (OBE) environment. The pro-ject is also an acknowledgement that the OBE approach has stimulated an awareness of hands-on approaches for language teaching and learning that were previously not thought of. Mother-tongue teaching and learning in Afri-can languages in general and Northern Sotho in particular was far less engag-ing for both teachers and learners. It is this distinctive feature (of being a 'sec-ondary' product) that has to be investigated for its further implications.

Mora (2001: 4) argues for a so-called 'lexical approach' to the teaching of vocabulary. She asserts:

> Task-based learning should help teachers to provide authentic, meaningful tasks in which students engage to achieve a concrete output, using appropriate lan-guage for the context.

In this case, as is clearly shown, the given task afforded learners a highly engaging context with, as concrete output, a collection of medical terms. Learn-ers were sent out as 'field-workers', to interview elderly members in their community about their medical knowledge. They were to ask them what medi-cines they use for what ailments, and how these are prepared. It turned out to be a demanding, yet interesting learning task as students had to work in groups, preparing a presentation in class for their peers. The case-study ap-proach is found to be more suitable to a project like this as lessons learnt in the

process of compiling this brief glossary will be easier to spell out. It is the exploration of these lessons that is considered to be another step in the process of working towards a possible and authentic model for the collection of other glossaries of this nature.

As was the case with the first case-study, this project also has a double potential. It will hopefully provide ways to supplement the corpus-based approach in the compiling and producing of dictionaries for African languages. It also has the possibility of contributing initiatives in IKS (Indigenous Knowledge Systems). A further factor makes this particular case-study significant to the field with which it deals, the health or medical field. There are a number of nationally initiated projects aimed at all three these directions: the lexical/terminological side, the health/medical side and the IKS side.

On the lexical side, examples would include the products of the Terminology Coordination Subdirectorate of the Department of Arts and Culture such as *Aids Terminology* <http://www.dac.gov.za/about_us/cd_nat_language/terminology/terminology.htm>, as well as *A Multilingual Glossary of Health/Medical Terminology* (Mawela 2003). In the foreword, Machwene Semenya, the Member of the Executive Committee: Sport, Arts and Culture of the Limpopo Province views the glossary as 'a fundamental step towards performing the expectations of the constitution in terms of promotion and usage of indigenous languages'. What is essential for projects like these is the recognition that indigenous knowledge is closely knitted and tied to indigenous languages. Any effort to develop indigenous knowledge systems will to a large extent depend on indigenous languages. It is this same recognition that found value in the school project reported in this case-study.

On the IKS side, the Traditional Health Practitioners Bill [B 66B-2003] legally appears as an interesting factor in this regard. Some of the objectives of the Interim Traditional Health Practitioners Council of South Africa established by this bill are to:

> promote and develop interest in traditional health practice by encouraging research, education and training; and

> promote contact between the various fields of training within traditional health practice … and … set standards for such training.

Research, education, training and communication are all functions dependent on language. The standardization of traditional health practice will contribute much in terms of extending an understanding of indigenous knowledge systems. But then, as argued in case-study I, there should be a recognition that indigenous languages have to be valorised to an extent that they will be instrumental in the development and modernisation of the technical vocabularies; the technical definitions necessary for the standardization of traditional health practice.

Through the use of metonymy (Carstens 1999), lexical equivalents for modernised practices will be more easily provided. An example in this regard

is the plant designated in Northern Sotho as *mošunkwane*, which, well known for its strong smell, is used as an insect repellent, especially for mosquitoes. Though originally it refers to a specific plant with specific properties, it has also come to be widely used to designate all plants with medicinal properties. The second meaning of the word *herb* given in the *Concise Oxford Dictionary* corresponds to this extended meaning of *mošunkwane* in Northern Sotho:

1. any non-woody seed-bearing plant which dies down to the ground after flowering; a herbaceous plant.

2. any plant with leaves, seeds, or flowers used for flavouring, food, medicine, scent, etc.

When it comes to medical/health terminology, the names of plants are often referred to. This is proven by the Glossary in which 31 of the 38 terms are concerned with plants. Van Wyk and Gericke (2000: 119) mention that there are 'somewhere in the region of 4 000 species of plants … used as medicines in southern Africa'. They indicate that 'indigenous African medicine coexists with Western allopathic medicine, Western herbalism, homeopathy, Ayurverdic medicine from India, and traditional Chinese medicine'. They further state that 'not all medicinal plant use can be explained from a reductionist perspective of Western science, and energetic, spiritual, ritual and symbolic aspects are fundamentally important from the perspective of other healing traditions and the world view and expectation of the patient'. African medicinal and health practices have suffered much criticism from those circles heavily influenced by this 'reductionist perspective of Western science'. Even from comments made on the school project reported here, it became evident that the influence of Western science is dominant as the learners and the teachers involved were accused by some of bringing 'demonic practices' into the school. This strong criticism was sparked by the preparations the learners were making for their class presentations about the information they managed to collect on traditional medicines. As a way to demonstrate to what measure they managed to obtain information from their elders, the learners went out of their way to bring along what could help to support their presentations. The 'energetic, spiritual, ritual and symbolic aspects' were well catered for in the class presentations.

Van Wyk and Gericke (2000: 119), accentuating that 'indigenous knowledge systems are dynamic and adaptive', say:

Over the last few hundred years many introduced plants have been incorporated into African medicine in southern Africa, including the extensive use of liquorice (*Glycyrrhiza glabra*), rue (*Ruta graveolens*), camphor (*Cinnamomum camphora*), calamus (*Acorus calamus*), bluegum (*Eucalyptus globulus)* and ginger (*Zingiber officinalis)*.

It is this dynamic and adaptive attribute that will extend the need to develop lexical equivalents, amongst others, for meaningful communication in the traditional medicinal sphere.

Most of the work produced so far has only gone a little beyond the raising of awareness, as Van Wyk and Gericke (2000: 71) admit. They say:

> Southern Africa is exceptionally rich in plant diversity with some 30 000 species of flowering plants, accounting for almost 10% of the world's higher plants. The region also has great cultural diversity, with many people still using a wide variety of plants in their daily lives for food, water, shelter, fuel, medicine and the other necessities of life.

The richness in plant diversity makes it easy to gather encyclopaedic information and compiling an encyclopaedia of this nature should be considered. More than two-thirds of the terms related to plants and given in the Glossary are not found in the extensive study of Watt and Breyer-Brandwijk (1962[2]) about medicinal and poisonous plants. More research is needed in both the directions discussed here, the lexical and the IKS.

Oral tradition/literature as a source for indigenous knowledge

This school project is proof that there is still much cultural knowledge only preserved through oral tradition; and yet the same cultural knowledge continues to shape attitudes and the overall way of living for the communities affected. This is also true of traditional health practices. A number of striking examples are worth citing. There are some prominent ailments or diseases in Northern Sotho culture for which it is difficult to find equivalents within the Western medical frame of reference. These diseases have recognisable symptoms as well as orally preserved ways of treatment. A common problem with the oral transmission of information is its level of variation and the different versions it is likely to produce.

One such ailment is a paediatric condition where a baby shows a certain number of symptoms, some of which include a red mark at the back of the neck and diarrhoea. In trying to find an equivalent, some have related this to meningitis. Although there are a number of versions, the condition is widely known as *bolwetši bja hlogwana.* Another such ailment is a condition in male patients where the symptoms include severe pain, restlessness and deterioration of health. This is called *go khutlega*. Both these conditions, affecting babies and males, are considered fatal. What is interesting is that communities are able to 'diagnose' the patients by recognising the set of symptoms and are able to offer treatment. It is, however, not any member of the community who possesses this expertise to diagnose and treat such ailments. It is only those considered specialists who can do so.

Besides the lack of equivalents demonstrated by the given examples, another interesting phenomenon is the contradictions that are found between traditional health practices and Western medical science and information. A case in point is the beliefs and practices associated with a child with measles. In Northern Sotho culture, it is believed that the condition of a child who is kept

in isolation because of measles may complicate further to the point of dying if the patient comes in contact with a pregnant woman. This will happen, it is believed, when the pregnancy is in its early stages and the woman is still keeping it secret. According to Western medical knowledge, a woman in the early stages of pregnancy is advised to stay away from children infected with measles because the virus can also infect the foetus.

Task-based approach to language teaching within OBE

Educators in South Africa, including those from rural communities, are faced with great changes, i.e. in the restructuring of the education system in the country including the introduction of 'outcomes-based education (OBE)'. Some view the changes as 'resource-hungry' especially in rural communities who are generally poorer, whereas others have exploited the opportunities the changes have opened for innovative teaching and learning.

Within the OBE framework, educators must, when planning assessment, identify the outcomes envisaged before approaching learners. What is really appreciable in this approach is that everyone must have clarity about the final expectations. This means that the educator must find ways and means of exposing learners to learning opportunities in order to show their full potential in terms of knowledge, skills, attitude and values. Van der Horst and Mac-Donald (1997) maintain that learners must take an active part in learning. Their critical thinking should be developed together with their reasoning and reflection. This is in line with the approach that Mora (2001) argues for in what she calls 'a lexical approach' to the teaching of vocabulary and language as such.

In this school project, a communicative approach in language teaching was adopted as it connects the learner with everyday life situations. Language as such is a means of communication; it depends on the skills of listening, observing, singing, speaking, reading and writing amongst others. Learners are expected to be proficient in responding to texts, including oral texts. This project was intended to guide the learners to respond to cultural and social values in texts, in order to contextualise learning through hands-on activities.

Learners were given the task of identifying various types of illnesses occurring in their communities. In the structural approach, learners would only name the illness, without a further description of how it affects people in everyday life. Example:

> *Bolwetši bja go ruruša marama ke*_____.
> (An illness that causes the cheeks to swell, is _____)
> The expected answer is *mauwe* (mumps).

In this project, learners were divided into groups to research types of illnesses and remedies traditionally used. Each group of learners was advised to choose one type of illness to research. They were encouraged to ask their parents, grandparents and the elderly in their communities for help with information. They

were given a week to complete the task. Each group brought the ingredients and prepared that particular medicine in class to show how it is mixed and how the dose is 'measured'. This included making fires so that each group could heat the required ingredients as part of the preparation. Learners were asked to submit a written version for further assessment.

Glossary

As in the first case-study, supplementary work was done to expand the glossary, which involved consulting informants able to provide the information needed. Although the focus of inquiry was effectuated by the initial ideas and terms originating from the school project, extra information was necessary in some instances. The following table is a list of 38 terms gleaned from the presentations. It is evident that this kind of project has the potential to serve as a model for the collection of an authentic glossary which can be of help in the compilation of dictionaries or encyclopaedias in the African languages.

Medical terms	Description of the term	Preparation of the medication	Symptoms or ailments that it treats
boloko bja tonki	Donkey's dung.	Leave the dung to dry, and smoke it like a cigarette.	Relieves headache.
boya bja kome	Fur of an animal, the *kome*.	Put the dry fur onto a fire, cover the patient with a blanket and let him/her inhale the smoke.	Relieves symptoms related to cancer.
disema	Tree of which the root is used.	Put the root onto a fire, cover the patient with a blanket and let him/her inhale the smoke.	Cures severe headache.
dithokolo tša dipudi	Goat's dung.	Leave the dung to dry, grind, mix with *lengana* (see Glossary), cook, let the infusion cool and drink.	Relieves symptoms of measles.
kgaši	Tree of which the root is used.	Chew the root while still fresh and swallow the liquid only.	Reduces abnormally fast heartbeat.
kgato	Tree of which the roots are used.	Cook the roots while still fresh, massage the back with the infusion while it is still hot.	Relieves backache.

lebake	Wild dagga.	Cook the leaves, leave the infusion to cool and drink.	Cures influenza.
lengana	Plant not growing very tall.	Cook the leaves, leave the infusion to cool and drink.	Cures influenza.
letlalo la kwena	Small piece of skin cut from a crocodile.	Cook the skin in water, leave the infusion to cool and drink.	Helps prevent (unintentional) abortion and premature delivery.
mahlare a mmilo	Wild fruit-producing tree of which the bark is used.	Dry the bark, grind, and pour it around the yard.	Protects the family from evil spirits.
modišo	Tree of which the roots are used.	Dry the roots, then grind and mix with sand from the crossroads; pour this into cold water and drink.	Administered to infants when they are generally unwell.
moditšana or *lejakolobe*	Tree of which the roots are used.	Cook the roots, leave the infusion to cool and drink.	Cleanses the blood.
mogwete	Tree of which the roots are used.	Cook the roots while still fresh, leave the infusion to cool and drink.	Helps against high blood pressure.
mogapu wa basadi	Tree not growing very tall of which the leaves are used.	Cook the leaves while still fresh, leave the infusion to cool and drink.	Cleanses the blood.
mokgoba	Tree of which the roots are used.	Cook the roots while still fresh, leave the infusion to cool and drink.	Relieves discomfort of the stomach.
molaka	Tree of which the bark is used.	Dry the bark, then grind it and pour a little in warm water and drink.	Cleanses the blood.
monee	Tree of which the bark is used.	Dry the bark, then grind and mix a little in porridge and eat.	Cures diarrhoea.
monokwana	Tree of which the roots are used.	Break the roots into small pieces, tie them together, cook, leave the infusion to cool and drink.	Cures stomach ulcers.
morola-kgomo-kgopha	Cow's skin.	Leave the skin to dry, grind, burn and lick the ashes.	Cures ulcers.

moroto wa ditšhoši	Tree of which the roots are used.	Cook the roots, leave the infusion to cool and drink.	Administered to children to relieve discomfort of the body.
moroto wa tšhwene	Tree of which the root is used.	Soak the root in water in a special dish, the *kgopa*, and drip the infusion into the ears, eyes and nose.	Relieves symptoms of measles.
mošalašopeng	Tree of which the leaves are used.	Put the leaves underneath the sheet when going to sleep.	Massages the whole body.
moselesele	Tree of which the roots are used.	Cook the roots while still fresh, leave the infusion to cool and drink.	Relieves general discomfort of the body.
moterebe	Grape tree.	Cook the roots while still fresh, leave the infusion to cool and drink.	Cures thrush in the vagina.
motopane	Tree of which the root is used.	While the root is still fresh, squeeze its juice into the eye.	Cures sore eyes.
mphara-tšhwene	Tree of which the roots are used.	Burn the roots to ash, cut the painfully swollen area of the patient's body, and smear the ash into the cut.	Helps against painful swellings.
naka la kgomo la medu	Root of a cow's horn.	Burn the root of the horn to ash, mix with softened Sunlight soap and smear it onto the wound.	Heals a swelling that has become a wound.
nkekologe	Tree of which the roots are used.	Put the roots onto a fire, cover the patient with a blanket and let him/her inhale the smoke.	Relieves symptoms of asthma.
seahlokwana	Tree of which the roots are used.	Cook the roots, leave the infusion to cool and drink.	Cures the kidneys, and purifies the urine.
sebale	Evergreen flowering plant, not growing very tall.	Crush the leaves while still fresh, put it in cold water and drink.	Relieves bloatedness and flatulence.

seefa-maeba	Tree of which the root is used.	Put the root onto a fire, cover the patient with a blanket and let him/her inhale the smoke.	Relieves stomach cramps.
sefala-badimo	Tree of which the roots are used.	Cook the roots while still fresh, leave the infusion to cool and drink.	Counteracts poison in the stomach.
sekanama	Tree of which the roots are used.	Cook the roots while still fresh, leave the infusion to cool and drink.	Cleanses the blood and cures dysmenorrhoea.
sekgoma	Mixture of different herbs: *sekanama*, *molaka* and *serokolo* (for all three see Glossary).	Grind together, put the mixture onto a fire, cover the patient with a blanket and let him/her inhale the smoke.	Relieves general discomfort of the body.
selomi	Tree of which the roots are used.	Grind the roots, mix with mealie-meal, cook and eat.	Cures stomach ulcers.
serokolo	Plant of which the root is left to dry before use.	Chew it and rub it onto the forehead. Crush it, mix it with *lengana* (see Glossary), cook it and let it cool before using.	Relieves condition following the attending of a funeral. Cures influenza.
thola-kgomo	Tooth of a certain animal.	Burn on coal and inhale the smoke through the open mouth.	Relieves toothache.
tšhikwane	Roots from the *mokgalo* tree.	Grind the roots and lick the powder.	Cures venereal diseases.

Linkages across related projects

The dictionary units sometimes use field-workers to enlarge their corpora. This case-study, together with case-study I, proves that schools and other institutions of learning can involve learners for this purpose of rapidly developing the corpora of languages through similar projects. This will minimise the use of adoptive words and increase the application of indigenous terms. Another advantage is that the elderly members of the targeted community will readily furnish information to their children, instead of to strangers visiting their homes with the aim of interviewing them, as is normally the practice.

If well coordinated and supported, similar projects can serve as 'tributaries' feeding bigger projects. Well-funded projects such as those that lead to the compilation of *A Multilingual Glossary of Health/Medical Terminology* in the Limpopo Province could benefit from and be supported by such initiatives. The methodology and expertise in the collection of the terminology can be transferred to projects in other African languages. Encyclopaedic information can be gathered through projects of this nature.

References

Carstens, A. 1999. *Lexicography and Terminography 876*. Lecture Notes. Pretoria: University of Pretoria.

De Schryver, G.-M. and D.J. Prinsloo. 2000. The Concept of 'Simultaneous Feedback': Towards a New Methodology for Compiling Dictionaries. *Lexikos* 10: 1-31.

Madiba, M., L. Mphahlele and M. Kganyago. 2003. Capturing Cultural Glossaries: Case-study I. *Lexikos* 13: 206-214.

Mawela, A. et al. (Eds.). 2003. *A Multilingual Glossary of Health/Medicinal Terminology.* Translated from English into Afrikaans, North Sotho, Tshivenda and Xitsonga. Polokwane: Language Services: Department of Sport, Arts and Culture, Limpopo.

Mora, S. 2001. Teaching Vocabulary to Advanced Students: A Lexical Approach [Online]. Available from: <http://www.seasite.niu.edu/Tagalog/Teachers_page/Language_Learning_Articles/teaching_vocabulary_to_advanced_.htm> [accessed 17 January 2005].

Thompson, D. (Ed.). 1995[9]. *Concise Oxford Dictionary of Current English*. Oxford: Clarendon Press.

Traditional Health Practitioners Bill [B 66B-2003]. 2003. As amended by the Portfolio Committee on Health (National Assembly).

Van der Horst, H. and R. Macdonald. 1997. *Outcomes-Based Education: Theory and Practice*. Pretoria: Kagiso.

Van Wyk, B.-E. and N. Gericke. 2000. *People's Plants. A Guide to Useful Plants of Southern Africa*. Pretoria: Briza Publications.

Watt, J.M and M.G. Breyer-Brandwijk. 1962[2]. *The Medicinal and Poisonous Plants of Southern and Eastern Africa*. Edinburgh/London: E. & S. Livingstone Ltd.

Developing a Learner's Corpus: The Case of a First-year Module in Mathematics

Christa van der Walt *(cvdwalt@sun.ac.za), Department of Curriculum Studies*, and
Hanelle Fourie *(hfourie@sun.ac.za), Department of Afrikaans, University of Stellenbosch, Stellenbosch, Republic of South Africa*

Abstract: A learner's corpus is a body of writing for use by a student whose first language is not (in this case) English to improve his/her use of (in this case) academic scientific terminology. In this case study, a learner's corpus was developed for a first-year mathematics module for students in the biological sciences. Lecturers struggle with big classes and a fairly high failure rate which they have addressed in a variety of ways. The learner's corpus is one of an array of support mechanisms built into the teaching-learning process and aims to support the development of academic literacy in this module in particular. In the process of developing and refining this learner's corpus it was compared to Coxhead's Academic Word List to determine whether a general academic word list may not include enough terms to render tailor-made learner's corpora unnecessary. The study concludes that the most frequent terms used in this module either do not appear in the Academic Word List or have such a specialised meaning that general academic support on the basis of the word list would probably not be very useful for students.

Keywords: LEARNER'S CORPUS, ACADEMIC WORD LIST, ACADEMIC LANGUAGE SUPPORT, MATHEMATICS WORD LIST, PEDAGOGIC APPLICATION OF CORPORA, CASE STUDY

Opsomming: **Die ontwikkeling van 'n leerderskorpus: Die geval van 'n eerstejaarswiskundemodule.** 'n Leerderskorpus is 'n versameling tekste vir gebruik deur 'n student wie se eerste taal nie (in hierdie geval) Engels is nie, om sy/haar gebruik van (in hierdie geval) akademiese wetenskaplike terminologie te verbeter. In hierdie gevallestudie is 'n leerderskorpus ontwikkel vir 'n eerstejaarswiskundemodule vir studente in die biologiese wetenskappe. Dosente sukkel met groot klasse en 'n redelik hoë druipsyfer, wat hulle op verskillende maniere benader het. Die leerderskorpus is een van 'n reeks ondersteuningsmeganismes wat in die onderrig-leer-proses ingebou is en beoog om die ontwikkeling van akademiese geletterdheid spesifiek in hierdie module te ondersteun. Tydens die ontwikkeling en verfyning van die leerderskorpus is dit vergelyk met Coxhead se Academic Word List om vas te stel of hierdie woordelys nie genoeg terme insluit om maatpas leerderskorpusse oorbodig te maak nie. Die studie kom tot die gevolgtrekking dat die frekwentste terme wat in hierdie module gebruik word óf nie in die Academic Word List voorkom nie óf so 'n gespesialiseerde betekenis het dat algemene akademiese hulp op grond van die woordelys waarskynlik nie vir studente baie nuttig sal wees nie.

Sleutelwoorde: LEERDERSKORPUS, AKADEMIESE WOORDELYS, AKADEMIESE TAAL-HULP, WISKUNDEWOORDELYS, OPVOEDKUNDIGE TOEPASSING VAN KORPUSSE, GEVALLESTUDIE

1.   Introduction

The impulse for the development of a learner's corpus in mathematics was two-fold. Firstly one of the authors wanted to continue an investigation started some years earlier, whereby the importance of *language* support for the teaching of mathematics, science and biology was investigated and specific recommendations were made for the use of English *and* the learners' home languages to facilitate understanding (Van der Walt, De Beer and Mabule 2001). In the second place a colleague's experiences with the development of a learner's corpus made us aware of the language teaching possibilities of this instrument, whereby "stress is laid on frequency of occurrence, a form of information which is for the first time starting to become widely and informatively available to the language teacher through corpora" (Leech 1997: 15).

A learner's corpus, as will be discussed in detail below, is a body of writing to be used by a student whose first language is not (in this case) English to improve his/her use of (in this case) academic scientific terminology. The notion of a corpus as an information source fits in very well with the shift in university teaching philosophy over the past twenty years or so, a trend to move from teaching as imparting knowledge towards teaching as mediated learning (Leech 1997: 2).

The main purpose of the project was firstly to determine whether a learner's corpus (that would appear useful to the *lecturers* involved) could be developed on the basis of the study material in a particular course. The second purpose was to see to what extent this corpus would agree with Coxhead's (2000) Academic Word List, since a comparison would give some idea as to the usefulness of using the Academic Word List in academic support environments. The question was, therefore, to what extent a *general* academic support measure such as the Academic Word List would be useful to students in a fairly specialised subject.

2.   Background

2.1   Origin and rationale for corpus linguistics

The term "early corpus linguistics" is often used to describe linguistics before it made any contribution to the field. Field linguists, for example Boas who studied American-Indian languages, and later linguists of the structuralist tradition all used a corpus-based methodology, although the term "corpus linguistics" did not necessarily appear in texts and studies from this era, and was only introduced later.

Roughly between 1876 and 1926, the so-called diary studies period of language acquisition, child language was studied based on diaries carefully composed by the children's parents. These early corpora are still used as sources of normative data in language acquisition research today. Corpus collection continued and diversified after the diary studies period: large sample studies covered the period roughly from 1927 to 1957 — analysis was gathered from a large number of children with the express aim of establishing norms of development. Longitudinal studies have been dominant from 1957 to the present, this time with a smaller sample of children (approximately 3) studied over longer periods of time.

Shortly after the diary studies period mentioned above, Kading used a large corpus of German — 11 million words — to collate frequency distributions of letters and sequences of letters in German in 1897. This corpus, by size alone, is impressive for its time, and is also comparable to modern corpora in terms of size (Leech 1997).

Fries and Traver, and Bongers are examples of linguists who used the corpus in research on foreign language pedagogy during the 1940s. There had been a strong link beween the corpus and second language pedagody in the early half of the twentieth century, with vocabulary lists for foreign learners often being derived from corpora. In 1921, Thorndike determined the relative frequency of each of the words in his corpus of 4.5 million words, which included classic works of literature and children's books. Later, in 1944, this early work was revised as *The Teacher's Wordbook of 30 000 Words*. This in turn had an influence on the *General Service List of English Words*, compiled in 1953 by Michael West, a work which has been described as the most well-known and persistent corpus-based description of the English lexicon for pedagogical purposes (Lancaster University n.d.).

2.2   Pedagogic application of corpus studies

2.2.1   Rationale for the pedagogic application of corpora

As indicated above, the notion of a corpus as an information source fits in very well with the shift in university teaching philosophy that places more emphasis on the autonomy of the learner and learner-centred teaching. The perceived distance between research and teaching is bridged as students are encouraged to do some research of their own (Leech 1997: 3):

> Teaching is a natural extension of research. The student-centred paradigm of 'discovery learning' […] can scarcely be better exemplified than through the use of the computer corpus. […] A corpus is, of itself, a rich resource of authentic data containing structures, patterns and predictable features that are waiting to be 'unlocked' by the human intelligence.

Although the purpose in this case was not to make students draw up their own corpus, it was argued that a corpus or word list derived from the study mate-

rial would have greater face validity as well as learning value for the students who are not, on the whole, enthusiastic about using general or even subject-specific dictionaries. Furthermore, the responsibility for making meaning by using the corpus would be theirs, especially if lecturers and tutors referred them to the source.

From the beginning of this project it was envisaged that the resulting corpus will be translated into Afrikaans and isiXhosa. In South Africa, there have been several calls for the development of terminology in the African languages, notably by Read and Ambrose (1999) and Carstens (1999). The usefulness of having explanations or terms available in the students' primary language has been demonstrated by Carstens and her Master's student Manyane in their 1997 survey (Carstens 1999). Most teachers who use a language of teaching and learning (a LOLT) that is not the students' primary language instinctively code switch and use either a home language or the dominant classroom language (see Van der Walt and Mabule 2001). However, a learner's corpus goes one step further and to some extent formalizes terms that teachers may or may not accept as scientific terms in a language other than English. As such, the corpus discussed in this article is very much in the spirit of what Carstens (1999: 3) describes as a specific philosophy towards terminologisation: "Use what you have. That which you lack, can be borrowed or adapted along the way."

2.2.2   Types of corpora applications

The Hong Kong University of Science and Technology developed a one-million-word corpus of English computer science texts, intended to assist the teaching of English for computer science students in Hong Kong. The corpus consists of three 2 000-word samples from each of some 166 English language textbooks used in computer science courses at the University in the early 1990s.

There are also other specialised corpora with applied linguistic purposes such as the *Jiao Tong University Corpus for English in Science and Technology* (JDEST) and the *Guangzhou Petroleum English Corpus* (GPEC), both produced in China. They are designed to help students analyse certain registers of language use and include counts of high frequency words. The JDEST Corpus was compiled in the 1980s and contains about one million words of written English texts from mainly the physical sciences, engineering and technology. The GPEC is slightly smaller and contains about 411 000 words in 700 texts from the petroleum industry. These texts originated from written American and British English sources of the mid-1980s (Kennedy 1998: 44).

2.2.3   The distinction between a corpus and a learner's corpus

The *New Collins Concise English Dictionary* defines a corpus as "a body of writings, especially by a single author or on a specific topic: *the corpus of Dickens' works*". A learner's corpus, therefore, is a body of writing, to be used by a

learner. However, the term "learner's corpus" is not to be confused with the widely used "learner corpus" — they are quite different corpora altogether. A learner corpus is a body of text, spoken or written, consisting of, for example, ESL learners' use of English. A corpus such as this is helpful to examine the kind of mistakes that learners make, and enable language teachers to target problem areas. "Learner's corpus" is often substituted by "LSP (Language for Special Purposes) corpus" to make the distinction a little clearer. Some linguists, however, feel that it should not be called a corpus at all, but a glossary or dictionary of specialist terms. See, for example, the *Cambridge Learner's Dictionary* (available from <http://uk.cambridge.org/elt/catalogue/0521663660/>) or the *Longman Dictionary of Contemporary English* (<http://www.pearsoned-ema.com/pdfs/elt_dictionaries.pdf>).

3.   Application of a learner's corpus in this study

3.1   Introduction

Few members of academic support units at South African universities and technikons will disagree that institutions of higher education continue to receive students whose language development, and therefore scholastic achievements, are "fragmentary and incomplete" (Skutnab-Kangas 1981: 240). The necessity of investigating how best to provide language support has led to a variety of techniques and materials, and the project reported on here is such an attempt. There can be little doubt that such a body of technical language, allowing students to draw from a bank of scientific terms and their appropriate use, would be of indeterminable value to bridge the existing language gap and provide the necessary language support that would enable students to follow lectures and tutorials with more understanding, read the textbooks and assignments with more insight and study more effectively. There is in fact an emerging worldwide need for corpora or term banks or subject lexicons such as the one that has resulted from this study. As Read and Ambrose (1999: 173) indicate, the key to accessibility in academic subjects is vocabulary and it is on this basis that the commonly-used academic and university word lists by Nation and Coxhead were developed.

    As a result of discussions with lecturers and academic support specialists[1] at the University of Stellenbosch, a first-year mathematics module for the biological sciences (Mathematics (Bio) 124) was identified as a module that might benefit from additional support in the form of a learner's corpus. This module is regarded as 'problematic' because almost 50% of the students, who are enrolled for a wide range of degree courses (including agricultural science and microbiology), fail the module. The classes are big but lectures are supplemented with tutorials. The perception that the module is 'difficult' is exacerbated by a big difference in interest and ability among students. Part of the reason for the high failure rate could be that the students, many of whom are not

English home language speakers, struggle to come to terms with the volume of specialised mathematical language with which they have to cope, particularly when using an English textbook of intimidating thickness and by listening to such language (which could be English or Afrikaans) in classes and tutorials.

We hypothesized that a corpus of mathematical texts would contain a high frequency of words that are different to those found in a standard academic corpus and argued that if this were the case it would be necessary to provide students with a study resource that clearly explains and even illustrates these specialised terms, so that students can easily look up problematic terms for better comprehension and use of academic texts.

The method of research, therefore, is that of a case study as "an intensive, detailed description and analysis of a single project, program, or instructional material in the context of its environment" (as defined by Education and Human Resources 2004). As such, the aim was to provide "[a] detailed analysis of an individual or group, especially as an exemplary model of a particular phenomenon" (Institute of Science and Technology, USA). This links up with the purpose of a learner's corpus, which is designed to serve the needs of a specific group of learners.

Since the design of the learner's corpus was to a large extent based on the prescribed textbook *Calculus and Its Applications* (Goldstein, Lay and Schneider 2004)[2], it was necessary first to negotiate and obtain copyright permission from the publisher Pearson Education. Then the study material generated by lecturers was obtained in electronic form. After this had been done, an electronic corpus was compiled that contained the entire prescribed text, tests, tutorials and homework assignments for the particular course. An electronic concordancing tool, Wordsmith Tools (version 3.00.00, 1999-02-18), was used to extract a frequency list of the words occurring in the texts. There are quite a few electronic corpus tools available at present, but Wordsmith Tools was chosen for its user-friendliness and affordability.

The resulting word list was 'cleaned up' and non-content words like *and* and *but* were removed. Statistics originally provided in the list about the relative frequency of the words, was also eliminated, and the remaining words were arranged alphabetically. The result was a word list of 13 pages of the most frequent mathematical/scientific terminology found in the corpus as described above.

3.2   Comparison with Coxhead's Academic Word List

According to Coxhead (2004: 1), the Academic Word List (AWL) was developed at the School of Linguistics and Applied Language Studies at Victoria University of Wellington, New Zealand. The list which contains 570 word families does not include words belonging to the most frequent 2 000 words of English, because it was primarily designed to be used by teachers as part of a programme preparing learners for tertiary level study or used by students

working on their own to learn the words most needed to study at tertiary institutions.[3] When one looks at the words from the AWL in the context of the mathematics word list, it soon becomes clear that even though these terms may be taught in an academic support course, students need to learn their meaning in the specialised, mathematical sense. For example, words like *integral, normal, range, area* and *complex* have a specific meaning in mathematics and a course that teaches them in a general, academic sense will probably not help students much.

A comparison of the mathematics word list with the AWL produced interesting results. In the mathematics corpus 1 018 words occurred with a frequency higher than one: of these only 114 appeared in the AWL, perhaps because the most frequently used, ordinary English words are not included in the AWL. However, words with a low frequency in the mathematics corpus were included in different categories of the AWL, e.g. *theorem, stress, strategy, specific, sequence, select, potential, precise, predict*, etc. Conversely, mathematical terms with the highest frequency did not appear in the AWL, as the short excerpt in Table 1 shows.

Table 1:
Excerpt from the most frequent items in the mathematics word list compared to Coxhead's AWL

N	Word	Freq.	%	Appearance in Coxhead's AWL
11	FUNCTION	129	0.94	AWL sublist 1
13	GRAPH	94	0.69	–
15	FUNCTIONS	84	0.61	AWL sublist 1
29	NUMBER	40	0.29	
30	EXAMPLE	39	0.29	
33	FIGURE	37	0.27	–
37	CALCULUS	34	0.25	–
39	APPLICATIONS	33	0.24	–
41	DOMAIN	33	0.24	AWL sublist 6
42	POINT	33	0.24	–
45	CURVE	30	0.22	–
46	GRAPHS	30	0.22	–
50	LINE	29	0.21	–

Excerpt from the least frequent items in the mathematics word list compared to Coxhead's AWL

N	Word	Freq.	Appearance in Coxhead's AWL
1821	RATIO	1	AWL sublist 5
1852	REQUIRED	1	AWL sublist 1
1860	REVENUE	1	AWL sublist 5
1862	REVISION	1	AWL sublist 8
1869	ROLE	1	AWL sublist 1
1897	SELECT	1	AWL sublist 2

1903	SEQUENCE	1	AWL sublist 3
1922	SITES	1	AWL sublist 2
1943	SPECIFIC	1	AWL sublist 1
1944	SPECIFYING	1	AWL sublist 3
1948	SQUARE	1	–
1958	STRATEGIES	1	AWL sublist 2
1960	STRESS	1	AWL sublist 4
1968	SUBSEQUENT	1	AWL sublist 4
1969	SUBSTITUTED	1	–
1971	SUBSTITUTION	1	AWL sublist 5
1983	SUMS	1	AWL sublist 4
1999	TANGENT	1	–
2019	THEOREM	1	–
2078	*VERGELYKING**	1	
2080	*VERHOUDING**	1	
2094	VISUALIZE	1	AWL sublist 8

*Afrikaans words from the Afrikaans study notes and assignments

3.3   Consulting with mathematics lecturers

In an effort to make the corpus seem manageable to first-year students an attempt was made to shorten it.[4] Although the AWL presented us with a basis for comparison, it did not solve our problem of which words could be omitted from the final word list. This problem was addressed in the next step of the research: consulting with course lecturers and tutors to obtain input about which terms are indeed more problematic than others, or which words are used in a mathematical sense that may be different from the more commonly used and understood one.

In a workshop, we presented lecturers and tutors involved in this module with a 'cleaned up' version of the word list — a list containing only the words and their frequency of occurrence. This was the first time that we as researchers met with the teaching staff because our initial meetings were only with the module coordinator. As language specialists, we had to explain why we thought the study was necessary and possibly useful. There was much discussion on the rationale for the study and we had to justify our viewpoint on the importance of language as the medium of instruction. Some lecturers felt that a general dictionary would be good enough, but others pointed out that students generally do not buy dictionaries and may actually use a shorter, tailor-made text. Some of the tutors in particular seemed to have a better understanding of the *language* problems that non-mother-tongue speakers of English and particularly students from a disadvantaged background might have with an English mathematics textbook. It was also clear that language problems and the perceived difficulty of the module were not the only reasons why students

failed. The fact that the module was compulsory for programmes which students felt should not require mathematics also led to negative attitudes towards it.

In the course of the workshop, we carefully went through the list, asking the teaching staff to identify those words which, in their experience, cause the most problems for students. There was some disagreement but after three hours we could finalise a list that everybody felt actually included the most important general and technical terms of relevance for successful learning.

In the final phase of the project the list was translated into Afrikaans and isiXhosa and English definitions were found for the terms included in the list. The list was edited and redesigned with hyperlinks so that students can click on an English term and find an English definition and the Afrikaans or isiXhosa translation. Since the final product will be used by first-year students who are relatively inexperienced and unsophisticated as far as academic reading skills are concerned, we decided to stick to the English terms as the basis for the word list and, in so doing, obviate the risk of it reaching "excessive proportions", as Read and Ambrose (1999: 174) caution.

The final document will be put on a website where students can find information about this particular module in 2005. A short excerpt is provided in Table 2 as an example of the final document.

Table 2: Excerpt from the final word list

English term	Afrikaans term	isiXhosa term
algebra	**algebra**	**ufundo-manani ngeesimboli**

The branch of mathematics that uses symbols to study numbers and the relations between them.

The use of **algebraic** symbols such as a, b, x, y having variable values makes for greater scope than is possible in arithmetic, which uses only constant numbers such as 5 and 5 ½ .

algorithm	**algoritme**	**i-algorithim**

A set of steps for finding the solution to a problem. Algorithms are especially important in programming a machine (e.g. a computer) to carry out computations.

It would be ideal if the definitions and descriptions can also be translated and the list designed in such a way that students can use it from an Afrikaans or isiXhosa list. In this process, the provision of examples in isiXhosa would probably also help to clarify the meanings of words currently rendered as suffixes, as in Table 3, for example.

Table 3: Examples of isiXhosa terms rendered as suffixes

English	Afrikaans	isiXhosa
appropriate	gepas	-fanelekile
approximate	by benadering	-kufutshane, -malunga

4.   Conclusion

The current project can only be declared a success if students actually use the word list. As a research project we managed to reach our goals, which were to extract a word list that would appear useful and valid to the lecturers concerned and we found valuable information after comparing the mathematics word list to Coxhead's AWL. The word list should be useful not only to the students enrolled in this module, but also to staff working in academic support, since mathematics is usually one of the subjects in which additional academic support is provided.

It is hoped that lecturers will further an awareness of the word list and that students will find it useful. It will be possible to trace students' use of the list because it will be possible to detect the number of visits to the website. Before similar projects are attempted in other problematic modules, a follow-up investigation will be done to determine whether students and lecturers use the word list.

Endnotes

1.   We would like to thank Proff. Kosie Smit from the Institute of Mathematics and Science Teaching at the University of Stellenbosch (IMSTUS) and Pieter Maritz of the Department of Mathematics, University of Stellenbosch, as well as all the lecturers and tutors involved in the teaching of Mathematics (Bio) 123 for their help and, above all, their precious time.

2.   We would like to thank Pearson Educational for selling us an electronic copy of the textbook which is not available commercially. They must also be thanked for their cooperation and permission to use the textbook for this study.

3.   For detail on the development and evaluation of the AWL, see Coxhead (2000).

4.   It is at this point that the learner's corpus can probably be described more aptly as a word list, since the actual corpus has now been adapted to such an extent that it probably cannot be said to reflect the 'completeness' implied by the word *corpus* anymore.

References

Carstens, A. 1999. Science through Sepedi: Is Terminologisation a Worthwhile Venture? *Lexikos* 9: 1-17.

Colorado State University, Writing Centre. 1997. *Case Study: Introduction and Definition* [Online]. Available from: <http://writing.colostate. edu/references/research/casestudy/pop2a.cfm> [viewed 2005-05-09].

Coxhead, A. 2000. A New Academic Word List. *TESOL Quarterly* 34(2): 213-238.

Coxhead, A. 2004. *The Academic Word List* [Online]. Available from: <http://www.vuw.ac.nz/lals/ research/awl/awlinfo.html> [viewed 2004-06-03].

Education and Human Resources. 2004. *Chapter 9. Glossary* [Online]. Available from: <http:// www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153/CHAP_9.HTM> [viewed 2004-11-01].

Goldstein, L.J., D.C. Lay and D.I. Schneider. 2004. *Calculus and Its Applications*. Tenth Edition. Upper Saddle River, N.J.: Pearson Education.

Institute of Science and Technology, USA. 2004. *Process Steps. Appendices: Glossary* [Online]. Available from: <http://www.labplan.org/glossary/> [viewed 2004-11-01].

Kennedy, G. 1998. *An Introduction to Corpus Linguistics*. London: Addison Wesley Longman Limited.

Lancaster University, Department of Linguistics and Modern English Language. *Early Corpus Linguistics* [Online]. Available from: <http://www.ling.lancs.ac.uk/monkey/ihe/linguistics/corpus1/1early.htm> [viewed 2004-06-02].

Leech, G. 1997. Teaching and Language Corpora: A Convergence. Wichmann, A., S. Fligelstone, T. McEnery and G. Knowles (Eds.). *Teaching and Language Corpora.* London: Addison Wesley Longman.

Read, J. and M. Ambrose. 1999. Towards a Multilingual Dictionary of Academic Words. *Lexikos* 9: 172-187.

Skutnab-Kangas, T. 1981. *Bilingualism or Not: The Education of Minorities.* Clevedon: Multilingual Matters.

Van der Walt, C., J. de Beer and R. Mabule. 2001. Letting the L1 in by the Back Door: Code Switching and Translation in Science, Mathematics and Biology Classes. *SAALT Journal* 35 (2-3): 123-134.

Van der Walt, C. and R. Mabule. 2001. Language Status and Covert Prestige in the Code Switching Practices of Mathematics, Science and Biology Teachers. *SAALT Journal* 35 (2-3): 257-268.

To Call or Not to Call a Spade a Spade: The Dilemma of Treating 'Offensive' Terms in *Duramazwi Guru reChiShona**

Emmanuel Chabata *(emmanuelchabata@yahoo.com)* and
Webster M. Mavhu *(vhezh2000@yahoo.com), African Languages Research
Institute (ALRI), University of Zimbabwe, Harare, Zimbabwe*

Abstract: As noted by Béjoint (2000: 6), the main objective in dictionary-making is to define words and terms. This is especially the case if the fact is accepted that dictionaries are mostly consulted for word meaning and that, in the consultation process, the user hopes to acquire and/or verify certain information. However, as again noted by Jackson (1988), Landau (1984), Svénsen (1993) and Zgusta (1971), among others, the description of word meaning is one of the greatest difficulties with which the lexicographer has to cope. This article discusses some of the challenges facing lexicographers when defining 'offensive' headwords in a monolingual dictionary. It is based on experiences of defining such headwords for the general-purpose, medium-sized, synchronic, monolingual Shona dictionary, *Duramazwi Guru reChiShona* (henceforth DGC) (Chimhundu et al. 2001). DGC was compiled and edited by a six-member team of mother-tongue speakers of Shona who are researchers at the African Languages Research Institute (ALRI). The article also discusses some of the strategies the team of editors adopted as ways of dealing with offensive words in DGC. One such strategy is the use of euphemism in defining.

Keywords: ALRI, CULTURE, DEFINITION, THEORETICAL DEFINITION, PRACTICAL DEFINITION, DICTIONARY, EUPHEMISM, LEXICOGRAPHY, MONOLINGUAL DICTIONARY, OFFENSIVE HEADWORD, SHONA

Opsomming: Om die kind by sy naam te noem of nie te noem nie: Die dilemma van die behandeling van "aanstootlike" terme in *Duramazwi Guru reChiShona*. Soos deur Béjoint (2000: 6) opgemerk, is die hoofdoel van woordeboekmaak om woorde en terme te definieer. Dit is veral die geval as die feit aanvaar word dat woordeboeke meestal geraadpleeg word vir woordbetekenis en dat, by die raadplegingsproses, die gebruiker hoop om sekere inligting te bekom en/of te kontroleer. Soos weer eens opgemerk deur onder andere Jackson (1988), Landau (1984), Svénsen (1993) en Zgusta (1971), is die omskrywing van woordbetekenis egter een van die grootste moeilikhede wat die leksikograaf moet hanteer. Hierdie artikel bespreek sommige van die uitdagings waarvoor leksikograwe te staan kom wanneer "aan-

---

stootlike" trefwoorde in eentalige woordeboeke gedefinieer word. Dit is gebaseer op die ervarings by die definiëring van sulke trefwoorde vir die meerdoelige, mediumgroot, sinchroniese, eentalige Sjonawoordeboek, *Duramazwi Guru reChiShona* (voortaan DGC) (Chimhundu et al. 2001). DGC is saamgestel en geredigeer deur 'n span van ses moedertaalsprekers van Sjona wat navorsers by die African Languages Research Institute (ALRI) is. Die artikel bespreek ook sommige van die strategieë wat die redakteurspan gekies het as maniere om aanstootlike terme in DGC te behandel. Een so 'n strategie is die gebruik van eufemismes by definiëring.

Sleutelwoorde: ALRI, KULTUUR, DEFINISIE, TEORETIESE DEFINISIE, PRAKTIESE DEFINISIE, WOORDEBOEK, EUFEMISME, LEKSIKOGRAFIE, EENTALIGE WOORDEBOEK, AANSTOOTLIKE TREFWOORD, SJONA

1.   Introduction

Defining is one of the most important stages in the compilation of a dictionary. Although it has become the norm in modern lexicography also to include information about the spelling, pronunciation, etymology, etc. of words, the most central information is about the description of their meaning. Meaning is even more central when it concerns monolingual dictionaries. In his description of monolingual dictionaries, Landau (1984: 8) 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 by native speakers, what other words mean.

The importance attached to meaning in monolingual dictionaries means that defining should be taken seriously. But what should a lexicographic definition giving the word meaning do or be like? According to Landau (1984: 120), a definition should be equivalent to or should capture the essence of the thing defined. It should be as informative and as enlightening to the reader as possible. To do this, a dictionary definition should be based on truth, clarity, accuracy and preciseness. It should say what a word means, explicitly. In other words, the description should both be pointed and direct, which is one of the basic principles of defining. A dictionary definition should be a description of a thing that enables the reader to recognise the thing when he/she sees it or feel it when he/she gets in touch with it, either physically or mentally. This kind of definition does not have any biases and is not conditioned by such factors as cultural, social, geographical, racial or historical context. Such is the theoretical or ideal dictionary definition, where a spade is called by its name.

The theoretical definition described above is objective. Objectivity is an important principle to achieve, especially if the fact is recognised that a dictionary is a normative and didactic reference tool, that it is an instrument primarily consulted when searching for information about a word in particular, about language in general or about the world at large. As noted by Béjoint (2000: 18-

19), dictionaries are instruments of self-teaching for, during the consultation process, their users learn or verify something that they did not know or that they had momentarily forgotten. It is because of this very important role that dictionaries have in society that the information they contain should be presented in an objective way.

However, dictionary-making is not entirely a theoretical enterprise; it is also a practical exercise. In addition to observing the principles guiding the compilation of dictionaries, lexicographers should also pay attention to the needs, expectations and cultural norms of their target users. Landau (1984: 121) notes:

> Lexicography is a craft, a way of doing something useful. It is not a theoretical exercise to increase the sum of human knowledge but practical work to put together a book that people can understand … Every lexicographer, like any good author, has his readers very much in mind.

The lexicographer's pen is sometimes not free to write its owner's views and convictions. The lexicographer is more often than not forced to write in a manner that does not invite controversies or disputes about his/her work or that does not disconcert a section of the society he/she intends to inform. In other words, he/she is forced to sacrifice some principles of his/her profession for the sake of pleasing the users for whom he/she is writing.

From this discussion, it can be noted that two aspects need to be balanced when defining, that is, the theoretical principles on which the lexicographer bases his/her definitions on the one hand and the cultural context in which the lexicographer is working on the other. A balance between these brings about good lexicographic practice. The process of trying to balance these two aspects makes defining a very complicated stage in dictionary-making. The reason for this is that the theory behind defining sometimes conflicts with the cultural norms as well as the expectations of the target users which the lexicographer should also observe.

In the consecutive sections of this article, the challenges the compilers of *Duramazwi Guru reChiShona* (henceforth DGC) faced in their treatment of offensive terms will be discussed. Their dilemma concerned their decision whether to be faithful to the theoretical aspect of defining (that is, being explicit and direct in their descriptions) or to follow the practical path (which implies being euphemistic). However, it may first be necessary to briefly look at the principles that guided the headword and sense selection processes of DGC.

2. Headword and sense selection in DGC

Headword and sense selection processes are of paramount importance in the compilation of any dictionary. These are the processes that determine the major contents of a particular reference work. As noted by Landau (1984: 185), the

decision to include or exclude a particular category of words and/or senses relates directly to the purpose of the dictionary. He argues that if a dictionary is conceived of as purely descriptive, then every word documented as being in sufficiently wide use should be included, no matter how objectionable. In the same vein, Trench (1857) (quoted by Béjoint 2000: 119) argues that, being an inventory of the language, a general dictionary should be descriptive; hence, it should include all the words used by speakers of the language. Trench rejects the idea that the dictionary-maker should only select words considered to be "good" or "appropriate" in a dictionary, at the same time omitting those that are felt "bad" or "inappropriate". He further accentuates that the task of the lexico-grapher, who he describes as a historian and not a critic, is to collect and ar-range all the words despite the fact that they may or may not sound good or appropriate. What Landau and Trench say about headwords to be included in a dictionary can also be said about the kinds of meanings the words should carry.

DGC was meant to be a comprehensive description of Shona. It was in-tended to be a general purpose dictionary describing all possible Shona words and their meanings and/or usages in the different Shona communities by speakers of all ages and in all possible contexts. As such, it was supposed to be as inclusive as possible of the whole vocabulary and complete usage of the lan-guage. No category of words or word meanings was regarded as unsuitable for inclusion in the dictionary. This was unlike the case with *Duramazwi reChiShona* (henceforth DRC) (1996), the forerunner of DGC, where some categories of words such as offensive and other terms were avoided for a number of reasons.

DRC was intended as a reference tool mainly for use by secondary school students, and the inclusion of offensive terms, for example, was believed to be "dangerous" to young users. In a few cases where such terms did find their way into the dictionary, the description of their meanings was left vague for fear of their sensitivity. However, as has already been noted, DGC was firstly in-tended to be a more comprehensive reference work also meant for adults who include Shona teachers and students studying at higher secondary and tertiary education levels. Secondly, the compilation of DGC was corpus-based, al-though, of course, it was not necessarily corpus-bound. The team of compilers of the dictionary decided to include all words and senses gleaned from the Shona linguistic corpus. The idea behind this decision was that the existence of the different words and word usages in the corpus was enough evidence that these words and senses were in use by Shona speakers. The implications of being corpus-based were that the majority of headwords, senses, citations and other relevant linguistic information required for the compilation of the dic-tionary would come from the Shona corpus. The corpus, which currently com-prises approximately three million running words, was built from oral and written materials focusing on a wide range of aspects of Shona life. Among the words that feature in the corpus and that had to be selected as headwords for inclusion in DGC, were offensive words.

3.   Offensive words in Shona

To clearly understand the position of offensive terms in Shona the reality that every society has a culture shaping its people's worldview needs firstly to be accepted. Since they have a certain way of viewing the world, the people tend to have some biases towards ideas that prevail in their society, at the same time being critical of those ideas that go against the way they are looking at life. A people's worldview can be seen through the people's attitudes towards factors such as, among others, social behaviour, dressing, eating habits, and of course language use. When it comes to language, people in any culture tend to have their preferred styles of speaking as well as words and expressions they believe to contain their distilled knowledge about life, acquired through their experiences over generations. On the other hand, they also have words and expressions they think should not be used, those they even discourage their children from uttering. This latter category includes offensive and/or taboo words a person should strive to avoid in speech. For purposes of discussion in this article, offensive and/or taboo words will be taken to refer to a variety of words that may be considered vulgar, obscene, impolite, insulting and derogatory, those referring to the physically deformed or mentally impaired and all those usually being avoided or normally found unacceptable in the Shona cultural context. Words in this rough category are considered unsuitable for use by any "normal" person. Their use goes against the norms of good behaviour for they are either injuring the dignity or hurting the feelings of an individual or social group in whose presence they may be used or to whom they may refer.

A large body of offensive words in Shona consists of those that may be called obscene. Obscenity has been described by Landau (1984: 183) as "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". As can be inferred from this quotation, obscene words are mostly those that refer to certain parts of the human anatomy as well as their functions. These are usually the private parts of the body that have to do with sex and excrement and whose exposure in public is seen as taboo. The following are a few illustrative examples of Shona words that may be regarded as obscene: *mboro* (penis), *beche* (vagina), *chindori* (clitoris) and *jende* (testicle). Obscene words also include those referring to the excretory organs and/or bodily processes. Examples of such terms in Shona are *mukosho* (anus), *-mama* (defecate), *-tunda* (urinate/ejaculate) and *-sura* (fart). Such words are socially so restricted that they may not be used in public without causing embarrassment either to the speaker or to the listener(s). However, it must be stressed that obscenity is circumstantial. Although these terms are discouraged or prohibited in speech they may be used freely in appropriate circumstances. For example, they are licensed for use when people are intimate or when maturing youths are taught about sex and sexuality.

Offensive words also include those terms generally considered insulting. Some examples of insulting words in Shona are *musatanyoko* (referring to the genitals), *nyamai* and *musvirinyoko* (both referring to the referent's mother's vagina) and *shiti* (shit). To offensive terms further belong derogatory racial or ethnic words used to refer to individuals or groups whose herid) appearance, geographical origin or social status are regarded as inferior. Among examples of derogatory terms in Shona are words such as *murungudunhu/ murungudondo* (albino), *mukarushu* (person of Mozambican origin), *bwidi* (person of Malawian origin) and *dzviti* (a Ndebele person).

In Shona culture, just as in most African cultures, the use of offensive words in ordinary speech is normally unacceptable. Although most speakers of Shona may be aware of the existence of these offensive words, they may not use them freely to refer to events, activities, objects or people. They are generally believed to be sensitive, thus they may not be used without offending those who care about cultural and social norms and good behaviour. In this case, therefore, sensitive words are not normally used for fear of offending other people. Instead, people would prefer the use of equivalent but indirect terms in place of such words, as this is usually believed to be culturally and socially polite, testifying to good manners. In the following section, the treatment of some of the examples of offensive terms in DGC is considered.

4.   Treatment of offensive words in DGC

In the previous sections, it has been noted that the objective of compiling DGC was an attempt to produce a comprehensive description of the Shona language in its possible totality. Part of this objective entailed the insertion of all possible words in the dictionary, including offensive terms which can either be obscene or insulting. The principles guiding the creation of definitions for dictionary headwords had also been considered. Particulary because general-purpose dictionaries such as DGC serve as instruments for self-education and social integration (Béjoint 2000: 21), dictionary definitions should be as informative, enlightening and explicit as possible. A close analysis of these few points shows that, as far as Shona monolingual lexicography is concerned, there is a conflict between producing an effective self-educative tool and producing a socially compatible tool. Given all this, one is tempted to ask: What is the position of the Shona lexicographer in all this, and how can good lexicographic practice be achieved in this situation?

Dubois and Dubois (1971) (quoted in Béjoint 2000: 20) describe the role of the lexicographer as that of a "mediator" between society and its people, as a mouthpiece of his/her society. If this description of Dubois and Dubois is anything to go by, then lexicographers (and hence the dictionaries they produce) are not free from ideological biases characterising the society/-ies from which they come. In other words, the social ideology of the speakers of the language being described is the one determining the way words should be defined in

dictionaries. This social ideology, for example, determines the extent to which the objectivity principle can be upheld in the compilation of a dictionary of any kind. The role of the lexicographer is to try to balance the dictates of his/her profession and the demands of the culture he/she is trying to portray and of the people he/she is writing for. In a bid to strike a balance between these two aspects, lexicographers face a number of challenges. The challenges the compilers of DGC faced with regard to the treatment of offensive words will be discussed in two stages, that is, (a) whether offensive words should be included in or excluded from the dictionary, and (b) if they are included as dictionary entries, how the meanings of these words should best be described.

Initially the editors had decided to entirely exclude offensive words from DGC. For a number of reasons, exclusion was to be the easiest option for the editors to handle them. As has already been stated, offensive words are not freely and openly used in Shona culture, whether in speech or in print. All the editors being mother-tongue speakers of Shona entails that their lives and backgrounds are deeply rooted in Shona culture. The cultural issue is one factor that made the editors of the dictionary experience some discomfort with the idea of including offensive words in the dictionary. This discomfort came from the realisation that, as lexicographers, they were expected to respect their society's cultural norms and morals with regard to the limits of what is communicable to the public. Another reason causing discomfort was the educational background of the editors who all had primary and secondary education at Christian mission schools where Christian morality was a priority. The emphasis on an adherence to the "good" and an avoidance of the "bad" indirectly affected the editors when it came to the selection of offensive terms as headwords in the dictionary. This was also compounded by the editors' consciousness of the fact that the dictionary would be used in Zimbabwe's educational system where there exists a policy discouraging writers and artists from explicitly using and describing offensive terminology in school books.

The editors were furthermore conscious of the criticism launched at DRC, DGC's forerunner, by some Shona conservatives. When DRC was published in 1996, one prominent scholar reviewed the dictionary. Although offensive headwords were entered with caution and were defined with discretion, keeping the target users in mind (Mberi et al. 1997: 24), the reviewer attacked the editors for entering and defining the terms in the dictionary. He commented, for example, that "the degree of specificity of most of the items goes beyond what the reader who is a native speaker wants to know" (Kahari 1996–97: 38). He further alleged that by containing offensive entries, the dictionary lacked "traditionally accepted taste and decorum". Thus, the editors of DGC had to be very careful when dealing with offensive words since some conservatives were likely going to focus on such words in their analysis of the dictionary.

However, the lexicographers had the principles of their trade to consider as well. Since the purpose of the dictionary was to give a description of Shona, the expectation was that all the words used in the language were to be fully

documented in it. Excluding some words, and more so, categories of words, could lead to some form of deficiency of the work since it would leave out a portion of the language as it is used by its speakers. The result, therefore, would not be a truthful picture of the way the language is used. Excluding them would also be an act of ignoring the Shona corpus since most of these terms appear in it. The existence of these words in the Shona oral corpus is evidence enough that the terms are used in communication and therefore qualify them for inclusion in the dictionary. Some people would argue that including them in a dictionary is a way of licensing them for use, even by young children who are often cautioned to avoid them. The argumentation accompanying this view is that dictionaries are authoritative reference works and whatever is found in them is public knowledge that can be used or referred to at any time without restriction. However, real life experiences have shown that even when children are not taught or encouraged to use offensive words, somehow they end up using them, whether they appear in the dictionary or not. The fact that children use these words is also evidence that they are used in the society. Otherwise, how else could they have acquired this vocabulary? Eventually, the decision was that offensive terms should be included in the dictionary. However, the challenge then was how to describe the meanings of these words.

With regard to defining offensive headwords, a deliberate decision was made to avoid using explicit and direct terms in descriptions. Instead, the editors would use euphemism as much as possible, without, however, losing the actual essence of a word's meaning. As noted in Holder (1995), "euphemism means (the use of a) mild or vague or periphrastic expression as a substitute for blunt precision or disagreeable truth". Sinclair (1995) says that "if you describe language as euphemistic, you mean that it uses polite, pleasant, or neutral words and expressions to refer to things which people find unpleasant, upsetting or embarrassing". From this, it can be concluded that euphemism is a language of evasion and avoidance. It tends to evade and avoid the explicit and direct and advocates the implicit and indirect. However, its use was regarded as a way of being courteous to users and of showing respect for societal norms. This is because it comprises words and expressions specially created or crafted to substitute those explicit and direct terms that cannot be used in public.

It should now be considered how a few of the offensive words given as illustrative examples above were treated in DGC. As far as grammatical information is concerned, these offensive words were treated similarly to nonoffensive lemmas, giving tone (D, *dzikisa*, low tone, K, *kwidza*, high tone, D-, low tone throughout, K-, high tone throughout), part of speech (z, *zita*, noun) and noun class (indicated by numbers). However, the subsequent treatment was supposed to be as discreet as possible. The arrangement of the entries of the group of terms referring to sexual organs and activities was such that after the grammatical information following each headword, a style marker, *nyadzo*, a shortened form for *chinyadzo* (something that is coarse, vulgar or obscene), was introduced, which was then followed by the description of meaning. Where

they exist, synonyms were given after the definition, and were indicated by *FAN*, which is a truncated form for *-fanana* (same as). The style marker *nyadzo* was meant to be a warning to the reader that the term may be offensive. By giving such a warning, the editors were simply trying to be courteous to their readers. To illustrate this, the treatment of *mhata*, *mboro* and *beche* in DGC can be considered.

> **mhata** K- Z 9>10. (nyadzo) *Mhata* mwena uri pamuviri wemunhu, unowanikwa pasi panoperera muzongoza, unoshandiswa kuburitsa tsvina mudumbu. FAN *horo 9, mukosho 3, mudhidhi 3, mutinhi 3, mupedzazviyo 3*.
> (The anus is a hole in the body of a human being, found below the hips, through which faeces are removed from the stomach.)
>
> **mboro** K- Z 9>10. (nyadzo) *Mboro* inhengo yemuviri wemunhurume kana mhuka inoburitsa weti neurete. FAN *chirombo 7, mbonausiku 9, chombo 7, chirema 7*.
> (The penis is a part of the body of a male person or animal through which urine and sperm pass.)
>
> **beche** [mheche] K- Z 5>6 map-. (nyadzo) *Beche* inhengo yemunhukadzi inowanikwa pakati pemakumbo, uye ine buri, panova ndipo panozobuda napo weti kana mwana ava kuzvarwa, asi nechomukati zvichinge zvambobva kwakasiyana. FAN *sikarudzi 9, mukana 3, kunzira 17*.
> (The vagina is a part of a female person found between the legs and having a hole, which passes urine and a child during birth, although internally these come from different locations.)

To strike a balance between being explicit and being euphemistic when it came to defining these words, the editors agreed on a policy explained in the style manual of the dictionary: that the male and female sexual organs, for example, were to be defined in terms of their location on a person's body and their use for non- or less sensational functions such as urinating and child bearing. This is evidenced from the definitions provided for a penis and a vagina, for example. Whilst urinating is one of the main functions of these two organs, it may not be the one people immediately think of when hearing these terms. Generally, people think of sex and related activities, those that are deemed more obscene than urinating. However, although the editors may be blamed for omitting some of the functions of these organs (and mentioning others) the editors may not be accused of misrepresenting these organs. Although the definitions do not mention everything about the penis or vagina, no one can fail to identify the organs once the above descriptions have been read. In the definition of *penis*, for example, the lexicographer tried to make the definition quite clear and complete whilst also doing so as unsensationally and neutrally as possible. The lexicographer gives a definition that touches only on the essential and ordinary functions of the penis, that is, passing urine and sperm but omitting, for example, the crucial fact that it is also used for having sex. It is true

that by leaving this last function unmentioned, the lexicographer did not do enough justice to the definition of the term. However, this should be understood against the factors already discussed above. In describing the functions of this organ, more euphemisms were used. For example, the use of the term *urete* for sperm is a euphemism, the explicit term being *uronyo*. Lastly, the lexicographer provided more neutral and euphemistic synonyms. When, for example, one refers to a penis as *chirombo* (a huge creature), *mbonausiku* (something that sees at night) and *chombo* (a weapon) or to a vagina as *sikarudzi* (one that creates a family) or *mukana* (a passage), no one can visualise the sexual act in the same manner than when the terms *mboro* and *beche* are used. By providing some euphemistic near-equivalents, the lexicographers actually made available to the users of the dictionary a number of other possible terms that can be used in place of the ones considered coarse, vulgar or obscene.

Another category of offensive terms which posed challenges with regard to defining was that consisting of quite a substantial number of terms used to refer to groups of people generally regarded as inferior in society. Under this category, can be found those offensive words referring to people of foreign origin, the coloured community, albinos and cripples. To illustrate this, the word *mubwidi* (a bwidi) can be taken as an example. Some Shona speakers use this term to refer to people of Malawian origin, who came to Zimbabwe during the colonial era to work as migrant labourers on mines and farms. The term is believed to come from the way these people speak. There is a general belief amongst people who use this term that the language of the migrant labourers is not comprehensible or understandable to native Shona speakers. The explanation for the origin of this name is the idea that the speech of these people sounds as if they were only pronouncing a meaningless *bwidi bwidi bwidi*. Because of this, anyone who speaks in this way has become known as a *mubwidi*. Whether this is correct or not, the fact remains that this term is being used in some Shona-speaking communities to mock and disparage the way the *vabwidi* speak and it offends anyone being referred to as such. This is also reinforced by the fact that people of foreign origin are often concentrated in very poor communities, usually on mines and farms, where they provide unskilled labour.

Another example whose nuances are closely related to those of *mubwidi* is the term *mukarushu* which refers to people of Mozambican origin. These are people, usually young boys, who fled the Mozambican civil war in the early 1980s to come and work as cattle herders on communal lands in some parts of Zimbabwe. Here *karushu*, the cashew nut, commonly found in some parts of Mozambique, has been used to identify these cattle herders who feel disliked and unwelcome when referred to by this name. Yet another example is the term *murungudunhu*, which is used to refer to an albino. This is a compound noun formed by joining *murungu* (white man) and *dunhu* (village), that is, a white man of the village. Of the same origin is *murungudondo* (white man of the forest), which is another name for an albino. The connotations behind both these names are that a person whose colour is white but who has been born of

black parents are mockingly and disparagingly referred to as a white man who is usually seen as very superior, but is now described as originating from the village or the forest, in this way being accorded an inferior status.

After some of the connotations characterising the use of these words had been looked at, the way the editors of DGC tried to describe the terms *bwidi*, *mukarushu* and *murungudondo* can now be considered.

> **bwidi** D- Z 5>6. (tuko) *Bwidi* izwi rokutuka rinoshandiswa kureva munhu wechi-Malawi.
> (*Bwidi* is a derogatory term used to refer to a person of Malawian origin.)
>
> **mukarushu** DDKD- Z 1. (manje) *Mukarushu* izita rinopiwa vanhu vaitiza hondo kuMozambiki vachiuya muZimbabwe.
> (*Mukarushu* is a name for people who fled to Zimbabwe from the war in Mozambique.)
>
> **murungudondo** D- Z 1. (tuko) Uyu munhu ane ganda jena asi akaberekwa navabereki vatema zvichikonzerwa nourema hwokushaya utema muganda make. FAN *musope 1*.
> (This is a person whose skin colour is white but whose parents are black. The white skin colour is caused by deficiency of black pigmentation in the skin.)

Just as in the case of terms referring to sexual organs where the style marker, *nyadzo*, has been used, the style marker, *tuko* (a shortened form of *chituko*, insulting word), has been used in the above examples to warn readers that the terms have derogatory connotations. (This is with the exception of the style marker for slang, *manje*, which has erroneously been given at *mukarushu*.) However, as can be seen from the definitions provided for these terms vis-à-vis the explanations given for the origins of these terms, the definitions seem to be deficient since they omit much detail. The idea behind the provision of very brief descriptions was to be courteous to the groups of people being referred to. But it could be asked whether this courteousness would help any user understand how these terms are actually used in certain communities. Some would argue that being forthright would show some measure of insensitivity, but, at the same time, it can also be argued that omitting much detail as in the case of the above examples is counter-productive for it leaves out much needed information, especially about the reason(s) why the term(s) should be avoided. For example, describing *bwidi* only as a person from Malawi may not be a complete enough description since not all people from Malawi are *mabwidi*, this term only being used specifically to refer to the unskilled labourers who are now permanently residing in Zimbabwe and whose children can no longer be identified as Malawians. Trying to balance these two important aspects is the problem with which the editors of DGC had to struggle, yet it is a challenge with which every responsible lexicographer has to deal.

5.   Conclusion

This article has tried to highlight some of the challenges lexicographers face when it comes to the treatment of sensitive vocabulary in their dictionaries. It has attempted to describe some of the difficulties dictionary-makers sometimes experience when they need to strike a balance between the demands of their trade and the expectations of their readers. Drawing concrete examples from Shona, the article has shown that it is sometimes very difficult to follow lexicographic principles of accuracy, explicitness and precision when describing offensive terms such as those referring to sexual and excretory organs and their products, as well as some derogatory words. The article has indicated that good lexicographic practice moves between being theoretically sound and being culturally sensitive. It has shown that being theoretically sound may lead to a socially unacceptable product whilst being culturally sensitive may result in a poor reference tool. Although illustrative examples were drawn only from Shona and from the class of offensive words, the challenges discussed in this article may be applicable to all defining situations, though with possible variations regarding different cultures and lexicographic traditions.

References

Béjoint, H. 2000. *Modern Lexicography: An Introduction*. Oxford: Oxford University Press.

Chimhundu, H. (Ed.). 1996. *Duramazwi reChiShona*. Harare: College Press.

Chimhundu, H. (Ed.). 2001. *Duramazwi Guru reChiShona*. Harare: College Press.

Dubois J. and C. Dubois. 1971. *Introduction à la lexicographie: Le dictionnaire*. Paris: Larousse.

Holder, R.W. 1995. *Oxford Dictionary of Euphemisms*. Oxford: Oxford University Press.

Jackson, H. 1988. *Words and Their Meaning*. London/New York: Longman.

Kahari, G.P. 1996–97. Review of *Duramazwi reChiShona*. *Southern African Political and Economic Monthly* 10(3-4): 33-38, December–January.

Landau, S.L. 1984. *Dictionaries: The Art and Craft of Lexicography*. New York: Cambridge University Press.

Mberi, N.E. et al. 1997. Response to Professor G. Kahari's Review of *Duramazwi reChiShona*. *The Herald*, 20 March, Features Section: B2–B4, B7.

Mberi, N.E. et al. 1997. Was the Dictionary Review Motivated by Hate? *Southern African Political and Economic Monthly* 10(7): 21-24, April.

Sinclair, J. et al. (Eds.). 1995. *Collins COBUILD English Dictionary*. London: HarperCollins.

Svénsen, B. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making.* Oxford/New York: Oxford University Press.

Trench, R.C. 1857. On Some Deficiencies in Our English Dictionaries. *Transactions of the Philological Society*: 1-70. London: Parker.

Zgusta, L. 1971. *Manual of Lexicography*. The Hague: Mouton.

Reflections on the Proposed Ndebele–Shona/Shona–Ndebele Dictionary[*]

Samukele Hadebe, *African Languages and Literature, University of Zimbabwe, Harare, Zimbabwe (samukeleh@yahoo.co.uk)*

Abstract: The master plan of the ALLEX Project includes a Ndebele–Shona/Shona–Ndebele dictionary in its proposed dictionary projects. Bilingual dictionaries are common in Zimbabwe, especially earlier ones with the language pairs English–Ndebele/Shona and *vice versa*. The proposed Ndebele–Shona/Shona–Ndebele dictionary, however, raises some interesting challenges. It will be a different kind of bilingual dictionary in which two African languages, Ndebele and Shona form the language pair. In this article, it will be shown how different dictionary types for both Ndebele and Shona reflect the intentions of Zimbabwean language planners from different periods. A Ndebele–Shona/Shona–Ndebele dictionary, unimaginable to many, raises several questions, among others: Who needs such a dictionary? Who are the target users of such a dictionary? In addressing some of these questions, it will be attempted to show how the proposed Ndebele–Shona/Shona–Ndebele dictionary reflects the language planning needs of present-day Zimbabwe.

Keywords: BILINGUAL DICTIONARY, MONOLINGUAL DICTIONARY, NDEBELE, SHONA, NDEBELE–SHONA/SHONA–NDEBELE DICTIONARY, LANGUAGE PLANNING, LANGUAGE POLICY, LANGUAGE DEVELOPMENT, SOCIOLINGUISTICS, ATTITUDES, REFERENCE NEEDS, USER NEEDS

Opsomming: Gedagtes oor die voorgestelde Ndebele-Sjona/Sjona–Ndebelewoordeboek. Die totaalplan van die ALLEX-projek sluit 'n Ndebele–Sjona/Sjona–Ndebelewoordeboek by sy voorgestelde woordeboekprojekte in. Tweetalige woordeboeke is algemeen in Zimbabwe, veral vroeëres met die taalpare Engels–Ndebele/Sjona en andersom. Die voorgestelde Ndebele–Sjona/Sjona–Ndebelewoordeboek stel egter 'n aantal interessante uitdagings. Dit is 'n andersoortige tweetalige woordeboek waarin twee Afrikatale, Ndebele en Sjona, die taalpaar uitmaak. In hierdie artikel sal aangetoon word hoe verskillende woordeboeksoorte vir sowel Ndebele as Sjona die bedoelings van Zimbabwiese taalbeplanners van verskillende periodes weerspieël. 'n Ndebele–Sjona/Sjona–Ndebelewoordeboek, ondenkbaar vir baie, stel verskeie vrae, onder andere: Wie benodig so 'n woordeboek? Wie is die teikengebruikers van so 'n woordeboek? By die beskouing van sommige van hierdie vrae sal probeer word om aan te toon hoe die voorgestelde Ndebele–Sjona/Sjona–Ndebelewoordeboek die taalbeplanningsbehoeftes van die huidige Zimbabwe weerspieël.

Sleutelwoorde: TWEETALIGE WOORDEBOEK, EENTALIGE WOORDEBOEK, NDEBELE, SJONA, NDEBELE–SJONA/SJONA–NDEBELEWOORDEBOEK, TAALBEPLANNING, TAAL-BELEID, TAALONTWIKKELING, SOSIOLINGUISTIEK, HOUDINGS, NASLAANBEHOEFTES, GEBRUIKERSBEHOEFTES

1.   Introduction

The proposed Ndebele–Shona/Shona–Ndebele dictionary raises some inter-esting challenges, especially to its compilers. The average Zimbabwean is familiar with either monolingual English dictionaries or with bilingual Eng-lish–Shona/Ndebele dictionaries. The African Languages Lexical Project (AL-LEX), a joint University of Zimbabwe–University of Oslo research project, intends to compile a bilingual Ndebele–Shona/Shona–Ndebele dictionary. The prospect of such a dictionary is unimaginable to many. One question that most people might ask is: Who needs the bilingual Ndebele–Shona/Shona–Ndebele dictionary? Once the potential users are known and their needs established, it would be relatively easy to produce such a dictionary. In attempting to answer the question on whether the Ndebele–Shona/Shona–Ndebele dictionary is nec-essary or not, one should go beyond lexicography into the sphere of language planning. This article wants to show that the proposed Ndebele–Shona/Shona–Ndebele dictionary reflects the language planning needs of Zimbabwe.

2.   Background to the project

The master plan of the ALLEX Project includes a Ndebele–Shona/Shona–Nde-bele dictionary as one of its proposed dictionary projects. According to this master plan, the proposed dictionary would be compiled soon after the com-pletion of both the Shona and Ndebele monolingual dictionaries. The two dic-tionaries, *Duramazwi reChishona* and *Isichazamazwi SesiNdebele*, have since been completed and published in 1996 and 2001 respectively. Work on the proposed bilingual dictionary has not begun yet as the researchers at the African Lan-guages Research Institute (ALRI) have still been working on other projects such as the trilingual dictionary of musical terms, the Shona dictionary of linguistic and literary terms and the advanced Ndebele dictionary.

As perhaps everywhere else, dictionary-making in Zimbabwe more or less reflects the language development needs of the people. For example, the early dictionaries compiled mostly by missionaries were bilingual, that is, English to Ndebele/Shona and *vice versa*. These bilingual dictionaries mainly targeted Nde-bele and Shona speakers learning English and Europeans wanting to learn Ndebele and/or Shona. The types of dictionaries compiled could be linked to historical periods. During the colonial period chiefly bilingual dictionaries were produced where English had always been one of the languages. The post-independence period mainly saw the publication of the ALLEX monolingual Ndebele and Shona dictionaries. ALLEX dictionaries target mother-tongue

speakers and attempt to redress the scarcity of reference books in the African languages of Zimbabwe.

When the ALLEX Project initiated monolingual lexicography in African languages, it was applauded as a very important development for nation-building. Although there were sceptics who felt that too much money was being spent on African languages instead of on what they perceived to be more pressing concerns like natural science and technology, the people at large were nevertheless appreciative. The monolingual dictionaries, especially for Ndebele, were felt to be long overdue. The potential users and their reference needs and skills were presumed to be known by both compilers and society in general.

The inclusion of the bilingual dictionary in the ALLEX master plan for dictionaries implies that the importance of the dictionary was already felt at that very early stage of planning. At the time of planning this dictionary which were supposed to be compiled within the next fifteen years or so, its target users and their needs must have been presumed known. For laypeople the target users would simply be Ndebele speakers who want to learn Shona and *vice versa*. Lexicographers and other language specialists would consider this to be too broad a category to be useful for planning purposes. Unfortunately, the ALLEX master plan does not spell out these pertinent issues. In the absence of these, there can only be speculated on the intended objectives and purposes and potential users of the Ndebele–Shona/Shona–Ndebele dictionary. However, speculation is useful as it may throw light on a project that has great potential for Zimbabwe's language planning.

3.   The language situation in Zimbabwe

Language matters easily trigger emotions and controversy in Zimbabwe. The publication of the first Shona dictionary by ALLEX was met with fierce criticism from one reviewer, activating a national debate on language, lexicography and dialects. This event and many similar ones demonstrate that dictionaries are taken seriously in Zimbabwe.

Zimbabwe recognises Ndebele, Shona and English as the three main languages to be used in education, administration and the media. However, as Chimhundu (1997: 132) indicates, "the national and/or official status of Shona and Ndebele is largely theoretical, as very little, if anything, is being done officially to develop and promote them and to diversify their functions". English dominates the two African languages in education, administration, industry, commerce and the media. The preponderance of English at the expense of the African languages in Zimbabwe is not based on the number of speakers, but brought about by historical, political and economic circumstances privileging English.

English is the official medium of instruction in Zimbabwean schools from the fourth grade upwards. Ndebele is only taught as a subject in the two Mata-

beleland provinces and in some schools in the Midlands province. Shona is taught in the rest of the country.

Zimbabwe also has a number of indigenous African languages that have been designated as "minority" languages. Some of these, namely Tonga, Venda, Nambya, Kalanga and Shangani, are taught only up to grade four. Thereafter either Ndebele or Shona takes over, depending on whether the surrounding area is Ndebele- or Shona-speaking. There are also unrecognised indigenous languages, especially the Khoisan languages. Furthermore there are non-Zimbabwean African languages like Chewa and other non-African languages.

The language situation in Zimbabwe can be seen as a hierarchical structure with English at the top as official language, followed by Ndebele and Shona, with the "minority" languages lowest both in prestige and in official recognition. For the same reason, the book industry is by far dominated by English with Shona and Ndebele lagging behind in second and third position respectively. Little is published in the "minority" languages, apart perhaps from some Bible translations and hymn books.

4.   Language planning needs

Currently, the Zimbabwean government is working on a new language policy, which might change the status of a number of the languages. Naturally, Ndebele and Shona are aspiring for the same status as English, while the "minority" languages similarly claim recognition as national languages. Language policy refers to a wide range of governmental and non-governmental actions designed to influence language acquisition and language use (Tollefson 2000: 13).

Hopefully the new language policy will reflect the wishes of the majority of ordinary Zimbabweans as well as take some sociolinguistic and historical realities in cognisance. During the outreach of the Constitutional Commission (1999–2000), the people of Zimbabwe expressed the desire to have Ndebele and Shona accorded official language status. Some in Zimbabwe feel the country is lagging behind in promoting African languages considering that neighbouring South Africa has eleven official languages of which nine are African.

Zimbabwe could break with its colonial language policy structure where English plays the pivotal *lingua-franca* role amongst the many African languages. Bilingualism in Zimbabwe is biased in favour of English, English being on one side and an African language on the other. Shona speakers speak and write English and Shona while with Ndebele speakers it is English and Ndebele. This is what Chimhundu (1993: 59) describes as "an example of neglected multilingualism and unbalanced, culturally unhealthy bilingual behaviour by its (Zimbabwe's) citizenry".

This unbalanced bilingualism, although traceable from the language policies of the colonial era, unfortunately persisted after Zimbabwe attained its independence. "Unlike in most countries outside Africa," Wolff (2000: 342) says, "the new African elites prefer to use the foreign language for many functions

which are normally reserved for mother tongues or national languages." The same elites have been unwilling to promote African languages, even in cases where the advantages of doing so are obvious to all. Chimhundu (1997: 130) writes that in Zimbabwe "it would seem that during the first fifteen years of independence, government policy was not to formulate or pronounce a language policy as such". As already stated, this state of affairs is rapidly changing.

Some of the language controversies at issue in Zimbabwe have been on whether Ndebele and Shona should be compulsory subjects to all students. Questions have been asked, for example, whether pupils could write both Ndebele and Shona at Ordinary Level Examinations and have these counted as different subjects. At present, it is either Ndebele or Shona and English. A new language policy sensitive to the challenges of the people should facilitate Ndebele–Shona bilingualism for the whole nation. Such an important language policy initiative could be implemented effectively if there is government support complemented by the consensus of the people.

Starting in 2003, the government of Zimbabwe introduced the National Strategic Studies at tertiary educational level. This subject has been viewed by some as partisan and meant to enhance the electoral advantages of one political party. Unless it is used to address fundamentals like the language question, it would be difficult for this important subject to be taken seriously. In addition to its requirement for all graduates at government colleges, the two national languages, Ndebele and Shona, should be added too. The official justification for this subject is the teaching of the country's history with a view to fostering, among others, patriotism, moral values and national unity. Fostering these in a language barely used and understood by the majority is a recipe for failure. Alienating Ndebele and Shona languages and at the same time attempting to draw national pride from the same cultures and people is self-defeating.

The media, especially the radio and to a lesser extent television have been instrumental in fostering the country's language policies. Actually, in the recruitment of radio announcers, it is usually the case that, apart from the requirement of English, which is mandatory for all government jobs, fluency in either Shona or Ndebele and knowledge of another African language is considered an advantage. The Zimbabwe Broadcasting Corporation (ZBC) has radically changed its language policy in favour of African languages. First, it propounded a 75% local content on radio and television. This had a positive effect in terms of promoting local music, drama, theatre and indigenous languages in general. Then it has gone further by converting the Montrose studio in Bulawayo into a studio 100% broadcasting in indigenous African languages.

Zimbabwe has long discoursed on indigenisation. Language indigenisation should accompany if not preceding economic indigenisation. What ZBC has pronounced and partly effected now need to be spread throughout the civil service. The education language policy should also be changed to suit the new challenges. With these language-planning needs for Zimbabwe, then, the bilin-

gual Ndebele–Shona/Shona–Ndebele dictionary is not only necessary but long overdue. However, important factors have to be taken into consideration before the dictionary is compiled. At national level, as noted, the conditions are conducive to such a dictionary project but there are specific concerns to be addressed as well.

In compiling a dictionary such as the proposed Ndebele–Shona/Shona–Ndebele dictionary, both lexicographic and sociolinguistic factors are in play. The question of who the potential users are must be treated more as a sociolinguistic than a lexicographic factor. It has already been mentioned that the potential users are mother-tongue Ndebele and Shona speakers. Lexicographic factors to be considered are the age and educational level of the potential users. Also of interest are the dictionary skills of the potential users as all these factors have a direct and/or indirect bearing on the lemmatisation, defining styles, exemplification and the overall density of the lemma.

5.   Some sociolinguistic factors to consider

Any dictionary project should clearly identify its target users prior to the compilation phase (Mavoungou 2002: 185). What have to be taken into consideration are the attitudes of the potential users. Hartmann (2001: 25) says: "The dictionary maker may well structure the text with the dictionary user in mind, but more often than not lexicographic work proceeds very much within the conventions of a given tradition." While this could be the case with the Ndebele–Shona/Shona–Ndebele dictionary, it is very important not to ignore the users. Actually, as there has been no prior Ndebele–Shona/Shona–Ndebele dictionary, there is no tradition to follow. According to Chimhundu (1997: 142), "in such a setting, everything that the ALLEX Project is doing or attempting to do is indeed pioneering work".

The users cannot be treated as a homogenous group. Even if the dictionary were to be targeted at secondary pupils or at young adults at tertiary institutions, these would still have different needs. The cultural and language background are important factors to address. For example, do the Ndebele-speaking users have the same user needs as their fellow Shona-speaking users? If it could be established that the two groups differ significantly on this matter, then this would have implications for the structure of the dictionary.

Another important consideration is the possible attitudes of the users towards the dictionary which will most likely be influenced by the following:

—   The users' attitudes towards their mother-tongue;

—   The attitudes towards the other language;

—   The prevailing socio-political situation; and

—   The perceived benefits of learning the other language.

5.1   The users' attitudes towards their mother-tongue

The users' attitudes towards their mother-tongue could have an effect on their willingness to learn the other language. If, for instance, a Ndebele speaker has a high esteem for his/her language, that speaker is likely to have a respect for the Shona language, and *vice versa*. Speakers who have a low regard for their mother-tongue and African languages in general, would not be expected to view a Ndebele–Shona/Shona–Ndebele dictionary positively.

Also those who hold a puristic view towards their mother-tongue might not see the wisdom of having a bilingual dictionary. For example, there are some Ndebele speakers who felt the monolingual Ndebele dictionary *Isichaza-mazwi SesiNdebele* (ISN) had unnecessarily lemmatised loanwords from Shona when there were equivalent words in Ndebele. The fear that Shona might dominate Ndebele may have a negative influence as the dictionary might be perceived as a means to spread Shona among the Ndebele people. The reverse would conveniently be forgotten.

5.2   The attitudes towards the other language

It is often said that one's attitude towards a language usually reflects one's attitude towards speakers of that language. Various non-linguistic factors such as one's education, experience and background, can influence one's attitude towards the other group and its language. As has been mentioned earlier, each historical period creates dictionaries characteristic of it. There were periods in Zimbabwe's history where such a dictionary would have met with resistance on both sides. As the situation has since improved, the ethnic relations are such that the proposed dictionary might win more support now than at any previous time.

5.3   The prevailing socio-political situation

The attitudes people hold are partly shaped and reinforced by the social, economic and political environment. Perceived differences are wider when there are visible conflicts and competition for resources by various groups.

The political rhetoric on the ground would actually favour any project that could be perceived as enhancing indigenous cultures and fostering national unity. The indigenisation of the economy has been accompanied, theoretically at least, with the policy of 75% local content on national radio and television. This local content policy favouring the growth of indigenous languages makes the compilation of the Ndebele–Shona/Shona–Ndebele dictionary timely in Zimbabwe.

5.4   The perceived benefits of learning the other language

A language is seldom learnt for its own sake. In most cases, there must be per-

ceived social, economic or political benefits for learning a language. All parents in Zimbabwe want their children to learn English because it is the way to higher education, better jobs and many other opportunities. English seems to be destined to enjoy this state of affairs for a long time to come. This is caused by the official status of English in Zimbabwe and its perceived status world-wide.

It would not be enough to urge bilingualism in Ndebele and Shona for purely national unity and societal coherence. While these are noble causes at national level, there is a need for immediate rewards at individual level. After all, Tollefson (2000: 10) says: "Learning a language is an extraordinarily difficult, complex, and time-consuming task, not to be undertaken lightly." If certain jobs in the civil service like the police, army, teaching, nursing, customs and immigration, for example, were to require a knowledge of the three languages Ndebele, Shona and English, this would be an important reason for people to learn them.

6.   Some lexicographic factors to consider

The lexicographic issues to consider are discussed under the following subtitles:

—   Bilingual or bilingualised dictionary

—   Equivalence of cultural terms

—   Corpus-based dictionary

—   User-friendliness

6.1   Bilingual or bilingualised dictionary

According to the *Dictionary of Lexicography* (Hartmann and James 1998: 14), a bilingual dictionary "relates the vocabularies of two languages together by means of translation equivalents", whereas a bilingualised dictionary "is based on a monolingual dictionary whose entries have been translated in full or in part into another language".

The compilers of the proposed dictionary should decide from the outset whether their aim is a bilingual or a bilingualised dictionary. This is a crucial decision because the dictionary has of necessity to follow the already published monolingual Ndebele and Shona dictionaries which will naturally form the bases of the proposed bilingual Ndebele–Shona/Shona–Ndebele dictionary.

6.2   Equivalence of cultural terms

In selecting and in defining entries, it is unlikely that serious problems of equivalence could arise regarding basic words and terms, for instance those

used in teaching grammar and literature. When, however, it comes to cultural terms or those peculiar to certain practices, caution should be exercised.

Although Ndebele and Shona cultures have a lot in common, their differences should be taken into consideration. Although these differences may not be many, they could establish or destroy the reputation of the dictionary depending on how they are handled.

6.3   Corpus-based dictionary

ALLEX dictionaries are corpus-based or corpus-aided. If the proposed dictionary has to follow this tendency, parallel corpora might need to be built to assist in the compilation. At the moment, there are no parallel Ndebele–Shona corpora yet, although building them would not be difficult. Some books which have been translated into both Ndebele and Shona might form the basis of these corpora. Also, the Bible translations of both languages have been scanned and included in the already existing ALLEX corpora.

The editors would have to be careful when using the corpora not to rely entirely on translated language. Similarly, caution has to be taken to avoid cultural controversies as mentioned under 6.2 above.

6.4   User-friendliness

From the planning stage, the users have to be clearly defined and their reference needs and reference skills known. As dictionary skills are rarely taught explicitly in Zimbabwean schools, reference skills are generally low.

In compiling the dictionary, editors must keep in mind that strategies used in previous dictionaries might not be relevant to this one. Similarly, it must not be assumed that potential users of the bilingual dictionary have had access to the monolingual Ndebele and Shona dictionaries. The success of this dictionary partly rests on its attunement to user needs and its accessibility to users. In this sense, user-friendliness refers to a combination of factors perceived to enable users easy access to the dictionary.

7.   Recommendations to compilers

From the foregoing, it can be noted that the compilation of the proposed Ndebele–Shona/Shona–Ndebele dictionary is fraught with lexicographic and non-lexicographic considerations. Although the editors of this dictionary would obviously be experienced lexicographers, serious attention has to be given to every aspect of the dictionary. Unlike the other ALLEX dictionaries whose target users were clearly defined, the proposed dictionary necessitates a study of the actual reference needs and reference skills of the potential users. The compilation of the dictionary cannot be based on assumptions about its potential users.

This background information could prove valuable in determining the size and scope of the dictionary as well as its style of presentation. Perceived as a pedagogical dictionary, the proposed Ndebele–Shona/Shona–Ndebele dictionary would be more useful as it would assist Ndebele speakers learning Shona and *vice versa*. A pedagogical dictionary is designed "for the practical needs of teachers and learners of a language" (Hartmann and James 1998: 107).

Lastly, sociolinguistic considerations should be constantly kept in mind throughout the compilation stage as this dictionary has a great potential for language planning needs in Zimbabwe. For this reason, controversial words and senses should be avoided as much as possible. The Shona and Ndebele monolingual dictionaries were embraced by their respective language communities as important symbols of community achievement. The bilingual dictionary needs a similar positive attitude from both language communities if it is not only to be viable commercially but also to be utilised by its intended users.

8.   Conclusion

The proposed Ndebele–Shona/Shona–Ndebele dictionary poses a number of challenges, some of which were highlighted in this article. The challenges are not insurmountable, although addressing them would call for the consideration of a combination of lexicographic, linguistic and sociolinguistic factors.

From the above reflections, the national significance of the Ndebele–Shona/Shona–Ndebele dictionary would have become clear. The successful completion of the proposed dictionary would be a landmark in Zimbabwean lexicography and a crucial contribution of lexicography towards language policy and language planning. It would be a turning-point for the African languages in Zimbabwe, as this would mark their reduced dependency on English. It has therefore been demonstrated that Zimbabwean society indeed needs the Ndebele–Shona/Shona–Ndebele dictionary.

References

Chimhundu, H. 1993. The Status of African Languages in Zimbabwe. *Southern African Political and Economic Monthly* 7(1): 57-59, October.

Chimhundu, H. 1997. Language Standardization without Policy or Planning: Zimbabwe as a Case Study. Royneland, U. (Ed.). 1997. *Language Contact and Language Conflict*: 129-150. Volda: Volda College.

Chimhundu, H. (Ed.). 1996. *Duramazwi reChiShona*. Harare: College Press.

Hadebe, S. et al. (Eds.). 2001. *Isichazamazwi SesiNdebele.* Harare: College Press.

Hartmann, R.R.K. 2001. *Teaching and Researching Lexicography*. Essex: Pearson Education Ltd.

Hartmann, R.R.K. and Gregory James. 1998. *Dictionary of Lexicography*. London: Routledge.

Mavoungou, P.A. 2002. Synopsis Articles in the Planning of a Trilingual Dictionary: Yilumbu–French–English. *Lexikos* 12: 181-200.

Tollefson, J.W. 2000. Policy and Ideology in the Spread of English. Hall, J.K. and W.G. Eggington (Eds.). 2000. *The Sociopolitics of English Language Teaching*: 7-21. Clevedon: Multilingual Matters.

Wolff, H.E. 2000. Language and Society. Heine, B. and D. Nurse (Eds.). 2000. *African Languages: An Introduction*: 298-347. Cambridge: Cambridge University Press.

Meaning Discrimination in Bilingual Venda Dictionaries*

Munzhedzi James Mafela, *University of South Africa, Pretoria, Republic of South Africa (mafelmj@unisa.co.za)*

Abstract: Venda, one of the minority languages in South Africa, has few dictionaries. All are translational bilingual dictionaries meant for dictionary users who are non-native speakers of the language. Dictionary users find it difficult to use the bilingual Venda dictionaries because they are confronted with equivalents which they cannot distinguish. In most cases, the equivalents of the entry-words are provided without giving meaning discrimination. Without a good command of Venda and the provision of meaning discrimination, users will find it difficult to make a correct choice of the equivalent for which they are looking. Bilingual Venda dictionaries are therefore not helpful for dictionary users who are non-native speakers of the language. Devices such as giving illustrative examples, indicating parts of speech and adding etymology could be used to solve the problem of meaning discrimination in bilingual Venda dictionaries. This article highlights the problem of the absence of meaning discrimination and suggests solutions to future Venda lexicographers in this regard.

Keywords: BILINGUAL DICTIONARY, MEANING DISCRIMINATION, EQUIVALENCE, ENTRY-WORD, LEXICOGRAPHY, CULTURE, TRANSLATION, SOURCE LANGUAGE, TARGET LANGUAGE, SYNONYM, POLYSEMY

Opsomming: Betekenisonderskeiding in tweetalige Vendawoordeboeke.
Venda, een van die minderheidstale in Suid-Afrika, het min woordeboeke. Almal is vertalende tweetalige woordeboeke bedoel vir woordeboekgebruikers wat nie moedertaalsprekers van die taal is nie. Woordeboekgebruikers vind dit moeilik om tweetalige Vendawoordeboeke te gebruik omdat hulle gekonfronteer word met ekwivalente wat hulle nie kan onderskei nie. In die meeste gevalle word die ekwivalente van die trefwoorde verskaf sonder om betekenisonderskeiding aan te dui. Sonder 'n goeie beheersing van Venda en die aanduiding van betekenisonderskeiding, sal gebruikers dit moeilik vind om die regte keuse van die ekwivalent waarvoor hule soek, te maak. Tweetalige Vendawoordeboeke is dus nie van hulp vir woordeboekgebruikers wat nie moedertaalsprekers van die taal is nie. Middele soos die gee van ophelderende voorbeelde, die aantoon van rededele en die byvoeging van etimologie kan gebruik word om die probleem van betekenisonderskeiding in tweetalige Vendawoordeboeke op te los. Hierdie artikel beklemtoon die probleem van die afwesigheid van betekenisonderskeiding en dien oplossings in dié verband aan die hand vir toekomstige Vendaleksikograwe.

---

* This article is a slightly revised version of a paper presented at the Eleventh EURALEX International Congress held at the Université de Bretagne-Sud, Lorient, France, 6–10 July 2004.

Sleutelwoorde: TWEETALIGE WOORDEBOEK, BETEKENISONDERSKEIDING, EKWI-VALENSIE, TREFWOORD, LEKSIKOGRAFIE, KULTUUR, VERTALING, BRONTAAL, DOEL-TAAL, SINONIEM, POLISEMIES

1.   Introduction

A bilingual dictionary entails the description of two cultures because it deals with the translation of entry-words from the source language to the target language. Bilingual dictionaries may serve both target language and source language speakers. When a target language speaker wants to comprehend a foreign language he/she will consult a bilingual dictionary, and when a source language speaker wants to express him-/herself in a foreign language, he/she will also consult a bilingual dictionary. As Swanson (1983: 193) puts it, a bilingual dictionary can be useful and desirable to several kinds of people: students, travellers and linguists. According to Zgusta (1971: 294), the basic purpose of a bilingual dictionary is to co-ordinate the lexical units of one language with those of another language which are equivalent in lexical meaning. In many instances such co-ordinations are accompanied by problems on the part of dictionary users, especially if they are not native speakers of the target language. A user is less likely to know the meanings of foreign language words in the entries, and will therefore need information on how these foreign language words differ in meaning. Gouws (2000: 102) writes that this problem is a direct result of the traditional approach which sees a bilingual dictionary as a mono-functional product in which the treatment is restricted to the mere listing of a number of translation equivalents.

The translation equivalents of entry-words in a bilingual dictionary are usually of two types, i.e. translational and explanatory. A translational equivalent is a lexical unit which can immediately be inserted into the language, whereas an explanatory equivalent is of a general nature functioning well if the target language is the user's native tongue because it may suggest or elicit in him/her some other equivalent which fits the particular context with which he/she is dealing (Al-Kasimi 1983: 60-61). Translational equivalence is favoured in a bilingual dictionary intended for speakers of the source language who want to express themselves in or translate into the foreign language.

Venda, one of the languages which were previously marginalised in South Africa, has very few dictionaries. Linguists did not pay much attention to Venda lexicography. In the past, the cultivation and development of a dictionary culture in the indigenous languages was minimal. in the past. As Van der Merwe (2003: 183) writes:

> In order to cultivate and develop a dictionary culture in a country the existence of lexicography has to be recognised as a subject field. Lexicographical research has to lay the foundation for sound dictionary projects. Lexicographers have to be trained to compile user-friendly dictionaries that are theoretically sound.

Hence there is a need to cultivate and develop a dictionary culture in Venda. At present, Venda has three English–Venda dictionaries (*Phindulano: English–Venda Phrase Book, Phrase Book for English and Venda*, and *Dictionary of Basic English–Venda)*, one Afrikaans–Venda dictionary (*Afrikaans–Venda Vocabulary and Phrase Book),* one Venda–English dictionary (*Venda Dictionary: Tshivenḓa–English)* and one Venda–Afrikaans–English dictionary (*Improved Trilingual Dictionary: Venda–Afrikaans–English).* All the above-mentioned are bilingual dictionaries, except one (*Improved Trilingual Dictionary: Venda–Afrikaans–English)* which is partly bilingual and partly trilingual. The equivalents of entry-words in almost all of them are translational. With the exception of one reference dictionary (*Venda Dictionary: Tshivenḓa–English)*, all are meant for foreigners, mainly missionaries (*Phindulano: English–Venda Phrase Book, Phrase Book for English and Venda, Afrikaans–Venda Vocabulary and Phrase Book)* and students (*Improved Trilingual Dictionary: Venda–Afrikaans–English),* who know very little or nothing at all about the Venda language or culture.

The objective of this article is to highlight the problem of the absence of meaning discrimination for the equivalents of the entry-words in the target language and the effect this has on dictionary users with reference to bilingual dictionaries in Venda. Suggestions about devices which can help solve this problem will be provided in the concluding remarks of the article.

2.   The choice of equivalents and equivalence discrimination

An equivalent is a word or phrase in one language which corresponds in meaning to a word or phrase in another language (Prinsloo and De Schryver 2002: 162). Al-Kasimi (1983: 58) sees the major task of a bilingual lexicographer to find appropriate equivalents in the target language to the units of the source language. Some Venda dictionaries, especially early ones, provide users with word-for-word equivalents that comprise one source language entry and one equivalent in the target language. Hereunder are examples from the *Phrase Book for English and Venda (*Marole 1932).

   (1) kidney   tswiyo
   (2) elbow   lukuḓavhavha
   (3) nail   ṋala

Newman (1980: 41) sees word-for-word equivalence as an ineffective second language learning method. However, a bilingual dictionary that does not provide more than one equivalent in the target language does not become a problem to the users because it does not involve meaning discrimination. Users of the *Phrase Book for English and Venda* do not encounter difficulty in choosing the equivalent of example (1). The equivalent of *kidney* is *tswiyo* in Venda. *Tswiyo* does not have synonyms or near synonyms. However, examples (2) and (3) can lead the user to wrong choices of equivalents because the entries in the source

language are polysemous, and only one translation equivalent has been provided in each case. Both *elbow* and *nail* should have more than one equivalent characterised by different meanings. For example, the entries elbow and nail are defined in *The Concise Oxford Dictionary of Current English* (1990) as follows:

> *elbow:* (a) joint between the forearm and the upper arm, (b) a short piece of piping bent through a right angle;

> *nail:* (a) a small usu. sharpened metal spike with a broadened flat head, (b) horny covering on the upper surface of the tip of the human finger or toe, (c) fasten with nails.

In the entry *elbow*, the Venda equivalent is *lukuḍavhavha* (joint between the forearm and the upper arm); and in the entry *nail*, the equivalent is *ṋala* (horny covering on the upper surface of the tip of the human finger or toe). The other meanings of the entry-words have been omitted. If the user is looking for an equivalent of an entry-word which is found in a different context than the equivalent provided, then he/she will make a wrong choice of the equivalent. The provision of more equivalents, accompanied by equivalent discrimination is vital in this situation.

In some Venda bilingual dictionaries, users are confronted with several equivalents for one entry. They cannot discriminate between the equivalents because of the absence of the necessary information for the purpose of meaning discrimination. They therefore fail to make a correct choice of the equivalent. The notion of equivalent discrimination applies to all dictionaries that present one or more translation equivalents as part of the lexicographic treatment of the lemma sign (Gouws 2000: 99). As Al-Kasimi (1983: 67) writes:

> When a person wants to say something in a foreign language, he might consult a bilingual dictionary. But instead of finding one word which expresses his meaning, he is frequently confronted with several words which he cannot distinguish one from another.

The equivalents themselves are the most important part of the entry. However, in the majority of cases, it does not suffice to indicate them alone, firstly because most of them have multiple meanings themselves, and secondly because they are only partial equivalents of the entry-word (Zgusta 1971: 329). Without additional information, the dictionary user will not be in a position to select the correct equivalent which fits the meaning in the context, unless he/she has a command of the target language.

The following are examples of dictionary entries from three English–Venda bilingual dictionaries which will assist in illustrating the above-mentioned problem. Examples (4)–(11) are from the *Improved Trilingual Dictionary*, examples (12)–(16) from the *English–Venda Vocabulary*, and examples (17) and (18) from the *Afrikaans–Venda Vocabulary and Phrase Book*.

(4) ape (n) (monkey)   ṱhoho(dzi); (vb) -edzisa

(5) amputate   -tumula, -ṱhukhula
(6) abduct   -ṱahisa (musadzi); -hwala
(7) abstain   -ḓidzima
(8) bellow   -kuma
(9) buffalo   ṋari
(10) console   -fhembeledza, -khuthadza, -fhumudza, -lilisa
(11) difficulty   -vhuleme (abstr.), tshikundisi (zwi)
(12) abandon (n)   laṱa, u furalela
(13) swarm (n)   murivha, gogo
(14) act (v)   mushumo, mulayo
(15) boundary (n)   mukano, ṱhanganyoni
(16) teacher (n)   mudededzi, mufunzi
(17) baard   dzindebvu, vhuṱambo
(18) berg   thavha

From the examples cited above, it can be noticed that some entries have one equivalent each, i.e. examples (7), (8) and (9), while the others have more than one equivalent each. In the case of entries having one equivalent each, the dictionary user does not have a problem of distinguishing equivalents in Venda because there is only one meaning. In the case of entries having more than one equivalent, it is difficult for the source language dictionary user to select the correct equivalent relevant to the context because of multiple meanings which are synonyms or near synonyms.

In example (4), the entry *ape* is a noun and a verb at the same time. Although the compiler has given one part of speech, the entry-word has equivalents both as a noun and a verb in Venda. The parts of speech (n) and (vb) have been used to distinguish the meanings of the equivalents. As Iannucci (1983: 179) noted, meaning discrimination is sometimes effected by the designation of the part of speech of the entry word. In this example, it is easy for the dictionary user to select the correct equivalent because there is only one noun and one verb in the definition, i.e. *ṱhoho* (monkey) and *-edzisa* (imitate). However, in examples (5) and (6), the dictionary user will have some difficulty in distinguishing the meanings of the given equivalents. When a user consults a bilingual dictionary, he/she wants to achieve an unambiguous identification and interpretation of the target language data and consequently the optimal retrieval of the target language information (Gouws 2000: 101). This is not possible in the examples given above. The lack of the inclusion of meaning discrimination impedes the user to choose the correct equivalent. In example (5), both equivalents *-tumula* and *-ṱhukhula* are synonymous verb stems. The verb stem *-tumula* is associated with cutting something with an instrument such as a knife, whereas the verb stem *-ṱhukhula* is associated with breaking something to

pieces (e.g. a rope) by way of pulling. The selection of the equivalent will therefore depend on the context. As Gouws (2000: 110) mentions, users often need a co-text or other contextualising information to achieve equivalents. He further suggests that to ensure this, the lexicographer has to complement the translation equivalents with ample additional data. Without this additional data, it would be difficult for the user to select the correct equivalent. Meaning discrimination in this regard will involve the power of observing differences between meanings of equivalents of the same entry-word.

In example (6), *abduct* is defined by *The Concise Oxford Dictionary of Current English* (1990) as 'carry off or kidnap (a person) illegally by force or deception'. Unlawfulness is an essential element of abduction as evidenced in the definition (Mtuze 1990: 30). The equivalents of this entry-word, *-ṱahisa* and *-hwala,* are both verb stems. Their meanings are not related. The stem *-ṱahisa* means 'to cause to elope', whereas the stem *-hwala* generally means 'to carry something'. When a girl is caused to elope, she might go of her own volition, especially if there is some consent between her and the man. The girl may also be forced to accompany the man. Even if there were an agreement between the girl and the man, the process of eloping is considered an improper marriage procedure in Venda society. The context in which the entry-word *abduct* is used, will therefore be the guideline in the choice of the correct equivalent because *-ṱahisa* and *-hwala* have different meanings. Meaning discrimination is therefore important in order to distinguish between the two equivalents which belong to the same part of speech and are near synonyms. For the user to be able to make the correct choice of translation equivalents in this situation, he/she has to be familiar with the Venda language and culture.

Worse still is the situation when one encounters an entry-word with four or more equivalents which all belong to the same word category and are all synonymous, as in example (10). The equivalents given are all verb stems.

-fhembeledza	persuade, coax
-khuthadza	appease; pacify
-fhumudza	silence; comfort, console
-lilisa	make cry; console

The stem *-fhembeledza* has a different meaning from the other equivalents. When one persuades a person, he/she does not console him/her. The word *persuade* is defined by *Webster's New World Dictionary* as 'to cause to do or believe something, especially by reasoning and urging'. Although the stems *-fhumudza, -khuthadza* and *-lilisa* are related, their meanings are not exactly the same. For example, the other meaning of *-lilisa* is 'to make to cry'. Dictionary users will find it difficult to choose the correct equivalent. The selection of the correct equivalent will depend on the context, and this will be elucidated by the provision of meaning discrimination.

In example (12), the entry-word *abandon* has two equivalents in Venda, *-laṱa* and *-furalela* which are both verb stems. Van Warmelo (1989: 61) translates *-furalela* as 'turn one's back upon', whereas Wentzel and Muloiwa (1982: 17) translate *-furalela* as 'turn one's back upon; ignore'. On the other hand, *-laṱa* means 'to throw away, discard or abandon'. From the definitions of the two equivalents shown above, it is clear that the equivalents are not absolutely synonymous. The same applies to example (13), *swarm*, which has two near synonymous equivalents, *murivha* and *gogo*, both nouns. The noun *murivha* refers to a group of birds flying together, while the noun *gogo* refers to many people gathered in one place. These meanings are not interchangeable, however. Therefore, the lack of additional information to define the equivalents impedes the correct choice within a given context.

The dictionary user will find example (14) problematic. The entry-word *act* has *mushumo* and *mulayo* as its equivalents. *Act* can be used as a noun or as a verb. However, it is only shown as a verb, but all its Venda equivalents are nouns, their word category not being indicated. *Mushumo* refers to 'work' or 'function', whereas *mulayo* refers to 'rule of law'. The context in which the entry-word *act* is used will guide the dictionary user to the choice of the correct equivalent. But without meaning discrimination, it will be difficult for a dictionary user to make the correct choice. The equivalent verb stems of the entry-word *act* in Venda should have been *-tamba* (play), *-edza* (mimic or imitate). However, these equivalents are not included in the definition of the entry-word. The incorrect entries provided will therefore mislead the user in his/her choice of the equivalents.

In example (15), *mukano* is a dividing line (boundary), whereas *ṱhanganyoni* refers to the intersection (e.g. of rivers). The equivalent *ṱhanganyoni* cannot be regarded as a boundary without considering the context in which it is used. In example (16), *mudededzi* refers to a teacher who teaches pupils or students in class, whereas *mufunzi* is a preacher in the church. Both, however, are near synonymous equivalents of the entry-word *teacher*. As with example (13), it would be difficult to make a correct choice of an equivalent without the necessary information about the two equivalents. For the dictionary user to choose correctly between these translations, meaning discrimination is important.

As in the dictionaries considered so far, co-ordination of source and target language items can impede the correct choice of a translation equivalent in the *Afrikaans–Venda Vocabulary and Phrase Book*. Users of this bilingual dictionary do not encounter difficulty in the choice of a translation equivalent in example (18), because the relation is characterised by one entry (i.e. *berg* which means 'mountain') and one equivalent in Venda (i.e. *thavha)*. There are no synonyms or near synonyms of the equivalent *thavha*. However, the user could find it difficult to make a correct choice in example (17). The entry-word *baard* is given its corresponding Venda equivalents *dzindebvu* and *vhuṱambo,* both near synonymous nouns. *Dzindebvu* is the plural form of *ndebvu*, which refers to *beard* (hair growing on the lower part of a man's face). On the other hand, *vhuṱambo* is a polysemous word which in this instance refers to pubic hair. No additional

information which would enable the dictionary user to discriminate these meanings has been provided for these equivalents. Therefore, the dictionary user may make a wrong choice for the equivalent of the entry-word *baard.* Both equivalents refer to hair, but hair found on different parts of the body. Furthermore, *dzindebvu* is a characteristic of males, whereas *vhuṱambo* is a characteristic of both males and females. Without additional information the user may choose the wrong translation equivalent for a particular context.

The above discussion of examples from three different dictionaries reveals that meaning discrimination is very important in bilingual dictionaries. As Gouws (2000: 99) accentuates, the more target language information included in a dictionary, the better the chances should be of an unambiguous transfer of information.

3.   Possible solutions

Dictionary users need entries which cover all the meanings they require, i.e. meanings which provide them with the necessary information to make appropriate well-informed lexical choices (Scholfield 1999). The ideal bilingual dictionary would anticipate every conceivable need of the prospective user. It would provide for each word or expression in the source language just the right translation in the target language, including, most importantly, the one needed for a particular passage in hand (Hartmann 1999: 7). As indicated above, Venda bilingual dictionaries are not helpful in this regard. The treatment of meaning discrimination in Venda bilingual dictionaries is unsatisfactory. In order to understand sentences, users need to know the meaning of words. The provision of more information about the equivalents in the target language is important, as stressed by Gouws (2000: 104):

> The average user of bilingual dictionaries needs much more information in the comment on semantics to utilise the dictionary as a practical instrument. Lexicographers will have to enhance the possibilities for equivalent discrimination.

This additional information will enable dictionary users to interpret the meaning of utterances in the context in which they are made.

According to Al-Kasimi (1983), meaning discrimination is necessary in the following cases:

(a)   When the source language has one meaning for which the target language has a polysemous equivalent.

(b)   The source language is polysemous, and for each of its senses the target language has two or more polysemous words.

Additional information can be given in different ways. To achieve meaning discrimination, use can be made of different devices such as providing short

definitions of equivalents, indicating parts of speech, adding etymology and usage labels, and giving context words and phrases, and illustrative examples. Of these, illustrative examples, parts of speech and etymology are vital devices for meaning discrimination in bilingual dictionaries.

Examples can be used to illustrate the meaning of the word defined. According to Al-Kasimi (1983: 91), the primary function of illustrative examples in dictionaries in general and bilingual dictionaries in particular is to capture the user's interest by showing the word in a live context, and to enhance his/her understanding of the grammatical and semantic rules governing the usage of the word by showing these rules in action. These illustrative examples can further give the dictionary user some notion of the foreign culture he/she is encountering. Illustrative examples, phrases and sentences can be utilised to show how the equivalents are used in context.

The designation of the part of speech of both the entry word and the equivalents can also help to distinguish between the equivalents. Jackson (1985: 55) says:

> If a dictionary gives no other information of grammatical nature, it is expected to indicate which part-of-speech or word class a lexical item belongs to, i.e. whether it is classed as a noun or as a verb or adjective, etc.

Although the word-class label does not provide sufficient information of a grammatical kind, it provides basic information about the syntactic operation of a lexical item. The grammatical identity of words such as noun, verb, adjective, etc., is given for all entries and derivatives to aid clarity.

Etymology is another device which can help users of bilingual dictionaries to obtain a clear understanding of the present meanings of words. In this case, the historically variable sources of the formation of a word and the development of its meaning are provided.

4.   Conclusion

With the introduction of lexicography units for the indigenous languages in South Africa by the Pan South African Language Board, it is hoped that future lexicographers of the indigenous languages in general and Venda in particular will attend to the problem of meaning discrimination in bilingual dictionaries to make them more useful to users. This could be achieved by training lexicographers to compile user-friendly dictionaries which will be of value to both students, travellers and linguists. As indicated above, the present Venda bilingual dictionaries are not user-friendly. Al-Kasimi (1983: 68) accentuates:

> The bilingual dictionary should provide meaning discriminations which enable the user to select the appropriate equivalent or the proper sense of an equivalent. And unless the problem of meaning discrimination is solved systematically, the bilingual dictionary cannot be a dependable guide to the proper equivalents.

Attempts to achieve equivalent discrimination will increase the communicative success of the bilingual dictionary.

References

Dictionaries

Allen, R.E. (Ed.). 1990. *The Concise Oxford Dictionary of Current English.* Oxford: Clarendon Press.

Guralnik, D.B. (Ed.). 1981. *Webster's New World Dictionary of the American Language.* New Jersey: Simon and Schuster.

Marole, L.T. 1932. *Phrase Book for English and Venda.* Sibasa: Marole Book Depot.

Marole, L.T. 1954. *English–Venda Vocabulary.* Sibasa: Marole Book Depot.

Marole, L.T. 1955. *Afrikaans–Venda Vocabulary and Phrase Book.* Sibasa: Marole Book Depot.

Van Warmelo, N.J. 1989. *Venda Dictionary: Tshivenda–English.* Pretoria: J.L. van Schaik.

Wentzel, P.J. and T.W. Muloiwa. 1982. *Improved Trilingual Dictionary: Venda–Afrikaans–English.* Pretoria: University of South Africa.

Other Sources

Al-Kasimi, A.M. 1983. *Linguistics and Bilingual Dictionaries.* Leyden: E.J. Brill.

Gouws, R.H. 2000. Strategies in Equivalent Discrimination. Mogensen, J.E. et al. (Eds.). 2000. *Proceedings of the Ninth International Symposium on Lexicography*: 99-111. Tübingen: Max Niemeyer.

Hartmann, R. 1999. Lexical Reference Books — What Are the Issues? *National Journal of Lexicography* 12(1): 5-12.

Iannucci, J.E. 1983. Meaning Discrimination in Bilingual Dictionaries. *Linguistics Honours: Volume 3: Equivalence, Bilingual Dictionaries and Translation*: 177-192. Pretoria: Unisa.

Jackson, H. 1985. Grammar in the Dictionary. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 53-59. Oxford: Pergamon Press.

Mtuze, P.T. 1990. Problems of Equivalence and Adequacy in a Trilingual Context. *South African Journal of Linguistics* 8(1): 30-32.

Newman, A. 1980. *Mapping Translation Equivalence.* Louvain: Acco.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2002. Reversing an African-Language Lexicon: The *Northern Sotho Terminology and Orthography No. 4* as a Case in Point. *South African Journal of African Languages* 22(2): 161-185.

Scholfield, P. 1999. Dictionary Use in Reception. *International Journal of Lexicography* 12(1): 13-34.

Swanson, D.C. 1983. The Selection of Entries for a Bilingual Dictionary. *Linguistics Honours, Volume 3: Equivalence, Bilingual Dictionaries and Translation*: 193-207. Pretoria: Unisa.

Van der Merwe, M. 2003. Towards the Creation of a Dictionary Culture in South Africa. De Schryver, Gilles-Maurice (Ed.). *Sixth International TAMA Conference: Conference Proceedings*: 183-188. Pretoria: (SF)[2] Press.

Zgusta, L. 1971. *Manual of Lexicography.* The Hague/Paris: Mouton.

The Impact of Translation Activities on the Development of African Languages in a Multilingual Society: *Duramazwi reMimhanzi* as a Case-study[*]

Gift Mheta, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe (gmheta@yahoo.com)*

Abstract: The article examines the impact of translation activities on the development of African languages in the multilingual Zimbabwean society. It analyses Shona musical terms created through translation processes and strategies such as borrowing, coining, compounding and derivation. Focus is on the way this ongoing term-creation is contributing to improving or hindering the development of Shona. The importance of such processes and strategies are discussed in the broader context of empowering African languages. The article also offers recommendations on how best to produce systematized terminology in music and other specialized fields.

Keywords: TRANSLATION, DEVELOPMENT OF AFRICAN LANGUAGES, MULTILINGUAL SOCIETIES, BORROWING, COINING, COMPOUNDING, DERIVATION, TERMINOLOGY, TERMINOGRAPHY

Opsomming: Die uitwerking van vertaalaktiwiteite op die ontwikkeling van Afrikatale in 'n veeltalige samelewing: *Duramazwi reMimhanzi* as 'n gevallestudie. Die artikel ondersoek die uitwerking van vertaalaktiwiteite op die ontwikkeling van Afrikatale in die veeltalige Zimbabwiese samelewing. Dit ontleed Sjonamusiekterme geskep deur middel van vertaalprosesse en -strategieë soos ontlening, munting, samestelling en afleiding. Die fokus is op die manier hoe voortgaande termskepping bydra tot die verbetering of belemmering van Sjona. Die belangrikeid van sulke prosesse en strategieë word bespreek in die wyer konteks van die bemagtiging van die Afrikatale. Die artikel maak ook aanbevelings hoe om op die beste manier gesistematiseerde terminologie in musiek en ander gespesialiseerde gebiede voort te bring.

Sleutelwoorde: VERTALING, ONTWIKKELING VAN AFRIKATALE, VEELTALIGE SAMELEWINGS, ONTLENING, MUNTING, SAMESTELLING, AFLEIDING, TERMINOLOGIE, TERMINOGRAFIE

Introduction

The first part of the article describes the general language situation in Zimbabwe that has resulted in translation and terminology development activities. The second part explores the nature of the ongoing term-creation in the music profession and shows its positive effects on Shona. The third section outlines the problems that result from the ongoing term-creation and the last part focuses on recommendations and probable solutions to the outlined problems.

The language situation

As many other African countries, Zimbabwe is a multilingual and multicultural society. However, its situation is less complex than that found elsewhere in the sub-Saharan region. The dominant languages are Shona, spoken by at least 75% and Ndebele, spoken by 10-16% of the country's estimated population of 11 376 676. These two Southern Bantu languages are sometimes referred to as national languages, along with English, which is used for most official purposes. Other small but significant minority language groups, which together account for 6% of the population, are Tonga, Nambya, Venda and Kalanga in the Ndebele-speaking area, and Kalanga and Shangaan in the Shona-speaking area. In addition, there are about eight other indigenous language minorities and about five African migrant minority language groups, with Chewa/Nyanja being by far the largest, plus another five or six non-African language minorities, including English and other European languages.

Whenever there are two or more languages in a society, translation and the resultant term-creation activities become inevitable. For communication between people of different languages, translation, which is a process by which the chain of signifiers that constitutes the source language (SL) text is replaced by a chain of signifiers in the target language (TL), is always a vital activity. In order to transfer ideas across different cultures the translator's tasks involve creating rather than merely finding equivalence (Venuti 1995: 38). In other words, translation involves not only finding equivalence but also creating terms when dealing with cultural and linguistic differences between languages. Terminology, therefore, is in some cases a by-product of translation activities. This is the reason why, in this article, translation and terminology are discussed simultaneously. It is this creative component which is of particular interest to this article because it contributes to borrowing and change (Hartmann and Stork 1983: 242) and therefore allows the TL to grow in both lexical capacity and utilitarian value.

Officially, not much attention, if any, has been paid to this multilingual and multicultural situation, because the development of language, translation activities and communication is generally a neglected issue. Even up to the present, Zimbabwe has no language policy, reflecting the government's lack of

seriousness on issues of language development. As a result, the standard of translation and the terms created through this process is low. The general attitude to translation and terminology in most sectors in Zimbabwe is rather casual and nobody seems to be desirous to appoint competent translators and terminologists.

Term-creation in the music profession

This article is based on observations made as a team member taking part in the ongoing project of compiling the Shona–Ndebele–English Dictionary of Music, *Duramazwi reMimhanzi*, undertaken by the African Languages Research Institute (ALRI) at the University of Zimbabwe. For this project, Shona, Ndebele and English musical terms have been collected from institutions offering music as a course or subject, and other organizations and persons involved in music. Close to 1 000 Shona musical terms have been collected and entered into the musical database.

An analysis of these terms shows that many of them have been created in various tertiary institutions teaching music. In Zimbabwe, music is a well-established discipline that uses specialized terms in the analysis and teaching of sound components. As in most specialized fields, it is taught in English and some of the existing Shona terms are equivalents created from English through translation. Chimhundu (1996: 449) aptly notes:

> A main trend in translation between international languages or languages of wider communication (LWCs) and indigenous languages or national official languages (NOLs) is unidirectional transfer from the LWCs as SLs to NOLs as TLs during the translation process. Both ideas and words are transferred as African societies modernize and change.

The same situation is found in the field of music. There is a unidirectional transfer of musical terms from English to Shona. This translation process often referred to as "borrowing" has resulted in the creation of many musical terms. Examples of such terms drawn from musicology are the following:

(1)   *aruto* (alto)
      bhesi (bass)
      bhiti (beat)
      korasi (chorus)
      sopurano (soprano)
      vhesi (verse)

The examples under (1) show that English terms are adopted and then adapted to the Shona phonological and morphological structures. However, term-creation does not end with borrowing but also includes other term-development processes such as coining. Examples of Shona musical terms that have been

created by different scholars or in different institutions of music as equivalents of existing English musical concepts and terms are the following:

(2)   *mavambo* (literally, beginning)
         musoro (literally, head)

The two terms under (2) are both used to refer to the beginning of a musical piece.

(3)   *dziro* (literally, wall (of a hut))
         muviri (literally, body)

The terms under (3) are both used to refer to the basic pattern of a musical composition.

(4)   *chimonauswa* (literally, pinnacle (of a hut))
         chisuwi (literally, pinnacle (of a hut))
         mabvuto (literally, climax)
         makukumidzo (literally, climax)

The terms under (4) are all used to refer to the climax of a musical composition.

(5)   *magumo* (literally, end)
         mhendero (literally, hem)
         mushwe (literally, tail)

The terms under (5) are all used to refer to the end of a musical composition.

The four sets of examples are good coinages which emanate from a sound understanding of the Shona environment and language. Under (2), the example *mavambo* is a well-established Shona word, denoting the beginning of something. The creator of this term selected a commonly used word that any Shona speaker/hearer can easily understand and identify with. The creation of this term might also have been influenced by the analysis of the structure of folktales. A folktale, like a musical composition, has a *mavambo* (beginning), a *muviri* (middle), and a *magumo* (end).

The examples under (3) and (4) equate the structure of a hut with that of a musical composition which has a *dziro* (basic pattern) and a *chisuwi*/*chimonauswa* (climax). However, the beginning of a song is not rendered by the terms that are currently being used. With the use of terms like *dziro* and *chisuwi*/*chimonauswa,* it would logically be expected to find a word like *hwaro* (foundation) being applied to refer to the beginning of a musical composition. This will also complete the hut imagery being applied to represent the full structure of a musical composition.

The terms *musoro*, *muviri* and *mushwe* under (2), (3) and (5) are drawn from animal imagery. A musical composition is thus equated to the structure of an

animal with a head (*musoro*), body (*muviri*) and tail (*mushwe*). The three Shona terms respectively refer to the beginning, basic pattern and end of a musical composition. The word *mushwe* is also a direct translation of the English term 'tail'.

As shown under (4), the end of a musical composition is also known as *mhendero,* a word from the Shona lexicon referring to the hem of a type of clothing. The purpose of a hem is to make the clothing appear finished and rounded off. Any musical composition requires a logical ending, showing *mhendero* an appropriate term to refer to the end of a musical piece.

The above examples demonstrate that the creation of musical terms is the result of the creators' understanding of their environment and the Shona language. Any speaker/hearer of the Shona language can easily identify with the terms for they are closely linked to real and appropriate images. This alone is a way of expanding and enriching the Shona language with new terminology. However, there are problems shown up by the given examples.

Problems

The discussed examples reveal that there are many inconsistencies in the creation and use of Shona musical terms. There are too many different terms referring to the same musical concept. The end of a musical composition for instance is represented by three different terms, namely *magumo*, *mushwe* and *mhendero* which are drawn from divergent images.

In some interviews and term lists obtained from institutions and organizations of music, there are inconsistencies that can be noted with regard to the Shona alphabet. On some lists of Shona musical terms the English term *alto* is retained despite the fact that the current Shona alphabet does not include /l/. On very few Shona term lists *alto* is represented as *aruto*. The lateral /l/ is replaced by the trill sound /r/. However, in most cases of speech /l/ instead of /r/ is used. If the term creators in the music profession coordinate their activities, they will come up with clear-cut policies on how to deal with terms like those given under (1). Decisions on how to treat the letter l and letter combinations such as pr in the English term *soprano* can easily be reached if term-creation is well planned and coordinated. Existing speech excerpts show that most Shona speakers use /pr/ and not /pur/ when speaking and to represent the letter combination pr as pur as in the given examples is to wholly misrepresent how the Shona terms are used in actual speech.

The other problem is the inaccessibility of some Shona musical terms already created. Some of the term lists that ALRI collected from the Zimbabwe College of Music contain Ndau terms coined by the late authority on ethnomusicology in Zimbabwe, Dr D.A. Maraire. The problem with some of these is that they are easily accessible to Ndau speakers and not to users of the other Shona dialects, Korekore, Manyika, Zezuru and Karanga. Terms like *mushwe* and *makukumidzo* are examples which can easily be understood by the Ndau but not

readily by users of the other Shona dialects. Therefore, the two terms should not be the standard musical terms for *end* and *climax* respectively. It would be better to use the terms *mushwe* and *chimonauswa* which are readily understood in most of the Shona dialects.

These examples indicate that there is a lack of uniformity of terms in the music profession which demonstrates that the term-creation that takes place is largely unplanned and uncoordinated. The chaotic situation in term-creation is not confined to the musical field but is a general problem affecting nearly all sectors of the Zimbabwean community. As noted by Chimhundu (1987: 142), term-creation is a growing phenomenon, particularly in the post-independence era in Zimbabwe. It is proliferating in business, central and local government, commerce, industry, mining, agriculture, broadcasting, education and other spheres of life. If this continues unchecked, it will eventually result in an undesirable terminological "flood". In other words, the result will be an influx of terms created outside the standards and principles that govern terminology development.

A situation whereby every institution has its own set of terms does not promote language development. As noted by Karl Marx, development comes through the effective control of resources, including language. In other words, language development should be well-ordered and -planned. He states that "language would be one of the things that in time would be taken under human control by the victorious proletariat" (quoted by Grinev 1987: 49). But the big question is who has to control language development in a country.

Recommendations

If the music profession is to produce terminological work that will promote the development of Shona, committees responsible for the development of musical terms should be formed. These committees should be established in the four main provinces of Zimbabwe namely Mashonaland, Matabeleland, Manicaland and the Midlands. The committees should comprise lecturers, academics and personnel from institutions and organizations dealing with musical terms. These provincial committees should collect Shona musical terms created in the areas they represent.

The chairpersons of the provincial committees should form the national board for the development of Shona musical terms. The board should be responsible for analysing the terms created so far in the music profession. To give credibility to the decisions of the national board, it should work closely with other committees, boards, institutions and organizations that have the development of indigenous languages at heart. The Shona Language Committee and ALRI are examples of groups with which such a national board for the development of Shona musical terms can cooperate in the implementation of the basic principles of terminological activities in the music profession. This recommendation, if implemented, should ensure uniformity in the creation and use of musical terms.

For the above recommendations to be effectively implemented, however, it is necessary for the Zimbabwean government to be actively involved by providing financial assistance. The above recommended committees need financial support from the government for them to effectively pursue the objective of systematically creating and storing musical terms. Terminological management requires adequate resources like buildings, research equipment and well-trained personnel. The result would be the implementation of the suggested recommendations not only for the music domain but also for other fields using specialized terms.

The government should also go a step further by forming a national unit for terminology development. Such a unit should be responsible for the ultimate documentation, standardization and publishing of term lists covering not only the music domain but also various other fields. The main objective of such a unit should be to reduce ambiguity and misunderstandings and thereby improving the exactness of technical communication. The proposed unit should deal with terminology not only of music, but also meteorology, basic health, HIV/AIDS, commerce, finance, banking, basic agriculture, information technology and other specialized fields. It should work in collaboration with linguists and specialists in these various subject fields. The unit should start with the two main indigenous languages in Zimbabwe, Shona and Ndebele, and then with time involve the recognized minority languages.

Another positive step would be to introduce tertiary training in the theory and principles of terminology. Currently, such training receives little attention. At the University of Zimbabwe in Harare, for instance, it is only taught as part of the translation and lexicography courses and as such does not receive a fair coverage. Both these courses are optional and consequently some graduates who have studied African languages have no knowledge of terminology. Furthermore, these courses are covered by only one module and the obvious result is that very little attention is given to terminological theory and principles. As a solution to this problem, it would be better for the Department of African Languages and Literature at the University of Zimbabwe and similar departments at various tertiary institutions in Zimbabwe to follow the example set by the Department of African Languages and Literature at the Midlands State University in Gweru where translation and lexicography are taught as different modules, the former in the second and the latter in the final year. The teaching of these two courses at different levels ensures that are they taught over a two-year period, hence giving more time for terminology, which is also included in the two modules. In addition, these courses are compulsory. If all tertiary institutions would follow suit, it would mean more students will receive training in terminology. However, teaching terminology as an independent subject will ensure training of terminologists and terminographers of a better quality. This is exactly what is happening in South Africa which in 2003 introduced a module on the practice and principles of terminology and terminography as part of the M.A. in the Department of African Languages at the University of Pretoria.

For quality terminological work to be realised in Zimbabwe, the same road as that taken in South Africa should be followed.

Terminological problems in the music profession and other specialized fields are not without solutions. It will only take patience and effort from the concerned stakeholders in implementing and developing the above recommendations. Effective implementation and development of these suggestions will in future enhance the communicative power of the indigenous languages of Zimbabwe.

References

Chimhundu, H. 1987. Pragmatic Translation and Language Elaboration. Pongweni, J.C.A. and J. Thondhlana (Eds.). 1990. *LASU Conference Proceedings 1987*: 140-152. Harare: LASU.

Chimhundu, H. 1996. Translation in African LLDs and Communication for Development. Malone, J.L. (Ed.). 1996. *XIV World Congress of the Fédération Internationale des Traducteurs (FIT) Proceedings, Melbourne, Australia, 9–16 February 1996*. Melbourne: The Australian Institute of Interpreters and Translators (AUSIT).

Grinev, S.V. 1987. Theoretical Foundations of Russian Terminology Work: Peculiarities and Perspectives. Grinev, S.V. (Ed.). 1987. *Applications and Implications of Current LSP Research. Volume 1*: 49-56. Bergen: Fagbokforlaget.

Hartmann, R.K.K. and F.C. Stork. 1988. *Dictionary of Language and Linguistics.* London: Applied Science Publishers.

Venuti, L. 1995. *The Translator's Invisibility.* London: Routledge.

Language Development or Language Corruption? The Case of Loan-words in *Isichazamazwi SesiNdebele**

Cornelias Ncube, *African Languages Research Institute, University of Zimbabwe, Harare, Zimbabwe (corneliasbncube@yahoo.co.uk)*

Abstract: This article discusses the loan-words in *Isichazamazwi SesiNdebele* (henceforth ISN), particularly looking at their acceptance and/or non-acceptance by the target users of ISN. In Zimbabwe, Ndebele shares the same linguistic environment with English, Shona and the official minority languages such as Kalanga, Tonga and Nambya. A historical heritage also links it with its Nguni sister languages such as Zulu and Xhosa spoken in South Africa. In selecting headwords for ISN, the Ndebele Lexicographic Unit used the frequency-list method, lemmatising words mostly found in the corpus. This method inevitably allowed the adoption of loan-words in the ISN with resultant public protest. The article is divided into two broad sections. The first section gives a general overview of comments from users of ISN about the inclusion of loan-words in the dictionary. The attitude towards loan-words in the ISN varies with different age groups, the younger generation freely accepting them as part of the Ndebele lexicon as opposed to the older generation. The second section analyses the justification by the editors of ISN for lemmatising loan-words against the views of target users. Reservations against the loan-words in ISN go beyond lexicographic principles. In the forefront is the users' attitude towards the source language. Language attitudes in Zimbabwe are mainly a result of the socio-political and economic power characterising the different tribal or ethnic groups in the country. The article concludes by discussing possible solutions to the problem of loan-words to be adopted in the forthcoming *Advanced Ndebele Dictionary*.

Keywords: LOAN-WORDS, CULTURAL BORROWING, DIALECT BORROWING, LEXICON, ADOPTION, LANGUAGE PURISM, LANGUAGE EMANCIPATION

Opsomming: Taalontwikkeling of taalbederf? Die geval van leenwoorde in *Isichazamazwi SesiNdebele*. Hierdie artikel bespreek die leenwoorde in *Isichazamazwi SesiNdebele* (voortaan ISN) deur veral te kyk na hul aanvaarding en/of nieaanvaarding deur die teikengebruikers van ISN. In Zimbabwe deel Ndebele dieselfde taalomgewing met Engels, Sjona en die offisiële minderheidstale soos Kalanga, Tonga en Nambya. 'n Historiese erfenis verbind dit ook met die Ngunisustertale soos Zoeloe en Xhosa wat in Suid-Afrika gepraat word. Vir die keuse van trefwoorde in ISN, het die Ndebele- leksikografiese eenheid die frekwensielysmetode gebruik deur die woorde te lemmatiseer wat die meeste in die korpus aangetref is. Hierdie metode het onver-

---

mydelik die opname van leenwoorde in die ISN toegelaat met die gevolglike openbare protes. Die artikel is in twee breë afdelings verdeel. Die eerste afdeling gee 'n algemene oorsig van die kommentaar van die gebruikers van ISN op die insluiting van leenwoorde in die woordeboek. Die houding teenoor leenwoorde in die ISN wissel by verskillende ouderdomsgroepe, met die jonger geslag wat hulle vrylik as deel van die Ndebeleleksikon aanvaar in teenstelling tot die ouer geslag. Die tweede afdeling ontleed die regverdiging deur die redakteurs van ISN vir die lemmatisering van leenwoorde teenoor die beskouings van die teikengebruikers. Voorbehoude teen die leenwoorde in ISN gaan verder as leksikografiese beginsels. Vooraan is die gebruikers se houding teenoor die brontaal. Taalhoudings in Zimbabwe is hoofsaaklik die gevolg van die sosiopolitieke en ekonomiese sterkte wat die verskillende stam- of etniese groepe in die land kenmerk. Die artikel sluit af met 'n bespreking van moontlike oplossings vir die probleem van leenwoorde wat in die toekomstige *Gevorderde Ndebelewoordeboek* opgeneem sal word.

Sleutelwoorde: LEENWOORDE, KULTURELE ONTLENING, DIALEKONTLENING, LEKSIKON, OORNAME, TAALPURISME, TAALEMANSIPASIE

1.   Introduction

The importance of making provision for target users' needs in the compilation of a dictionary has received considerable attention from lexicographers, among others De Schryver and Prinsloo (2000). They argue that feedback from the target users should be carried out simultaneously with the compilation of the dictionary. The process should involve publications of small-scale parallel dictionaries, stimulating feedback, which is eventually fed into the main dictionary. Reference is made here to the indispensability of the 'simultaneous feedback' process because quite often it is bypassed in the dictionary-making process. When it is eventually assessed, De Schryver and Prinsloo correctly assert that such feedback can 'only be implemented in forthcoming editions of these dictionaries'. Although the Ndebele Lexicographic Unit (henceforth NLU) bypassed the process, various lexicographic researches by students at Honours and Masters level had laid important groundwork, which was at the disposal of the editors of the first Ndebele monolingual dictionary, *Isichazamazwi Sesi-Ndebele* (henceforth ISN), to implement. The research work include those on defining formats by Ndlovu (1998), and entry ordering in Ndebele dictionaries by Maphosa (1997). The problematic area of loan-words in Ndebele received attention from Moyo (1999). Nong et al. (2002) also investigate the problem of loan-words versus indigenous words in Northern Sotho. These researchers concur that the problem with loan-words lies in language attitude which varies from generation to generation. In this article, it is argued that to comprehend the reasons for and/or against loan-words in ISN, it is important to explain both the historical tribal or ethnic relations in Zimbabwe before and after independence, and the socio-political power characterising these different tribal or ethnic groups. The need to contextualise the problem of loan-words in the socio-political milieu of Zimbabwe was realised when the NLU conducted its

seminar series to invite views on and opinions about ISN. The inclusion of loan-words expectedly generated much interest and debate.

2.   Origins and Sources of Loan-words in Ndebele

In order to understand the origins of loan-words in the Ndebele language, some information on the historical development and social position of the Ndebele community is needed. The Ndebele speech community is a mixture of the Ndebele, Kalanga, Xhosa, Sotho, Venda and Nambya tribal groups. During the precolonial period, the different ethnic groups had a hierarchical coexistence. At the zenith of the social scale was the royal family (Khumalos and those of Nguni descent). This group was called *abeZansi*, a designation that reflects its prestigious status in the earlier Ndebele community. The Xhosa and the Sotho, originally from South Africa, as well as those incorporated into the Ndebele state in the previous Transvaal area, were perceived as part of the mainstream Ndebele ethnic group. In terms of social status, the group called *abeNhla* ranked second after the royal class. Following their defeat, the Kalanga, Tonga, and Nambya groups were incorporated into the larger Ndebele group when it settled in present-day Zimbabwe. Adjacent to this amalgamated Ndebele speech community was an equally large Shona ethnic group over which the Ndebele had jurisdiction. The Shona, Kalanga, Tonga and Nambya groups were derogatively called *amaHole*. The relevance of these labels shall be referred to in detail in section 3.2

The historical heritage linking the Ndebele language to its Nguni sister language Zulu was conserved and maintained throughout the above hierarchical social arrangement and later through the education policy which allowed the introduction of Zulu literature in the school and college syllabi. Each ethnic group in the region, however, continued to speak its native language, albeit increasingly adopting the language of the conqueror, the Ndebele. Emerging as the language of the privileged in the region, Ndebele has been forced to develop in a multilingual environment composed of Kalanga, Xhosa, Sotho, Tonga, Nambya, Venda and later, on a national level, Shona and English. The above scenario inevitably resulted in both cultural and dialectal borrowing between languages. It is a generally accepted and undisputed opinion and observation that where cultural exchange and dialectal code switching are involved, lexical borrowing simultaneously takes place. In tracing the origins of loan-words in Ndebele, Zulu is certainly the primary source. In fact, Zulu continues to facilitate the entrance of loan-words from other languages into the Ndebele lexicon. A historical inquiry of the origins of many words that today are assumed, superficially, to be authentic Ndebele words, shows either English or Afrikaans as the source languages. Hadebe (2000: 229) notes:

> Ndebele has adopted more terms from English and Afrikaans than from a combined sum of all the several African languages in daily contact with the Ndebele people.

This observation shows the underlying attitude associated with the source language(s) of loan-words. This point will be discussed in detail under section 3.2. The table below shows examples of words whose source can be traced to either English or Afrikaans:

ENGLISH	AFRIKAANS	ZULU	NDEBELE
table	tafel	itafula	ithebuli / itafula
pills	pille	amaphilisi	amaphilisi
window	venster	ifasitele	iwindo / ifasiteli
store	winkel	ivinkili	isitolo / ivinkili

Most of these loan-words, which today have been naturalised as authentic Ndebele words convey ideas, concepts and/or objects previously unknown to or inexpressible in Ndebele (or Zulu itself). Because certain loan-words had not yet entered Zulu, borrowing sometimes took place directly from other languages such as English, instead of indirectly by the adoption of the rephonolised Zulu form. Thus words like *ireza* (razor) and *idotibhokisi* or *ibhini* (dustbin) now appear as part of the Ndebele lexicon, but the Zulu variants (cf. Doke et al. 1990) such as *impuco*, *ingego* or *incoshoba* and *ibhokisi lezibi (likadoti)* respectively have not been borrowed. The other variant *insingo,* though appearing in both Zulu and Ndebele, no longer belongs to the active lexicon of Ndebele. The phrase *ibhokisi lezibi* appears in everyday spoken Ndebele, but has not been entered as a lemma in ISN. This signals a shift from taking every adoption from Zulu as authentic and therefore correct. In resorting to coining as another source of words, the editors seem to have heeded the warning of Moyo (1999: 32) that selecting lexical entries simply because they are found in Zulu paints 'a rather negative picture … about Ndebele because Ndebele will be presented as being inferior to Zulu'.

It has also proved difficult for Ndebele (like all languages) to become an 'isolated language' in a multilingual society like that characterising Zimbabwe. As can be expected, from a linguist's point of view, Ndebele has enriched its vocabulary by borrowing words like *inopi* (pumpkin porridge) and *ishamari* (a person involved in an illicit love affair) from Kalanga and Shona respectively. Using the frequency-list method, editors of ISN selected and lemmatised headwords found in the Ndebele corpus. The lemma carrying the definition is that most frequently used in the corpus. But, as Kennedy (1998: 34) observes, 'it is easy to overestimate the probability of occurrence of the low frequency words of a language', because frequency depends on whether real life contexts where the words are generally used were ever mentioned during interviewing. The method nonetheless presented the possibility for the inclusion of loan-words in the dictionary. Does the inclusion of such loan-words imply that ISN adds linguistic evidence of a developing language? Or does ISN provide lexicographers (and general language practitioners) with a case-study of a language corrupted in the name of academic correctness?

3.   Language Development or Language Corruption?

3.1   Loan-words and attitude: the target users' perspective

The inclusion of loan-words in a dictionary generates much interest from the mother-tongue speakers of a language. The mother-tongue speakers of Ndebele are generally concerned with the written and relatively less with the spoken form of the language. Thus second language speakers of Ndebele are 'unconsciously' spared the criticism of mispronunciation and even distorting the syntactic features of Ndebele in spoken form. When it comes to the written form of Ndebele, both first and second language speakers are expected to correctly record the language. A person is judged to have a good command of written Ndebele, if he/she uses the words perceived as original to the language.

It is against this background that the editors of ISN were criticised for lemmatising loan-words. The argument is that if lexicographic activities at ALRI are aimed at promoting the use and at the same time preserving of the indigenous languages, then Ndebele should be recorded and included in its earlier state. Ascertaining the earlier composition of a language's authentic lexicon, however, is a daring task for a lexicographer to undertake. The biggest challenge is to circumscribe a point in history when a language like Ndebele, for instance, existed as an 'isolated language', unadulterated by loan-words. Words today perceived as original to Ndebele were once also new and probably also caused controversy, as is the case with newly adopted loan-words. This points to the difficulty of determining the number of loan-words in Ndebele. As can be expected, however, the target users of ISN complained that too much borrowing could corrupt 'their' language. The inclusion of words having the sounds /r/ and /dzw/ was thus criticised by most mother-tongue speakers of Ndebele. The general feeling was that editors should not have included loan-words in ISN, especially where Ndebele has equivalent terms. The following examples can be noted:

NDEBELE	LOAN WORD	SOURCE LANGUAGE
isijeza	inopi	Kalanga
ukuhlukuluza/ukuhlupha	ukudzwinyisa	Kalanga
umbangazwe	ipolitiki	English
ubabhemi	idonki	English
ishende	ishamari	Shona

This list represents a sample of loan-words considered unnecessary and hence should have been excluded since most of them have equivalents original to Ndebele. These favoured words, however, are not actively used in everyday conversation. This reinforces the point that those who directly experience the borrowing and are themselves the borrowers object to the standardization of these same borrowed words. The objections came despite the research team

having explained how the collection of words was carried out. This was accomplished through a rigorous process of recording everyday conversations from different Ndebele communities scattered throughout the Matebeleland and Midlands provinces, in this way creating a corpus that was used as basis for compiling the dictionary. The dictionary, it was argued, is a reflection of how the Ndebele speak the language at present. This argument was objected to on the grounds that Ndebele's status was being corrupted and that ALRI was furthering the corruption by large-scale borrowing without 'separating the chaff from the grain', as one observer complained. It is within this context that creativity is called upon in coining 'better' words in place of such terms as *ifayinifesi* (fine face) and *ikheshithokhu* (cash talk).

It should be noted, however, that the non-acceptability of the /r/ sound is not consistent in all instances where it occurs in loan-words. For example, some users of ISN are comfortable to accept the /r/ sound in words such as *irula* (ruler) and *irobhothi* (robot), but are not comfortable when the same sound appears in words such as *ishamari.* It is apparent that some words are denied a place in the Ndebele lexicon according to the attitude towards the source language. For instance, words such as *inopi* and *ishamari* were criticized because they are borrowed from Kalanga and Shona respectively. A suggestion was made that for *ishamari* the editors should have borrowed its Zulu equivalent *ishende.* Similarly, others argued that the word *ukudzwinyisa* (bullying) is the slang form of *ukuhlukuluza/ukuhlupha* which are the formal Ndebele equivalents. Objections against the inclusion of *ukudzwinyisa* are motivated by the fact that the sound /dzw/ is not found in Ndebele and the word itself is borrowed from Kalanga.

It is clear that the contention about the loan-words included in ISN is the attitude towards the source language and not the concept of borrowing itself.

3.2   Factors determining the users' attitude towards loan-words in ISN

Sebba (1997: 4) stresses:

> Almost as obstinate and damaging an idea is the notion that languages can be, and should be, 'pure'.

Language attitude generally reflects ideological/cultural contradictions amongst different ethnic groups. Mparutsa et al. (1992: 237) argue that these contradictions are 'motivated by different socio-economic interests'. In Zimbabwe, the root cause of these contradictions, which inevitably foster language attitude dates back to the historical ethnic/tribal relations before and after the country's independence.

In section 2, it has been noted that during the precolonial era, the Ndebele community was socially divisive. An association with the royal family *(abeZansi)* of the Ndebele state bestowed high and prestigious social status on an individual. As a result, the Ndebele language ended up reflecting the divisions in

the society. Herbert (1992: 5) notes that the potentiality of a language to be socially divisive 'is readily observable in the words that have been used to label population groups'. It is also a generally acceptable view that the power to name derives from social dominance which is true of the precolonial Ndebele speech community. Consequently, the language of the privileged was regarded as the language of officialdom. The *lingua-franca* status of the Ndebele language was further reinforced by the colonialists, who made it a policy that Ndebele had to be taught in place of the region's indigenous languages such as Kalanga and Nambya. In Bulilimamangwe, Kalanga chiefs were even replaced by Ndebele chiefs, in line with the dominance of the Ndebele language. The effects of these changes are today reflected in language attitude. Although Ndebele and the region's local languages continue to be in a close diglossic relationship, the stereotypical language status created before independence still continues. To date, minority groups such as the Kalanga favour Ndebele in place of their mother tongues. In such a situation, borrowing is conceived of as an anomaly when it occurs from a perceived language of 'low status' to a perceived language of 'high status'. McMahon (1994: 202) concurs:

> In cases of close contact two languages may not be perceived as equivalent in status within their speech communities … In such linguistic relationships of unequal prestige borrowings generally move from the more to the less prestigious language.

Accordingly, Ndebele mother-tongue speakers find it difficult to accept Kalanga words such as *inopi* (pumpkin porridge) and *isiki* (syphilis) in ISN, even though the corpus reveals that these words are also active in strictly Ndebele-speaking areas. The original Ndebele words, *isijeza* and *ingulamakhwa* respectively, are becoming obsolete. After the country's independence, Shona and Ndebele were declared national languages while English continued as the '*de facto* language of power and economic advancement' (Mparutsa et al. 1992: 238). The shift in policy was recommendable as it acted as the liberation's counter-ideology, which Mparutsa et al. (1992: 238) argue put 'emphasis on the pride of indigenous culture and language' in order to fight the colonial ideological hegemony. However, the political instability that emerged soon after independence between the country's major political parties of (PF) ZAPU and ZANU (PF), consciously or unconsciously, triggered a rise in the levels of contradictions existing between language usage and language attitude in present-day Zimbabwe. The two national languages, Ndebele and Shona, thus emerged out of the political conflict with their own linguistic struggle — a struggle for purism. It is against this background that it should be understood why the /r/ sound in a word such as *ishamari* is denied its place in the Ndebele lexicon while it is accepted in a word such as *irula*. A similar degree of reluctance to accept the /l/ sound even in words Shona borrowed from Ndebele has been observed among Shona speakers.

Age is also an important factor in users' attitude towards the inclusion of loan-words in ISN. Nong et al. (2002: 17) observe that 'younger respondents

seem to accept loan-words much more easily than the older generation'. Resistance to loan-words by the older generation should be understood within the context of the high social status Ndebele *(abeZansi)* had over other ethnic groups during the precolonial period. It is subsequently a result of the perceived prestigious linguistic status of Ndebele over the regional languages such as Kalanga, Tongo, Nambya and Sotho. Add to this the post-independence political conflict, which apparently took tribal dimensions, and it becomes clear why the older generation who lived through the experience rejects any influence from Shona. However, the conservatives propagating a pure Ndebele language, suppress the political causality by arguing for 'language emancipation' from Shona and English. Tribal causality is also suppressed by arguing for 'language purism' free from corruption by minority languages such as Kalanga, Tonga, Nambya and Sotho. Instead, 'tradition' is used to justify the demand for the 'pure' state of Ndebele, the merits of which arguably lie in its preciseness, musicality, logicality, lyricism etc. Sebba (1997: 4), however, notes that purist attitudes 'are cultural phenomena which can, and do, change with time and which are not invariant from generation to generation or society to society'. The younger generation, who only has a remote recollection of the high status of the Ndebele language and a hazy politically informed comprehension of the linguistic struggle in Zimbabwe, has a moderated attitude towards loan-words. Their attitude is conditionally determined by the problems loan-words pose in the education curriculum. Otherwise, the younger generation, in particular those of school-going age, is favourably inclined towards the inclusion of loan-words in ISN. The dilemma of the younger generation is whether to proceed and use the loan-words notwithstanding the strong resistance by teachers who, for instance, prefer the original words *umkambo* (flea market) and *esilindweni* (bus stop) instead of the commonly used *ifilimakhethi* and *ibhasitopu* respectively. A student/pupil who uses the controversial loan-words risk being evaluated as and penalized for lacking a good command of Ndebele.

The problem reveals the insufficient liasing between Zimbabwe's main examination regulating board, ZIMSEC (Zimbabwe Schools Examination Council) and the lexicographic activities at ALRI (including other language research departments at the University of Zimbabwe). New developments in the structure of a language are first and foremost noticed by lexicographers and linguists in general. With a view to this, it would be expected of the Ministry of Education and ZIMSEC to work closely with academics at universities when designing school syllabi and setting examinations. The young generation agrees that ISN has been useful as far as Ndebele grammar is concerned. Attention should now be directed towards ensuring that the language as reflected in ISN becomes acceptable in examinations. At this point, a close working relationship between ALRI and the Ministry of Education, Sports and Culture, under whose auspices ZIMSEC falls should be fostered as a matter of urgency. One of the functions of a dictionary is to popularize newly introduced words. It is apparent, however, that in as much as the editors of ISN want to popularize loan-words, this might not succeed unless the words are acceptable in exami-

nations in both teachers' colleges and schools. There is a big chance that loan-words such as *ishamari* which are actively used today may not stand the test of time. The reason for this is that some words are introduced into the lexicon of a language for the sake of style. The lexicographer therefore has the difficult task of distinguishing between loan-words used colloquially and those used formally in written or spoken form. Failure to distinguish between colloquial and formal words seems therefore to be one of the factors influencing the negative attitude towards loan-words. This shows that popularization through lexicography alone is not enough, especially where a dictionary culture amongst the general users is, as observed by Maphosa (1997), very low.

3.3   Loan-words and corpus-based lexicography: the editors' perspective

In a multilingual society such as the one characteristic of Zimbabwe where the language attitude is very much pronounced, it is imperative for the lexicographer to authenticate or justify his/her dictionary entries so as to avert criticism for lemmatizing controversial loan-words. Authentication or justification in ISN is found as part of the front matter and in some instances sentences retrieved from the corpus for purposes of clarifying definitions. In the front matter, the editors explicitly state that the corpus was the basis for the selection of headwords for lemmatization. From a lexicographic point of view, synchronic attestation by the speech community is accountable for the increase in loan-words in Ndebele. This sheds light on the merits of corpus-based lexicography. Landau (1984: 280) accentuates the advantage of using a corpus for lexicography:

> One enduring value of such a computer-generated file for lexicography is, theoretically, to provide a solid basis for reexamining the definitions of the standard lexicon. It can also be a treasure of information on word frequency, punctuation, preferred written form (spelling, capitalization, and compounding), word order, verb patterns, and other grammatical data.

A corpus, therefore, is a useful tool to which generalisations and hypotheses on the language can be referred for verification. From the editors' point of view, the inclusion of loan-words in ISN is justified. Instances where perceived 'original' Ndebele words like, for example, *umkambo* and *umlola* (vs *isepa* (English *soap*)) appear in the corpus, they have a probability usage of 15%. The usage of the original words was mostly taken from older literature or from interviews involving older people. Loan-words, on the other hand, are found in interviews involving both the older and younger generations. The probability usage of loan-words represents 85% of the total Ndebele lexicon. The increasing use of loan-words reflects that Ndebele 'is adapting to the dynamic world's technological and scientific innovations' (Moyo 1999: 42). The inclusion of loan-words from Kalanga and Shona such as *inopi* and *ishamari* is not aimed at corrupting the purity of Ndebele. Instead, from a lexicographic viewpoint, ISN is synony-

mous with culture speaking to its people. What the dictionary is saying to the people is that their cultural values have undergone a cycle of transformation. The dictionary captures signs of slight changes in attitudes towards other languages. These changes are today reflected in the loan-words in Ndebele.

4.   Possible Solutions

Moyo (1999: 42) argues that the problem with loan-words in Ndebele can be solved by a change in 'the social attitudes the Ndebele people have with regards to language'. Moyo further argues that the Ndebele's acceptance of the universal principle that language changes will make the inclusion of loan-words 'an acceptable move characteristic of all languages'. Although the importance of a change in language attitude may be seen as one of the possible solutions, additional ways of authenticating the Ndebele lexicon can be proposed.

The present authentication as it appears in the front matter is inadequate for laypersons or general users who rarely consult it. In its present form, the ISN assumes selfauthentication of loan-words. It is for this reason that additional information about the lemma is proposed for inclusion in the *Advanced Ndebele Dictionary* (AND). Consequently, editors of AND consider the inclusion of the etymology of controversial loan-words as a solution. The etymological field, it must be noted, might perhaps be helpful to a language specialist such as a lexicographer or linguist, but hardly to the layperson or general user. A language specialist has both a practical and a theoretical interest in the language. A layperson or general user, on the contrary, is solely practically concerned with definitions and the conservative form of the language. To a language specialist therefore, AND can become a source for both the diachronic and synchronic study of Ndebele and for comparative studies regarding the stability of varieties and the general nature and direction of change in Ndebele. For the layperson or general user to know the source language from which the loan-lemma had been borrowed could be sufficient.

The inclusion of etymology therefore will not necessarily entail a change of attitude towards the source language from which the borrowing took place. Because a negative attitude towards the source language can mirror the same attitude towards the speakers of that language, it will probably take a political rather than a linguistic solution to resolve this. The acceptance of loan-words and the attitude towards the source language of the loan-words will always overlap in the view of the layperson. For purposes of the acceptance of loan-words, it might help to bridge the gap between the lexicographic activities at ALRI and ZIMSEC. There is an urgent need for the Ministry of Education, Sport and Culture to formulate a policy that recognizes the loan-words in ISN as acceptable in examinations. At present, the incorporation of loan-words into the Ndebele lexicon is prevented by the examiner by penalizing students using them. It might also be helpful to introduce a lexicography course in teachers' colleges since classroom practitioners are the very people who teach the lan-

guage. This will go a long way in helping them appreciate the science of language change and language development.

5.   Conclusion

This article has shown that Ndebele develops in a multilingual environment. The increasing adoption of loan-words can be accounted for by taking into consideration the diglossic environment in which Ndebele develops. Despite the increasing use of loan-words by the speakers in everyday speech, their inclusion in ISN has been viewed as a process of 'language corruption'. There are, however, some inconsistencies, with some loan-words with the same problematic sounds apparently being accepted but others being denied a place in the Ndebele lexicon. In instances where loan-words are resisted, this article has shown that the aspect of language attitude is at play and not the principle of borrowing itself. The deep-rooted problem of language attitude among the speakers of Ndebele can be explained by taking into consideration the 'divisive' nature of the Ndebele community in precolonial Zimbabwe. It also emanates from the political conflict, which took a tribal bias soon after independence. The younger generation, however, appreciates the inclusion of loan-words, with the only problem lying in the gap between lexicography and the education curricula in Zimbabwe. Bridging this gap is seen as a positive step towards ensuring the acceptability of these loan-words in examinations. From a lexicographic point of view, loan-words vindicate the widely accepted opinion that there is no static language. Every language grows, Ndebele being no exception to the rule. As a step forward, the inclusion of etymology is proposed for the *Advanced Ndebele Dictionary*, although it might benefit the language specialist more than the layperson or general user.

References

De Schryver, G.-M. and D.J. Prinsloo. 2000. The Concept of 'Simultaneous Feedback': Towards a New Methodology for Compiling Dictionaries. *Lexikos* 10: 1-31.

Doke, C.M. et al. 1990. *English–Zulu/Zulu–English Dictionary*. Johannesburg: Witwatersrand University Press.

Hadebe, S. 2000. Developing Terminology in African Languages with Special Reference to Ndebele. Chiwome, E.M. et al. (Eds.). 2000. *Indigenous Knowledge and Technology in African and Diasporan Communities: Multi-disciplinary Approaches*: 225-231. Harare: Mond Publishers.

Hadebe, S. et al. (Eds.). 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.

Herbert, R.K. (Ed). 1992. *Language and Society in Africa: The Theory and Practice of Sociolinguistics*. Johannesburg: Witwatersrand University Press.

Kennedy, G. 1998. *An Introduction to Corpus Linguistics.* London: Longman.

Landau, S.I. 1984. *Dictionaries: The Art and Craft of Lexicography*. London: Cambridge University Press.

Maphosa, M. 1997. *The Morphological Structure of the Noun in Ndebele and its Implications on the Ordering of Entries in Ndebele Dictionaries*. Unpublished B.A. Hons. Thesis. Harare: University of Zimbabwe.

McMahon, M.S. 1994. *Understanding Language Change.* Cambridge: Cambridge University Press.

Moyo, P. 1999. *Headword Selection for the General Ndebele Dictionary: Problems and Probable Solutions.* Unpublished B.A. Hons. Thesis. Harare: University of Zimbabwe.

Mparutsa, C. et al. 1992. An Initial Investigation into Language Attitudes of Secondary School Students in Zimbabwe. Herbert, R.K. (Ed.). 1992. *Language and Society in Africa: The Theory and Practice of Sociolinguistics*: 235-245. Johannesburg: Witwatersrand University Press.

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.

Nong, S. et al. 2002. Loan-words versus Indigenous Words in Northern Sotho: A Lexicographic Perspective. *Lexikos* 12 : 1-12.

Prinsloo, D.J. et al. 2000. *The Dictionary-Making Process with 'Simultaneous Feedback' from the Target Users to the Compilers.* Paper presented at the Fifth International Conference of the African Association for Lexicography at the University of Stellenbosch, Stellenbosch, 3–5 July 2000.

Sebba, M. 1997. *Contact Languages: Pidgins and Creoles.* London: Macmillan Press.

From *TshwaneLex* to *TshwanePedia*: Creating and Flexibly Maintaining Online Encyclopaedias[*]

Gilles-Maurice de Schryver, *Department of African Languages and Cultures, Ghent University, Ghent, Belgium and TshwaneDJe HLT, Pretoria, Republic of South Africa (gillesmaurice.deschryver@UGent.be)*, and David Joffe, *TshwaneDJe HLT, Pretoria, Republic of South Africa (david.joffe@tshwanedje.com)*

Abstract: The addition of a restricted number of features to the dictionary (compilation) software *TshwaneLex* suffices to turn this application into a tool for the creation and maintenance of encyclopaedias. This article gives a brief overview of those extra features, using the online encyclopaedia of the *James Randi Educational Foundation* (JREF) as case study.

Keywords: LEXICOGRAPHY, DICTIONARY, ENCYCLOPAEDIA, SOFTWARE, ONLINE, TSHWANELEX, TSHWANEPEDIA, JAMES RANDI EDUCATIONAL FOUNDATION

Samenvatting: Van *TshwaneLex* naar *TshwanePedia*: het samenstellen en flexibel herzien van on line encyclopedieën. De toevoeging van een beperkt aantal uitbreidingen aan de woordenboek(aanmaak)software *TshwaneLex* is voldoende om dit programma met succes voor de samenstelling en herziening van encyclopedieën in te zetten. Dit artikel geeft een kort overzicht van die extra uitbreidingen, en gebruikt de on line encyclopedie van de *James Randi Educational Foundation* (JREF) als illustratie.

Sleutelwoorden: LEXICOGRAFIE, WOORDENBOEK, ENCYCLOPEDIE, SOFTWARE, ON LINE, TSHWANELEX, TSHWANEPEDIA, JAMES RANDI EDUCATIONAL FOUNDATION

1. The off-the-shelf dictionary (compilation) software TshwaneLex

In South Africa, the dictionary (compilation) software *TshwaneLex* is well-known. Development of the application started in Pretoria in mid-2002, and already one year later a first release was in use at the Sesotho sa Leboa *National Lexicography Unit* (NLU). Since then, all members of the eleven NLUs have come into contact with TshwaneLex, either through training sessions organised by the *Pan South African Language Board* (PanSALB) and/or simply as a result of

---

the fact that they use TshwaneLex on a daily basis in their respective units. Several commercial dictionary publishers in South Africa, including *Oxford University Press* and *Pharos Dictionaries*, also use or are in the process of acquiring TshwaneLex. Reports of the first South African products compiled and placed online with TshwaneLex may be found in *Lexikos* 13 (De Schryver 2003: 10-12) and *Lexikos* 14 (De Schryver et al. 2004: 56-57, 66).

Over the years, TshwaneLex has also been well-received at all major international lexicography conferences, including TAMA 2003 (Johannesburg), AFRILEX 2003 (Windhoek), DWS 2003 (Brighton), EURALEX 2004 (Lorient), AFRILEX 2004 (Libreville), DWS 2004 (Brno), ASIALEX 2005 (Singapore), PALMA 2005 (Kuala Lumpur), LEXICOM 2005 (Brno), COMPLEX 2005 (Budapest), and AFRILEX 2005 (Bloemfontein). At each of those meetings, the then-latest features of TshwaneLex were introduced, features about which one can read more in the proceedings of each of those conferences.

Today, there are TshwaneLex users in the four corners of the world: from Papua New-Guinea and China in the East, to the United States in the West, from Estonia and Ireland in the North, to South Africa in the South. The dictionary projects are either government-sponsored (e.g. at the *Royal National Academy of Medicine* in Spain, or at the *Research Centre of African Languages and Literatures* in Congo), commercial (e.g. at *Van Dale Lexicografie* in the Netherlands, or at *Macmillan* in Botswana), or private (with users in Japan, Macao, Afghanistan, Albania, Slovenia, the Czech Republic, Germany, Luxembourg, France, the United Kingdom, Kenya, etc.).

Clearly, in order to cover such a wide variety of projects and languages, each with its own unique dictionary structure and needing its own script(s), TshwaneLex had to be a truly off-the-shelf application. To attain this, the software was built around three core concepts: user-friendliness, language-independency, and full customisability. User-friendliness is achieved by means of close cooperation between the developers of the software and numerous beta testers around the world. The language-independent nature of the application is realised thanks to full Unicode support on all levels, which also allows for the simultaneous use of various left-to-right and right-to-left scripts. Customisability is brought about by, among others, a powerful Document Type Definition (DTD) editor and linked styles system. This third aspect, customisability, turned out to be so powerful that it led to two adaptations of the basic TshwaneLex code: *TshwanePedia* for the production of encyclopaedias, and *TshwaneTerm* for the management of terminology. In this article we will be concerned with the former, and in a subsequent one (cf. Joffe and De Schryver 2005a) we will look into the latter.

2.   From TshwaneLex to TshwanePedia

One of the most important aspects we felt had to be in TshwaneLex was a high degree of built-in customisability, as each dictionary project has its own struc-

ture and styles, or "style guide". To this end, we built functionality into TshwaneLex to allow end-users to customise the DTD. The DTD defines the structure of articles in the dictionary, and the fields that appear in a specific dictionary. Tied in with the DTD is the styles system, which allows one to customise the entire formatting for all fields (e.g. bold/italics, Times New Roman/Arial, as well as common punctuation to appear before, after or between fields). An in-depth (technical) discussion of the multilayered TshwaneLex DTD editor may be found in Joffe and De Schryver (2005).

What is important here is that this customisability allows for the creation of other types of reference works with TshwaneLex, not just 'dictionaries'. For example, several TshwaneLex users have (ab)used the software for the creation of bibliographies, address databases, and even diaries.

When the *James Randi Educational Foundation* (JREF) approached us to place their 'Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural' (Randi 1995) online, we realised that TshwaneLex was indeed flexible enough to handle such a project. At the same time, however, we seized the opportunity to add a string of additional features to turn Tshwane*Lex* into Tshwane*Pedia*. The extra features, although predominantly useful for the compilation of encyclopaedias, have been 'fed back' into TshwaneLex, thus becoming available for dictionary compilation as well. Three issues will be focused upon, viz. 'window layout', 'multimedia' and 'export' features.

The first difference one notices when comparing a typical dictionary with a typical encyclopaedia, is that encyclopaedia entries are generally much longer than dictionary articles. Additionally, whereas the data of a single dictionary article is normally broken up into many different chunks, with each chunk being placed in a separate and carefully thought-out field in the DTD, encyclopaedia entries are more straightforward. Although it remains important for the compilers of an encyclopaedia to be able to see the 'structure' of the entries they are compiling (in the Tree View), more (horizontal) space is thus often needed for the various input boxes. For that reason, a so-called optional 'Wide Tools window layout' was implemented, which is accessible with a single 'hotkey'. Addendum 1 shows a screenshot of the TshwaneLex interface with the wide view enabled. (Note that when the wide view is not enabled, the entire right side is taken up by a preview of the encyclopaedia entries.)

Secondly, encyclopaedias also typically contain far more illustrations throughout. A new (multimedia) data type 'Image file' was added to that intent. In the DTD for the encyclopaedia shown in Addendum 1, images (and their captions) may be added following any paragraph. All images are stored in a central place, and whenever compilers want to add a new image, they can simply use the 'Browse …' button to select a stored image. This has been taken one step further. Given that the corpus of the future is the Web, TshwaneLex already had a hotkey to launch a *Google Web* search for the lemma sign one is working on, the idea being that one can simply select/adapt corpus lines from the Web. This functionality has been extended to the images, with another hotkey now also launching a *Google Images* search.

The third extra feature concerns flexibility of the export, especially with online encyclopaedias in mind. TshwaneLex already provided several methods for placing reference works online. The online dictionaries described in *Lexikos* 13 (De Schryver 2003: 10-12) and *Lexikos* 14 (De Schryver et al. 2004: 56-57, 66), for example, were placed on the Web with the 'TshwaneLex online software module'. This is a customisable set of PHP scripts that provide functions for creating a search interface where the user can enter words, to perform searches on a TshwaneLex file stored in a MySQL database, and to generate HTML output. In order to decrease the load on the web server, one may rather wish to generate 'static' output, where the reference work is placed online as a pre-generated file or set of files. In this regard the 'Export HTML' features were extended, with options to create one file per alphabetical category or even one file per encyclopaedia entry, in addition to one single large file. In the screenshot shown below, for instance, the output will be generated as one file per entry, with the data for each of those entries being 'dropped' into a template file.
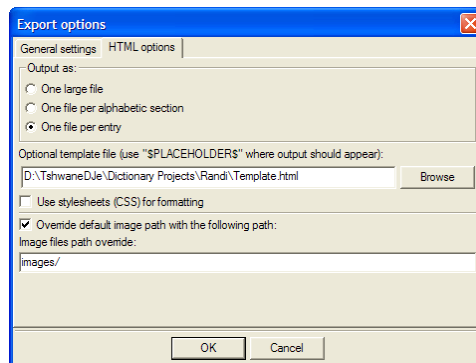
3.   Placing the JREF encyclopaedia online

The first edition of James Randi's encyclopaedia was published as a hardcopy in 1995, by St. Martin's Press in New York. In 1996 the not-for-profit *James Randi Educational Foundation* (JREF) was founded "to promote critical thinking by reaching out to the public and media with reliable information about paranormal and supernatural ideas so widespread in our society today" (JREF 1999–2005). When the JREF website was launched in 1999, the idea rose to link the *entire* contents of James Randi's encyclopaedia to the site. Various attempts produced mixed results over the years, and in mid-2005 *TshwaneDJe HLT* (the company which created TshwaneLex) agreed to undertake the task.

The encyclopaedia was received as a set of WordPerfect files, and these were parsed and then imported into TshwaneLex. The material was proofread, corrected and extended, and a TshwaneLex plug-in was created to transform the (implicit) cross-references into hyperlinks. In the process, the three adaptations mentioned above were made to the software. The encyclopaedia pages,

one per alphabetical category, were uploaded to the JREF site on July 28, 2005, and James Randi 'announced' this one day later in his weekly column.

In just four days' time, from August 1 to 4, no less than *sixty* 'bloggers' referred to and commented on the encyclopaedia — an overwhelming response. Today's bloggers clearly complement the feedback strands used so far in our research on (online) dictionary use: "a well thought out log file has been unobtrusively keeping track of all aspects of dictionary use, while an online feedback form has allowed for a more traditional and open way of receiving feedback" (De Schryver and Joffe 2004: 188). One blogger pointed out that large HTML files are cumbersome (= 'implicit feedback'); in the update that went live on August 9, 2005, seven hundred pages were automatically exported and uploaded, one per entry, instead of one per alphabetical category (= 'reaction'). See Addendum 2 for an example of the online encyclopaedia in this regard.

As one can see, *creating* and subsequently flexibly *maintaining* an online encyclopaedia has now become available at every compiler's fingertips thanks to TshwanePedia (just as was already the case for online dictionaries thanks to TshwaneLex).

References

De Schryver, Gilles-Maurice. 2003. Online Dictionaries on the Internet: An Overview for the African Languages. *Lexikos* 13: 1-20.

De Schryver, Gilles-Maurice and David Joffe. 2004. On How Electronic Dictionaries are Really Used. Williams, G. and S. Vessier (Eds.). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6–10, 2004*: 187-196. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud.

De Schryver, Gilles-Maurice, Elsabé Taljard, M.P. Mogodi and Salmina Maepa. 2004. The Lexicographic Treatment of the Demonstrative Copulative in Sesotho sa Leboa — An Exercise in Multiple Cross-referencing. *Lexikos* 14: 35-66.

Google. 1998–2005. Google Search Engine [online]. Available: <http://www.google.com/>.

Joffe, David and Gilles-Maurice de Schryver. 2005. Representing and describing words flexibly with the dictionary application TshwaneLex. Ooi, V.B.Y., A. Pakir, I. Talib, L. Tan, P.K.W. Tan and Y.Y. Tan (Eds.). 2005. *Words in Asian Cultural Contexts, Proceedings of the 4th Asialex Conference, 1–3 June 2005, M Hotel, Singapore*: 108-114. Singapore: Department of English Language and Literature & Asia Research Institute, National University of Singapore.

Joffe, David and Gilles-Maurice de Schryver. 2005a. From *TshwaneLex* to *TshwaneTerm*: Tailoring Terminology Management for South Africa. *Lexikos* 15: 312-315.

JREF. 1999–2005. James Randi Educational Foundation [online]. Available: <http://randi.org/>.

Randi, James. 1995. *An Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural: James Randi's Decidedly Skeptical Definitions of Alternate Realities*. New York: St. Martin's Press.

TshwaneDJe HLT. 2003–2005. TshwaneDJe Human Language Technology [online]. Available: <http://tshwanedje.com/>.

Addendum 1:   Screenshot of TshwaneLex, in 'Wide Tools window layout'
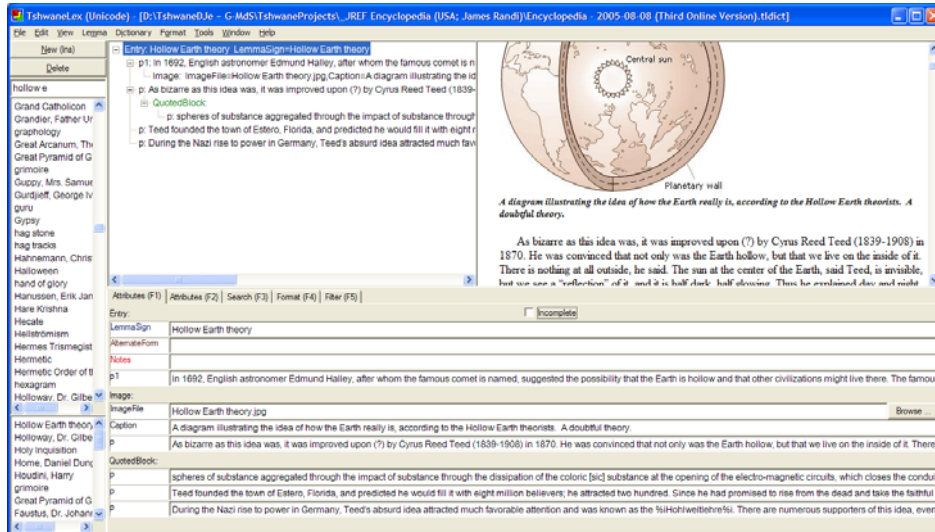
Addendum 2:   Screenshot of one entry of the JREF's online encyclopaedia
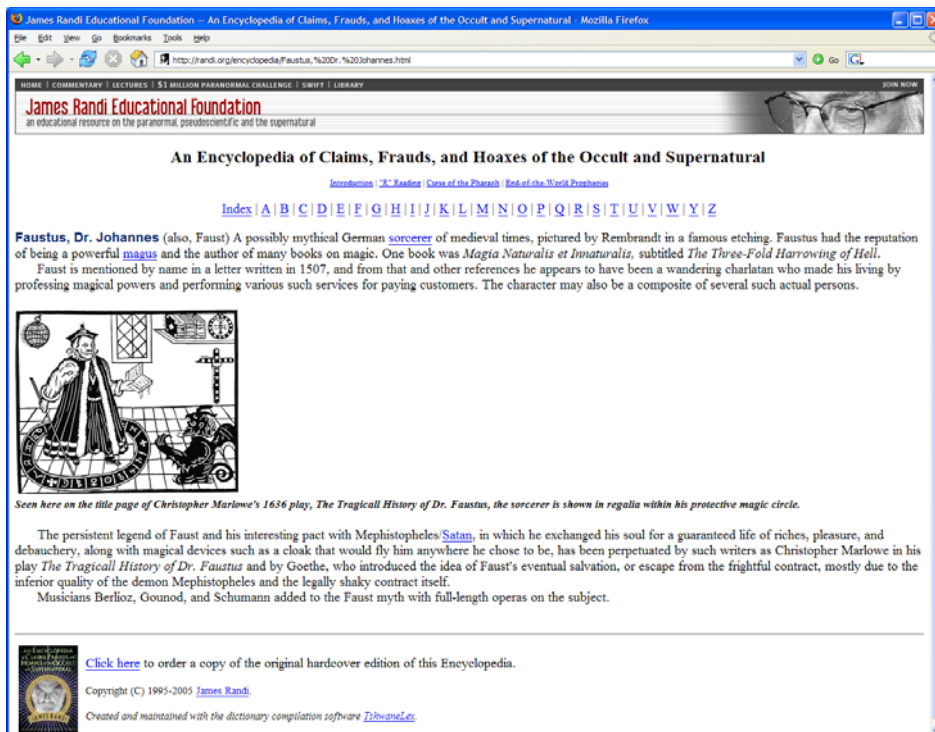
From *TshwaneLex* to *TshwaneTerm*: Tailoring Terminology Management for South Africa[*]

David Joffe, *TshwaneDJe HLT, Pretoria, Republic of South Africa*
(david.joffe@tshwanedje.com), and
Gilles-Maurice de Schryver, *Department of African Languages and
Cultures, Ghent University, Ghent, Belgium and TshwaneDJe HLT, Pretoria,
Republic of South Africa (gillesmaurice.deschryver@UGent.be)*

Abstract: The addition of a restricted number of features to the dictionary (compilation) software *TshwaneLex* suffices to turn this application into a terminology management system. This article gives a brief overview of those extra features, using the *Department of Arts and Culture* (DAC) AIDS list as case study.

Keywords: LEXICOGRAPHY, TERMINOLOGY LIST, TERMINOLOGY MANAGEMENT SYSTEM, SOFTWARE, TSHWANELEX, TSHWANETERM, DAC AIDS LIST

Samenvatting: Van *TshwaneLex* naar *TshwaneTerm*: terminologiebeheer op maat voor Zuid-Afrika. De toevoeging van een beperkt aantal uitbreidingen aan de woordenboek(aanmaak)software *TshwaneLex* is voldoende om dit programma met succes als terminologiebeheersysteem in te zetten. Dit artikel geeft een kort overzicht van die extra uitbreidingen, en gebruikt de *Department of Arts and Culture* (DAC) AIDS lijst als illustratie.

Sleutelwoorden: LEXICOGRAFIE, TERMINOLOGIELIJST, TERMINOLOGIEBEHEER-SYSTEEM, SOFTWARE, TSHWANELEX, TSHWANETERM, DAC AIDS LIJST

1. The need for a South African terminology management system

This article is a sequel to De Schryver and Joffe's (2005a) 'From *TshwaneLex* to *TshwanePedia*', in which it is pointed out that the high degree of customisability found in the dictionary (compilation) software TshwaneLex may be exploited for the production of a wide variety of reference works. The variety under discussion here are terminology lists, and by extension the move from an environment to compile dictionaries with, Tshwane*Lex*, to a terminology management system, Tshwane*Term*.

---

[*] An earlier version of this article was presented at the Tenth International Conference of the African Association for Lexicography, organised by the Sesiu sa Sesotho Lexicography Unit, University of the Free State, Bloemfontein, Republic of South Africa, 13–15 July 2005.

When Mariëtta Alberts, the 'Focus Area Manager: Standardisation and Terminology Development' of the *Pan South African Language Board* (PanSALB), pointed out that the South African terminographers who work at the *Department of Arts and Culture* (DAC) and elsewhere in the country do not have adequate tools at their disposal, *TshwaneDJe HLT* decided to produce a solution for them. In analogy with the design of TshwanePedia, and in order to create a successful application, the extra features that were added to the TshwaneLex base revolve around the 'window layout', 'multimedia' and 'export'. For the illustrations below a draft AIDS terminology list compiled by DAC, which was parsed and imported into TshwaneTerm, is used.

2.   From TshwaneLex to TshwaneTerm

To begin with one can point out that terminology databases typically contain a rather large number of languages in parallel. Creating a Document Type Definition (DTD) in TshwaneLex that can handle multilingual or interlingual data is straightforward, as the customisability already allows for the addition of fields to entries for more languages, and in fact *any number* of languages. In South Africa this usually means that one will want to work with an eleven-lingual database, and that each terminographic entry will contain data in all eleven official languages. Compared to dictionary articles, terminology entries are less complex, in the sense that, at least in theory, only one equivalent is supposed to be provided per language. To terminographers, seeing the 'structure' of the entries (in the Tree View) is thus a matter of lesser importance, as that structure is repetitive and basically the same throughout. Instead, more (vertical) space is needed to see more languages in parallel. For that reason, a so-called optional 'High Tools window layout' was implemented, which is accessible with a single 'hotkey'. Addendum 1 shows a screenshot of the TshwaneTerm interface with the high view enabled. (Note that when the high view is not enabled, the top section is taken up by the Tree View. If one for example wants to add a second translation equivalent for a particular language, as is the case for siSwati in the screenshot, then this can simply be done by adding the extra field(s) in the Tree View first.)

Secondly, a multimedia feature available in TshwaneTerm is enhanced support for audio files, through a data type 'Sound file', that can be added at *any level* in the DTD. Sound files are thus not restricted to the head-term (and, say, its morphological derivations) but can truly appear anywhere in the DTD, and they can be of any length. For the terminology list shown in Addendum 1, an audio recording was added of the pronunciation of the equivalents in each of the eleven languages. Just as for the images, sound files are stored centrally, and the compiler simply uses 'Browse …' to select one.

The third and most important feature, however, is the ability to be able to generate multiple 'editions' from a single database (cf. De Schryver and Joffe 2005: 56-59). This is done by creating different sets of styles, each of which basi-

cally 'lets through' different parts of the data. In the screenshot shown below, for example, four style sets have been defined: a default style (which corresponds with the view in Addendum 1), and then three 'masks' to output only the isiZulu, or only the Tshivenda, or only the Sotho-group languages.
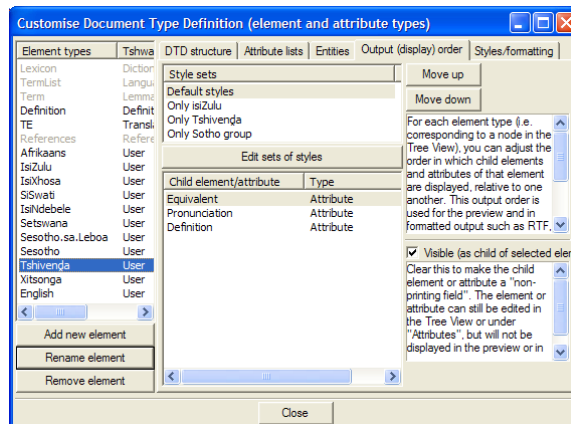
In Addendum 2 screenshots of samples of the 'isiZulu mask' and the 'Sotho mask' are shown. One can easily toggle between the different editions, and thus views, with a single key press.

The use of sets of styles within a single database is very powerful indeed. Not only can subsets of the data be output in various formatted ways, as illustrated in Addendum 2, but one can also successfully prepare, say, both a pocket edition and a full edition of the same data simultaneously. Or one can generate different views of the same data, such as a print view and an electronic/online view, or one can generate monolingual, bilingual and semi-bilingual reference works from the same database, etc. All of these possibilities are especially useful when one deals with large amounts of data, as in multilingual terminography — South Africa's richness situation par excellence.

References

DAC. 2003–2005. Department of Arts and Culture [online]. Available: <http://www.dac.gov.za/>.

De Schryver, Gilles-Maurice and David Joffe. 2005. One database, many dictionaries — varying co(n)text with the dictionary application TshwaneLex. Ooi, V.B.Y., A. Pakir, I. Talib, L. Tan, P.K.W. Tan and Y.Y. Tan (Eds.). 2005. *Words in Asian Cultural Contexts, Proceedings of the Fourth Asialex Conference, 1–3 June 2005, M Hotel, Singapore*: 54-59. Singapore: Department of English Language and Literature & Asia Research Institute, National University of Singapore.

De Schryver, Gilles-Maurice and David Joffe. 2005a. From *TshwaneLex* to *TshwanePedia*: Creating and Flexibly Maintaining Online Encyclopaedias. *Lexikos* 15: 306-311.

TshwaneDJe HLT. 2003–2005. TshwaneDJe Human Language Technology [online]. Available: <http://tshwanedje.com/>.

Addendum 1:   Screenshot of TshwaneTerm, in 'High Tools window layout'

Addendum 2:   Generating multiple 'editions' (views) from a single eleven-lingual terminology database with TshwaneTerm. Left: extracting isiZulu (and English). Right: extracting the three Sotho equivalents (*with a nod to D.J. Prinsloo*).

The African Association for Lexicography: After Ten Years

Mariëtta Alberts, *PanSALB, Pretoria, Republic of South Africa*
(marietta@pansalb.org.za)

Abstract: This year it has been ten years since the establishment of AFRILEX. After an explanation of what led to its forming, this article gives a short overview of its composition and activities.

Keywords: AFRICAN ASSOCIATION FOR LEXICOGRAPHY (AFRILEX), AFRICAN LANGUAGES ASSOCIATION OF SOUTH AFRICA (ALASA), BUREAU OF THE WOORDEBOEK VAN DIE AFRIKAANSE TAAL (WAT), DICTIONARY OF SOUTH AFRICAN ENGLISH (DSAE), LINGUISTIC ASSOCIATION OF SOUTHERN AFRICA (LSSA), NATIONAL LEXICOGRAPHY UNITS (NLUs), PAN SOUTH AFRICAN LANGUAGE BOARD (PanSALB)

Opsomming: Die African Association for Lexicography: Ná tien jaar. Vanjaar is dit tien jaar sedert die totstandkoming van AFRILEX. Na 'n verduideliking wat tot sy stigting gelei het, gee die artikel 'n kort oorsig van sy samestelling en werksaamhede.

Sleutelwoorde: AFRICAN ASSOCIATION FOR LEXICOGRAPHY (AFRILEX), AFRICAN LANGUAGES ASSOCIATION OF SOUTH AFRICA (ALASA), BURO VAN DIE WOORDEBOEK VAN DIE AFRIKAANSE TAAL (WAT), DICTIONARY OF SOUTH AFRICAN ENGLISH (DSAE), LINGUISTIC ASSOCIATION OF SOUTHERN AFRICA (LSSA), NASIONALE LEKSIKOGRAFIESE EENHEDE (NLE's), PAN SUID-AFRIKAANSE TAALRAAD (PanSAT)

1. Background

1.1 Feasibility study

In 1992, a feasibility study was done to determine the need for a Southern African Institute for Lexicography. The idea for such an Institute originated at the Board of Control of the Woordeboek van die Afrikaanse Taal (WAT). The then Editor-in-Chief of the Bureau of the WAT, Dr D.J. van Schalkwyk, requested Dr M. Alberts to do the research on behalf of the Board of Control of the WAT. Prof. W. Branford was the other member of the research team. The Gencor Development Trust provided funding for the research.

Although the idea originated at the Board of Control of the WAT, the members of the Board realised that an Institute could not be established without the collaboration and consent of all stakeholders in the Southern African lexicography fraternity. It was therefore decided to do a feasibility study in this regard (see Alberts 1993).

1.2   Problems related to lexicography practice

Dictionaries play a vital role in communication and communication is vital for the well-being of all citizens of a country. Dictionary users may experience communication needs due to a lack of information in dictionaries and other lexicographical information sources (e.g. word and term banks). This could be related to insufficient dictionary products and/or lexicographical practice.

The South African dictionary practice was fragmented. Different private or governmental components were established as a result of historical, cultural or organisational reasons, had the right to exist, and continued to exist due to these reasons. An Institute as proposed could serve to unite all these components.

A big problem related to insufficient dictionary products was (and still is) financing. At that stage (1992), only the official Afrikaans and English dictionary offices (the Bureau of the WAT and the Dictionary for South African English (DSAE)) received funding from the government. This was due to the bilingual policy of the country at that stage. The African languages did not receive official funding. Some dictionaries were compiled in the African languages, but these products were either university-based or private initiatives. Proper funding was needed to produce more dictionaries for better communication.

Another problem was the lack of trained lexicographers and terminologists in the country. It was thought that a Southern African Institute for Lexicography could utilise available expertise in this regard and provide in-house training.

1.3   Research areas

Three research areas were determined:

—   the need for a Southern African Institute for Lexicography,

—   collaboration between various stakeholders, and

—   the structure of such an Institute.

1.4   Target group

All known individuals, institutions and bodies involved in lexicography and related professions formed the target group for the research. The various stakeholders forming the target group received a covering letter, an information document explaining the purpose of the research, and a questionnaire.

1.5   Research findings

The main concern from respondents was a lack of coordination of lexicographi-

cal efforts. The respondents wanted collaboration, coordination, training and sharing of expertise in the field but they did not agree on a formal structure such as an Institute.

There was a big concern regarding another formal bureaucratic controlling structure. That an Institute might hamper initiatives and private activities was an underlying fear.

Several respondents suggested that a coordinating body, clearing-house or association should be established as interim structure before deciding on an institute. An association for lexicography could unite interested parties or components, allaying underlying fears regarding individual initiatives or projects. An association could provide expertise, training, information, news, etc.

1.6   Recommendations

1.6.1   Institute for Southern African Lexicography

The feasibility study indicated that the time was not ripe for the establishment of a lexicographical institute. More in-depth discussions in this regard were needed as indicated by the research findings.

1.6.2   Professional association

It was felt that a professional association for Southern African lexicography should be established as soon as possible to address the needs of the respondents regarding communication and the coordination of activities.

1.6.3   Linguistic Association of Southern Africa (LSSA)

It was suggested that the association to be established could function for the time being under the auspices of the Linguistic Association of Southern Africa (LSSA) which formed a task group for lexicography during its annual general meeting of July 1992 in Pretoria.

1.6.4   Aims of a professional association

The professional association should, in order to ensure autonomy, have its own constitution and should formulate its own aims and projects. Some of the aims could include:

—   the establishment of a liaison office or clearing-house to coordinate projects,

—   an e-mail network,

—   a quarterly newsletter,

— an accredited magazine (e.g. *Lexikos* in the AFRILEX Series),

— an annual conference to share professional information, and

— a general meeting (e.g. an annual general meeting or conference) to formulate a national policy regarding lexicography.

1.6.5 Funding

An association would require less funding than an Institute. Membership subscriptions only would not be able to sustain such an association. The people involved in the planning of such an association will have to look into funding possibilities.

1.6.6 Imbalances

An important aspect to be taken into consideration was the existing imbalance between public spending on Afrikaans and English in comparison with the African languages and how this imbalance should be addressed.

1.6.7 Publishers

Lexicography depends to a great extent on publishing houses. Some of the publishers have an excellent record regarding the publication of dictionaries. Others try to make a profit without taking the overhead and research costs involved in the lexicographical process into consideration. Publishers should be encouraged to effectively finance the research and compilation process. Should an association be established, publishers should be encouraged to support the association financially.

1.6.8 The way forward

It was suggested that

— all respondents should be informed of the envisaged professional association;

— all respondents who indicated interest in an association should be contacted to take part in the process of planning such an association;

— all respondents and other interested parties should be invited to become members of the envisaged association;

— a meeting should be called as soon as possible to gather stakeholders and interested parties for the establishment of such a professional association (e.g. at an annual general meeting of LSSA);

—   feedback regarding the outcomes of the feasibility study should be given to all respondents; and

—   the report on the feasibility study should be available to all stakeholders and decision-makers in the field of lexicography.

1.7   Concluding remarks

The results indicated a keen interest in a unifying body among lexicographers and members of related professions. It was clear that there was not as yet a need for an Institute. It was strongly felt that whatever body is established, such a body should not be bureaucratically structured and should not restrict individual freedom, *inter alia* with regard to management and control.

The research team therefore suggested that a professional association for lexicography should be established to meet the needs of lexicographers as well as those of other interest groups.

In 1992, the report on the feasibility study was presented to the Bureau of the Woordeboek van die Afrikaanse Taal. In 1993, the Bureau of the WAT published both an Afrikaans and an English version of the report (see Alberts 1993).

2.   Testing the viability of an Association for Lexicography

Apart from the existing dictionary units, the interests of lexicography in South Africa had up to the beginning of 1995 been served by the Linguistic Society of Southern Africa (LSSA) and the African Language Association of Southern Africa (ALASA). Lexicographers, however, felt the need for the establishment of an association dedicated to lexicography.

In March 1995, Dr M. Alberts and Prof. D. Prinsloo finally decided to draft a questionnaire to test the viability of an Association for Lexicography. The questionnaire was sent to members of LSSA, ALASA, publishers, government departments and even political parties. More than 800 questionnaires were mailed. The reaction and enthusiasm expressed by the respondents was overwhelming.

3.   The establishment of AFRILEX in 1995

There were respondents who seriously considered the outcomes of the 1992/93 feasibility study. With the overwhelming response from the viability study, there was no option but to establish an Association for Lexicography.

On 14 July 1995, several lexicographers, academics and stakeholders came together at the Eighth International Conference of the African Languages Association of South Africa (ALASA) to establish a professional association for lexicography. At the closure of the ALASA conference, the big moment arrived,

and at 11:00 AFRILEX, the African Association for Lexicography, was established.

Dr R.R.K. Hartmann chaired the inaugural meeting, facilitated the whole process as well as the election of office bearers of the new association. Prof. R.H. Gouws was elected Chairperson, Mrs I. Dippenaar and Prof. S. Satyo were elected Vice-Chairpersons, Dr M. Alberts Secretary-Treasurer, Prof. D.J. Prinsloo Conference Organiser, and Mr P. Harteveld, Editor of *Lexikos*, was co-opted to the Board.

The Bureau of the WAT granted permission to this new Association to adopt the name AFRILEX. After the establishment of the African Association for Lexicography, *Lexikos* also became the official mouthpiece of AFRILEX.

In the same year, the Pan South African Language Board (PanSALB) was established in accordance with Act 59 of 1995. The amendment Act of 1999 made provision for the establishment of eleven national lexicography units (NLUs). These NLUs were established as Section 21 Companies, under the supervision of Boards of Directors. The NLUs function under the auspices of PanSALB and is financed by PanSALB. Each NLU is managed by an Editor-in-Chief. The Bureau of the WAT and the DSAE became the NLUs for Afrikaans and English respectively, remaining where they were seated, e.g. in Stellenbosch and Grahamstown respectively. Nine African language NLUs were established and these are hosted at tertiary institutions in the geolinguistic area where the majority first-language speakers live. Employees of the NLUs are members of AFRILEX and participate regularly in AFRILEX activities.

4.   Present situation

AFRILEX promotes and coordinates research in and the study and teaching of lexicography in Southern Africa in its broadest sense. AFRILEX aims to train lexicographers, terminologists and other language practitioners in various aspects relating to lexicographical principles and practice. It also aims to organise regular international conferences and local seminars on topics relevant at a specific time; and to publish a journal and other appropriate literature. This creates an environment conducive to the exchange of ideas and to a mutual stimulus to researchers and practitioners in the field of lexicography. AFRILEX also facilitates tutorials and training courses (e.g. Salex'97 and AFRILEX–Salex'98).

Membership of AFRILEX is open to individuals who and institutions that have an interest in lexicography.

AFRILEX is managed by a Board elected biannually by postal ballot. It holds its annual general meetings during the annual international conferences. The first AFRILEX Board drafted a constitution that was adopted at the first such meeting. Over the years some amendments were made to the constitution so as to keep it updated and relevant.

The AFRILEX Board aims to coordinate conference dates with that of other local associations such as ALASA and LSSA, fostering an interest in the study of language. AFRILEX seeks cooperation with international associations such as the European Association for Lexicography (EURALEX), the Asian Association for Lexicography (ASIALEX), the Australian Association for Lexicography (AUSTRALEX), the Dictionary Society of North America (DSNA), etc. AFRILEX members also attend conferences of these associations whenever possible.

AFRILEX holds its international conferences by invitation at tertiary or lexicographical institutions. The Sesiu Sesotho Dictionary Unit hosted the Tenth International Conference of AFRILEX at the University of the Free State in Bloemfontein from 13–15 July 2005. The theme of this conference was "Monolingual Dictionaries for Africa". This was also the tenth birthday celebration of the establishment of AFRILEX.

The AFRILEX Board and Executive work according to portfolios and each Board member has an allocated task to fulfil, e.g.

> The President oversees the activities of the association, and liaises with other associations.
>
> The Vice-President performs the responsibilities of the President when the President cannot do so.
>
> The Secretary keeps minutes, writes letters, and liaises with Board and general members.
>
> The Registrar updates the address list, manages and maintains the AFRILEX database and sends the address list to *Lexikos* at the Bureau of the WAT.
>
> The Treasurer changes the signatories, processes membership fees, and prepares the auditor's report.
>
> The Organiser organises seminars and tutorials, organises the conference, and liaises with other associations (e.g. for the coordination of dates).
>
> The Webmaster maintains the AFRILEX web page.
>
> The Editor of *Lexikos* serves on the Board as *ex offico* member.

Each Board member tries to promote the Association at various conferences, seminars, symposiums, etc. by taking AFRILEX flyers to such gatherings. AFRILEX members also contribute to the membership drive.

Dr Alberts made the pottery trophies mounted on wood that are usually awarded to keynote speakers and the organiser of international AFRILEX conferences. She also made a big pottery AFRILEX emblem to be used by the Association during conferences or other events.

AFRILEX recognises special contributions that were made by members in honour or on behalf of AFRILEX (e.g. being a chairperson for a number of years). Deserving members are presented with a certificate of merit in recognition of their contribution. The first recipient of the certificate was Dr J.C.M.D. du Plessis, Editor of *Lexikos*. The next two recipients were the two previous chair-

persons of AFRILEX, Proff. R.H. Gouws and D.J. Prinsloo. Their certificates were handed to them at the Tenth International Conference in July 2005.

At this stage, AFRILEX has only one honorary member, namely Prof. A.C. Nkabinde. He was presented with a certificate for exemplary service for his valuable work as lexicographer and as past vice-chairperson of AFRILEX.

AFRILEX maintains a biannual postal voting system. This system was decided upon because not all members are in a position to attend the annual general meeting. With a postal ballot, all members of the Association are able to bring out a vote.

When AFRILEX was established, it was decided to have a regular newsletter. Unfortunately the editor of this newsletter usually has to write all the articles him-/herself. In 2004/05, there was a new effort to revive the AFRILEX Newsletter.

Over several years, some of the AFRILEX members were fortunate to receive Kernerman grants. These grants enabled the recipients to continue research in lexicographical matters or to complete masters or doctoral studies in lexicography. AFRILEX members also successfully apply for grants from EURALEX, ASIALEX, AUSTRALEX, DSNA, the Lawrence Urdang Award for Lexicography, etc.

AFRILEX is in the fortunate position that it has an accredited journal such as *Lexikos* as its official mouthpiece. In 1991, the Bureau of the WAT had taken the initiative to start publishing a journal of lexicography called *Lexikos* in a publication series called AFRILEX, acronym for "lexicography in and for Africa". After the establishment of AFRILEX, *Lexikos* (from volume 6, published in 1996) also became the official mouthpiece of the Association. *Lexikos* serves to promote lexicography in its broadest sense. Members of AFRILEX are invited to share, through *Lexikos*, in lexicographical discussion. AFRILEX members receive *Lexikos* free of charge or at drastically reduced rates. With the Bureau of the WAT, AFRILEX shares responsibility for the future existence of *Lexikos*.

Since the establishment of AFRILEX, Pharos Publishers, one of the leading local publishers of dictionaries, sponsors or co-sponsors the Conference dinner. This is now regarded as a Conference tradition that is much appreciated by the Association.

5.   Conclusion

Members of AFRILEX are all in one way or another involved in dictionary compilation, in different kinds of dictionary work or projects, or in an advisory capacity, or for the mere love, or addiction to lexicography:

—   Some are members of the lexicography teams for the eleven National Lexicography Units (NLUs);

—   Some are compiling terminology lists or technical dictionaries for Language for Special Purposes (LSP);

—   Some are members of advisory bodies (e.g. members of the Pan South African Language Board (PanSALB), members of PanSALB Sub-Committees for various focus areas, members of the Boards of Directors of the National Lexicography Units (NLUs), members of National Language Bodies (NLBs), and members of Provincial Language Committees (PLCs)) ;

—   Some are members of language offices (e.g. the National Language Service (NLS), the Terminology Coordination Section (TCS), the Language Research and Development Centres (LRDCs), Language Units (LUs), etc.);

—   Some are lecturers at tertiary institutions who teach courses in metalexicography and terminology;

—   Some are students of lexicography and terminology; and

—   Some are language practitioners such as translators, editors, interpreters, language teachers and journalists — those who cannot do their specific jobs without the assistance of dictionaries.

None of the AFRILEX members are impartial towards lexicography and they are all interested to learn more about lexicography, dictionaries and how to compile dictionaries or use them to their best advantage.

The members of AFRILEX have a responsibility towards the various speech communities they serve. A dictionary culture needs to be created and speech communities should be aware of AFRILEX. Should the lexicographical endeavours of AFRILEX and its members be known to the general public, and all the work of the South African lexicographers could be available on the envisaged Human Language Technology virtual network, then this region would be in a position to preserve the languages and develop them into functional languages in all spheres of life. The various language communities will only then be in a position to develop their communication skills to the fullest.

If all members of AFRILEX could further the aims of AFRILEX in promoting lexicography in the Southern region of Africa, AFRILEX will have a bright future.

Bibliography

AFRILEX webpage: <http://www.up.za/academic/libarts/afrilang/homelex.html>.

Alberts, M. 1993. *Feasibility Study: Institute for Southern African Lexicography/Lewensvatbaarheidstudie: Instituut vir Suider-Afrikaanse Leksikografie*. Stellenbosch: Bureau of the WAT.

A.P. Cowie. *English Dictionaries for Foreign Learners: A History.* 1999, 232 pp. ISBN 0-19-925084-7. Oxford: Oxford University Press.

In the mid 1990s the evolution of the English learner's dictionary reached a zenith with the appearance on the market of four advanced-level monolingual learners' dictionaries. Three of these were existing works, i.e. the *Oxford Advanced Learner's Dictionary of Current English* (henceforth ALD), the *Longman Dictionary of Contemporary English* (henceforth LDOCE) and the *Collins CO-BUILD English Language Dictionary* (henceforth COBUILD), which marked, to differing extents, significant departures from the lexicographical procedures followed by their predecessors. The last one, the *Cambridge International Dictionary of English* (henceforth CIDE), was a completely new dictionary. This highly productive burst of lexicographical energy was also the catalyst for an even more productive and sustained body of metalexicographical research and writing that dealt with the topic of learners' dictionaries.

The plethora of books on learner lexicography can broadly be categorised as belonging to two main groups: those dealing with "research on dictionary use" (Wiegand 1984: 15) and those providing "criticism of dictionaries" (Wiegand 1984: 15). The latter can then be extended to include those resources that move beyond criticism and into the realm of speculation, such as the book *The Perfect Learners' Dictionary (?)*. There is, however, a third and very important element of metalexicography, which is broadly categorised by Wiegand (1984: 15) as "history of lexicography". This has unfortunately been a sparsely populated niche in the general body of metalexicographical work.

A.P. Cowie's *English Dictionaries for Foreign Learners: A History* is a valuable contribution to the field of metalexicographical research and writing that deals with the topic of learners' dictionaries, as it provides a point of departure for both research on dictionary use and dictionary criticism. Furthermore, it treats the dictionaries not as projects in isolation, but rather as results of a fascinating evolutionary process.

A previous reviewer, however, reacted negatively to the scope of the book, by stating that it only presents a "description of advanced-level learner dictionaries from their inception to 1989" (Nichols 2001: 205):

> That a scholar as respected as Mr. Cowie should choose to gloss over a decade in which the impact of corpus lexicography became fully evident in dictionaries produced with the aid of corpora calls into question the ultimate usefulness of this book as a complete history of the subject.

In my opinion, though, this restriction of its scope is not necessarily a weakness, as a detailed discussion of the new generation of learners' dictionaries would not only lengthen an already formidable read, but also be unnecessary, because many other sources have discussed the methods and structures of these dictionaries (as was indicated in the first paragraph). As it stands, Cowie's work has made an important contribution by filling a vacuum that has existed in current research in this field, by providing a comprehensive (though

not complete) and insightful history that traces the innovative and pedagogical spirit that underpins EFL lexicography back to its roots.

The criticism of the book expressed by Nichols and discussed in the previous paragraphs is counterbalanced, in part, by the following statement in that review (Nichols 2001: 205):

> Rather than in breadth, then, the strength of the *History* lies in its detail: in Mr. Cowie's descriptions of the scope of the task facing the earliest pioneers, and in his painstaking comparisons of the ways various early works treat the myriad elements that make up a learner's dictionary.

This attention to detail is already evident in the first chapter of the book, in which Cowie seeks to describe what he regards to be "The Genesis of the Learner's Dictionary".

The book begins with a brief introduction in which each of the major early role players — Harold Palmer, Michael West and A.S. Hornby — are introduced and cursorily profiled, particularly with regard to their contributions to EFL teaching and applied linguistics. Hereafter, Cowie confines chapter 1 to their lexicographical input and the seminal dictionaries or related books their work led to. It establishes the link between the vocabulary control movement (of which these men were part) and the eventual coming of age of the general monolingual learner's dictionary. Significantly, it also shows the association between vocabulary control and "work on phraseology and grammar" (Cowie 1999: 14), a point that is explored further in the second chapter of the book. All these influences, though, are shown by Cowie to be underpinned by a real concern for — and intimate knowledge of — the problems faced by foreign learners of English (mainly in India and Japan). These three innovators drew from wells of practical experience, which is probably why "nearly all of the elements that are now familiar to learner lexicography can be traced to the work of Palmer, West and Hornby" (Nichols 2001: 205).

From the beginning of the book, Cowie shows the commitment by EFL lexicographers to accommodate the productive needs of their target users alongside the receptive needs that are usually well met by most dictionaries. In chapter 1 the concept of the effective use of grammar representation in the EFL dictionary, especially in aid of text production, is shown to be at the heart of pioneering efforts in this field. Chapter 2 sees a narrowing of this focus and mainly explores the profound influence that the emphasis placed on the natural and adequate reflection of phraseology had on shaping the EFL dictionary as an effective tool for text production. It deals in great detail with the distinctions between different types of combinations and their placement and subsequent treatment in early EFL dictionaries. The distinction between idioms and collocations on the one hand, and collocations and so-called "open" compounds on the other, as well as the polemic criterion of transparency, as applied to collocations, are just some of the issues Cowie deals with in a systematic and very structured, analytical way. The pre-corpus gathering of collocations also merits some discussion. These issues are not, however, viewed in isolation and are always discussed within the broader parameters of the book, i.e. their relevance

in the historical discourse and therefore the manner in which they were addressed by Palmer, West and Hornby, as well as by subsequent compilers of the earliest reference works for learners.

The focus on syntax and phraseology is carried through to the next chapter in which the second generation of learners' dictionaries is discussed and which "spans the period between the compilation and publication of *ALD 2* (in 1963) and the first appearance, in 1978, of the … *LDOCE 1*" (Cowie 1999: 82). However, Cowie expands this focus in order to provide a blueprint for dictionary criticism. The lexicographical treatment of "pronunciation", "verb and adjective patterns", "definitions and glosses", "illustrative examples" and "idioms and phrasal verbs" is evaluated (with minor alterations to these broad categories, depending on the dictionary under evaluation). In chapter 3, this classification is applied to, in turn, ALD 2, ALD 3 and LDOCE 1, with detailed critical analyses being provided within the broader framework for each of these dictionaries. Interestingly, Cowie changes tack in chapter 5 and, instead of discussing each dictionary in turn, identifies broad categories (altering his blueprint somewhat), i.e. "microstructure and macrostructure" (dealing especially with issues of main vs sublemmatisation), "grammatical schemes and codes", "definitions", "style, attitudinal, and register labels" and "sense relations and lexical fields". He then proceeds to give brief critical discussions within these categories of the treatment of each in LDOCE 2, COBUILD 1 and ALD 4, all of which are classified as belonging to the "third generation of learners' dictionaries" and published within a few years of one another in the late 1980s.

In both these chapters, which can arguably be considered to be, as a unit, the crux of the book, the analyses are careful and detailed, plotting with great aplomb the balance between adhering to theoretical insight and paying attention to practicalities. Such analyses are usually most evident in the writings of a metalexicographer with extensive practical experience, as Cowie clearly illustrates. There is, however, a major point of possible criticism, which is again succinctly revealed by Nichols (2001: 206-207):

> Particularly in the third and fifth chapters, we find a bias toward Oxford dictionaries showing itself in Mr. Cowie's comparisons of works that compete with each other, despite his obvious efforts to be fair. Certain opinions that resulted in choices that he made as … editor … are presented as absolutes: his criticism of tightly controlled defining vocabularies, for instance, or his defense of parenthetical elements in definitions to indicate possible objects of a transitive verb.

This criticism is certainly valid and in line with my own observations regarding these chapters, but it can be argued that bias is, in the metalexicographical musings of practical lexicographers, inevitable. It is certainly not unique to this book and is evident in many of the other writings on learners' dictionaries. It is, I would further argue, an inevitable by-product of the commercial polarisation of the learner's dictionary market and the corresponding (commercially motivated) alliance of academics and lexicographers to particular publishing houses. In fact, the previous reviewer (commendably) declares a bias too, having been "trained at Longman as a corpus lexicographer" (Nichols 2001: 207).

The fundamental question is whether the bias is strong enough to detract substantially from the value of Cowie's analyses. I would argue that it is not, and that, despite the points mentioned by Nichols (which should also not be accepted uncritically), the bias mainly has the following result: the analyses of Oxford dictionaries are truly engrossing, the prose comes alive and one is left with the sense of having been afforded a rare glimpse into the mind of the lexicographer, as not only the decisions, but also the thought-processes that motivated those decisions are laid bare. Far from being the weakest parts of the book, these are, in my opinion, the strongest. Furthermore, Cowie does try "to be fair", as is attested to by Nichols, and generally succeeds, which is why "these chapters contain highly useful summaries of the different approaches both of successive and competing editions" (Nichols 2001: 207).

Wedged between these excellent chapters though is a curious attempt to give a nod to the influence of the computer on learner lexicography, which is sadly unsuccessful. Nichols' (2001: 207-209) estimation of this chapter and particularly the deficiencies of its treatment of corpus-based lexicography, is, in my opinion, valid and extremely useful, and cannot be improved upon. I therefore defer to that account.

I would like to add some criticism regarding an area which Cowie rather disappointingly undertreats, i.e. the importance and value of dictionary-making systems. Cowie's observations in this regard are limited to a very brief treatment of "the lexicographical workstation" and little mention is made of the fundamental obstacles that faced the evolution of computer-based lexicography in the time period within which this chapter is situated. These obstacles are, in brief, the transition from elementary word processing to working within structured tag sets, and, eventually, the debate whether to continue working in an SGML- or similarly based text editor, or to make the paradigm shift to working within database-driven dictionary-making software. The eventual evolution of dictionary-making systems has had a profound influence on the editorial process, allowing for faster, more accurate and better integrated compilation of dictionaries, improvements to and more comprehensive automisation of final editing and quality control processes, and the ready export into data formats for paper-based or electronic publishing. That a comprehensive historical overview of these developments is omitted from a book of this extent is a great disappointment.

With the last chapter Cowie seems to be on more solid footing again. It provides a critical and very informative overview of research into dictionary use and presents an especially insightful discussion on the balance that needs to be maintained between using dictionaries for encoding (productive) and decoding (receptive) purposes. His concern for a perceived neglect of the encoding function of learners' dictionaries is especially evident here and he (Cowie 1999: 176) states in this regard:

> This equilibrium, however, has been increasingly challenged since the early 1980s, and, side by side with a general increase in 'user-friendliness', the 1990s have witnessed a shift in favour of designs which prioritize decoding.

He acknowledges that there is ample evidence from user-related studies to motivate such a priority-shift, but proceeds to critically analyse these results and to state a solid case for returning to a more balanced approach when dealing with a larger scale learner's dictionary. It is his focus on the need for the improvement and structured teaching of dictionary skills that is especially compelling. He also comes out in favour of the use of so-called "bilingualized" dictionaries to "serve as a bridge between an introductory bilingual dictionary and the advanced monolingual work" (Cowie 1999: 198). The chapter ends with a summary of the most salient findings and issues regarding user-related research into learners' dictionaries, one which serves as a very useful introduction into the branch of metalexicography dealing with research on dictionary use.

In conclusion, Nichols (2001: 209) aptly credits the book as follows:

> In the end, Mr. Cowie's ability to summarize lexicographic issues and controversies that are, more often than not, quite complex is of help to a reader coming to this topic for the first time.

This finding has been in keeping with my experience in reading this book. Though it is obviously not the last word on this intriguing and vastly important topic, it is a detailed, enlightening and essential introduction to the evolution of learners' dictionaries. If read in unison with sources such as the critical review by Nichols often quoted here, it can provide a solid platform for the critical study of the latest generation of these influential reference works that target such a huge segment of the international lexicographical market.

References

Herbst, T. and K. Popp (Eds.). 1999. *The Perfect Learners' Dictionary (?).* Lexicographica. Series Maior 95. Tübingen: Max Niemeyer.

Nichols, W. 2001. Review of A.P. Cowie. *English Dictionaries for Foreign Learners: A History. Dictionaries* 22: 205-210.

Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K (Ed.). 1984: 13-30. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9–12 September 1983.* Lexicographica. Series Maior 1. Tübingen: M. Niemeyer.

Phillip Louw
Bureau of the Woordeboek van die Afrikaanse Taal
Stellenbosch
Republic of South Africa
(pal@sun.ac.za)

Madaleine du Plessis (Hoofredakteur/Chief Editor). *Pharos 1, Afrikaans–Engels/English–Afrikaans Woordeboek/Dictionary.* Eerste uitgawe/ First edition. 2005, xx + 1 552 pp. ISBN 1868900444. Kaapstad/Cape Town: Pharos Woordeboeke/Pharos Dictionaries. Prys: R450.

Die Pharos van Alexandrië was 'n 152 m-hoë vuurtoring van wit marmer wat in 160 v.C. deur die ou Egiptenare aan die monding van die Nyl op die eiland Pharos gebou is. Pharos was, as vernaamste rigtinggewer en ligbaken, die voorloper van alle latere vuurtorings en een van die sewe wonders van die antieke wêreld. Sedert 1996 het Pharos Uitgewers, met sy bekende ligbakenikoon, as spesialis in die uitgee van woordeboeke en taalhulpmiddels, lig gewerp op talle taalkwessies van die Suid-Afrikaanse taalgemeenskap. Vanjaar was Pharos Woordeboeke verantwoordelik vir 'n nuwe, langverwagte vlagskippublikasie — die omvangryke, enkelvolume *Pharos 1, Afrikaans–Engels/English–Afrikaans Woordeboek/Dictionary.*

Sedert werk aan dié woordeboek in 1994 begin het, moes dit tred hou met die snelveranderende, moderne wêreld met sy gepaardgaande dinamiese korpus kontemporêre Afrikaans en Suid-Afrikaanse Engels. Hierdie publikasie is die eerste tweetalige Afrikaans–Engelse woordeboek na die verskyning van die agtste en laaste uitgawe van die *Tweetalige Woordeboek/Bilingual Dictionary* van Bosman, Van der Merwe en Hiemstra in 1984. Na 21 jaar was hierdie uitgawe vanweë ingrypende maatskaplike, kulturele, sosiale, politieke en tegnologiese veranderinge, plaaslik en oorsee, erg verouderd. 'n Akute behoefte aan 'n kontemporêre Afrikaans–Engelse woordeboek het ontstaan om aan die nuwe taalbehoeftes in 'n meertalige land, wat groter taaleise aan die algemene publiek begin stel het, te voldoen.

Die *Tweetalige Woordeboek/Bilingual Dictionary* beslaan 1 344 bladsye met altesaam 300 000 woorde, frases en uitdrukkings. *Pharos 1* bevat vergelykenderwys 1 552 bladsye met meer as 200 000 trefwoorde en 120 000 frases en uitdrukkings, met ander woorde 'n totaal van 320 000 woorde, frases en uitdrukkings in 'n heelwat langer en breër bladsyformaat. Wat die uitleg betref, het die driekolomindeling van die woordeboekgedeelte behoue gebly. In *Pharos 1* volg sowel die Afrikaanse as die Engelse Gebruiksriglyne/Usage guidelines en Redaksionele afkortings/Editorial abbreviations aan die begin, direk na die voorwoord. In die 1984-uitgawe het die gebruiksriglyne en redaksionele afkortings van elke taal afsonderlik die betrokke taal voorafgegaan. Net die tyd sal leer of woordeboekgebruikers hierdie ruimtelike aanpassing as gebruikersvriendelik sal ervaar en of hulle nie die gebruiksriglyne en redaksionele afkortings van elke taal logieserwys by die betrokke taalhelfte sal gaan soek nie.

Selfs die redaksionele afkortings het die afgelope twee dekades aansienlik vermeerder en word in vier kolomme (teenoor 1984 se driekolomindeling) aangebied — 'n onteenseglike bewys dat die mens se denk- en leefwêreld die afgelope twintig jaar op veral vaktaal-, tegnologiese en sosiale vlak (om maar enkeles te noem) ingrypend verander het en dat dit nie bloot maar as nuwe lem-

mas deur redaksies van woordeboeke verreken word nie, maar alreeds as leksi-kografiese instrumente ingespan word. Dit word weerspieël in nuwe redaksio-nele afkortings in die Afrikaans–Engelse deel, soos *bacteriol. (bacteriology), Jud. (Judaism), palaeontol. (palaeontology), parapsych. (parapsychology), pop mus. (popular music), sci-fi (science fiction),* ensovoorts. Nuwe redaksionele afkortings in die Engels–Afrikaanse deel sluit in: *igt. (igtiologie/viskunde), ing. (ingenieurswese), metafis. (metafisika), MW (maatskaplike werk), Prot. (Protestant, Protestants(e)), ver-losk. (verloskunde)* en *wynb. (wynbou).* Die lyste redaksionele afkortings verleen nuttige gebruiksriglyne by verskeie aspekte van die mikrostrukturele aanbie-ding, waaronder woordsoortkategorieë, grammatikale inligting, etikette en die afkortings- en akroniemlyste.

Die gebruiksriglyne beslaan nege dubbelkolombladsye (Afrikaans links en Engels regs, met 'n relevante voorbeeld in 'n oorvleuelende kolom in die mid-del tussen die twee tale). Die betrokke gebruiksriglyn onder bespreking word dan in die woordeboekartikel, wat as toepaslike voorbeeld in die voorbeeld-kolom gebruik word, teen 'n grys agtergrond uitgelig. Die gebruiksriglyne kom onder die volgende opskrifte aan bod: die struktuur van woordeboekartikels, trefwoorde, vertalings, fleksievorme, woordsoortkategorieë, lettergreepverde-ling en klemaanduiding, grammatikale inligting, etikette, kruisverwysings en afkortings, gevolg deur vier bladsye redaksionele afkortings vir beide taalge-deeltes onderskeidelik.

Die woordeboek bevat ook 'n uitgebreide alfabetiese lys algemene afkor-tings en akronieme aan die einde van sowel die Afrikaans–Engelse deel (pp. 721-745) as die Engels–Afrikaanse deel (pp. 1530-1552). Daar word 'n skeidings-teken gebruik om die Afrikaanse en Engelse vertalings van die woorde, waar-voor die afkorting staan, te skei. In enkele gevalle bestaan daar geen erkende afkorting vir 'n spesifieke vertaling nie en word die afkorting na die vertaling weggelaat, byvoorbeeld:

Joh. Johannes · John

Verwante woorde verskyn alfabeties na die afkorting as dieselfde afkorting vir albei geld. 'n Kommapunt skei verwante woorde met verskillende woordsoort-like funksies, terwyl 'n komma sinonieme skei, byvoorbeeld:

prov. provinsie; provinsiaal, =siale · province; provincial **prov.**

As *Pharos 1* met betrekking tot die gedeelte oor afkortings en akronieme ver-gelyk word met die *Tweetalige Woordeboek/Bilingual Dictionary,* blyk die om-vang van die uitbreiding van veral nuwe tegnologiese afkortings (byvoorbeeld **DOS** skyfgebaseerde bedryfstelsel; → SBS · disk operating system **DOS** en **WAP** koordloseapplikasieprotokol; → DTP, KTP · wireless application proto-col **WAP**) en afkortings vir nuwe organisasies en instansies in 'n veranderde politieke en maatskaplike bedeling (byvoorbeeld **SETA, seta** Sectorial Educa-tion and Training Authority · Sektorale Onderwys- en Opleidingsowerheid **SOOO** en **SAQA** South African Qualifications Authority · Suid-Afrikaanse Kwa-

lifikasieowerheid **SAKO, Sako**). Die *Tweetalige Woordeboek/Bilingual Dictionary* (met sy kleiner formaat) het, wat albei tale betref, altesaam 24 bladsye afkortings. *Pharos 1* met sy groter formaat beskik in totaal oor 48 bladsye afkortings.

'n Nuwigheid in *Pharos 1* wat beïndruk, is die gebruik van etikette, selfs by afkortings, byvoorbeeld: **RWB** Rand Water Board *(hist.)* … en **N.Germ.** Noord-Germaans *(taaln.)* …

As die afkortinglys geografies beskou word, is dit opvallend dat selfs streekspesifieke instansies (soos **NALN**, of te wel die Nasionale Afrikaanse Letterkundige Museum en Navorsingsentrum in die Vrystaat) vermeld word. Daar is ook tred gehou met naamsveranderinge, soos **UV** (Universiteit van die Vrystaat · University of the Free State **UFS**; volledigheidshalwe word die ou naam, **UOVS** Universiteit van die Oranje-Vrystaat · University of the Orange Free State **UOFS** egter nie nagelaat nie, maar ter verduideliking met 'n etiket *(hist.)* gemerk.

Op mikrostrukturele vlak word lemmas alfabeties in vet druk aangebied, byvoorbeeld:

> **speler** *=lers* player; gambler, punter; musician; actor; performer …

Daar word een (of meer) vertalings vir elke betekenisonderskeiding van lemmas gegee. Leestekens speel hierby 'n sleutelrol, deurdat algehele of gedeeltelike sinonieme deur 'n komma geskei word, terwyl 'n groter betekenisverskil deur 'n kommapunt aangedui word. Om verwarring te voorkom, word gebruikliker betekenisse voor minder gebruiklike (waaronder onder andere sleng-, informele en tegniese) betekenisse aangebied. Let veral op die gebruik van die komma en kommapunt in die volgende voorbeeld:

> **folder** omslag; *(rek.)* gids; *(Am.)* voublad, *=blaadjie, =pamflet, =biljet*; vouer, voumasjien; voubeen; vougleuf *(van 'n naaimasjien); (i.d. mv. ook)* knypbril.

Homonieme word apart opgeneem met 'n boskrifnommer na elke lemma.

> **saal**[1] *sale, n.* hall; auditorium; ward *(in a hospital)*; saloon; …
> **saal**[2] *saals, n.* saddle, tack (up) *(a horse); (machinery)* swage; …
>
> **prank**[1] *n.* grap, poets; kaskenade, streek; …
> **prank**[2] *ww.* uitdos, optooi, windmaker(ig) aantrek; versier, dekoreer; spog.

Die funksies van ander leestekens in die woordeboek, soos die weglaatteken (=), die hakies (( )), die skuinsstreep (/), die tilde (~) en dubbele tilde (~ ~) word ook uitvoerig in die gebruiksriglyne uiteengesit. Woordelemente word in die 2005-uitgawe meer konsekwent as in die 1984-uitgawe as aparte trefwoorde opgeneem. Kombinasievorms soos *bio-* ontbreek byvoorbeeld as 'n aparte trefwoord in die 1984-uitgawe, maar verskyn wel as 'n aparte trefwoord in die 2005-uitgawe. Dieselfde geld sommige prefikse soos *ko-*.

Kruisverwysings word aangetoon met behulp van 'n pyltjie (→) of gelykteken (=), gevolg deur die betrokke woord(e) waarna verwys word, en dui steeds

alternatiewe spelvorms vir lemmas aan, soos **aalwee** → ᴀᴀʟᴡʏɴ. Dit lei die woordeboekgebruiker na 'n ander artikel elders in die woordeboek. Selfs die lettertipe (vet, kursief of klein kapitale) dui aan waar die verwysingswoord opgespoor kan word. In die geval van klein kapitale na die pyltjie of gelykteken word die kruisverwysing onder 'n ander artikel in die woordeboek behandel. Daarteenoor kom die verwysingswoord in dieselfde artikel as die kruisverwysing voor in die geval van vet kursiewe letters na die pyltjie of gelykteken. Sommige woordeboekartikels dui selfs soms méér as een kruisverwysing aan om die woordeboekgebruiker na nog meer moontlike vertalings van die lemma te lei.

Pharos 1 bevat nuttige uitspraakleiding wat by elke lemma gegee word: onderstreping van die lettergreep wat die hoofklem dra en gesentreerde punte wat die lettergrepe aandui. Daar word egter nie in die woordeboek gebruik gemaak van fonetiese transkripsies wat veral niemoedertaalsprekers kan help met die uitspraak van woorde nie. Daarvoor vergoed die baie vertaalkeuses in *Pharos 1* egter ruimskoots.

Wat die struktuur van woordeboekartikels betref, bevat sommige woordeboekartikels inligting oor trefwoorde se fleksievome *(intermezzo =mezzi)* en/of woordsoortlike funksies, bv. *n.* (nomen), *ww.* (werkwoord), *adj.* (adjektief), ensovoorts. Waar 'n trefwoord meer as een woordsoortlike funksie het, word die naamwoord eers behandel; dan die adjektief, werkwoord en/of bywoord. Sintaktiese en morfologiese leiding word gegee deur die trefwoord na die vertaalkomponent deur middel van kursief gedrukte frases te illustreer. Interessant genoeg word afleidings van die lemma as afsonderlike trefwoorde opgeneem, terwyl samestellings met die trefwoord as subtrefwoorde in dieselfde woordeboekartikel behandel word. As 'n trefwoord egter slegs afleidings bevat, word dit alfabeties in dieselfde artikel aangebied.

Morfologiese inligting ten opsigte van fleksievorme word in kursief gedrukte vorm direk na die trefwoord vermeld.

> **huis** *huise*
> **huisie** *=sies*

Die manlike en vroulike vorme en verkleinwoorde van selfstandige naamwoorde word soms as afsondelike lemmas hanteer en soms, waar nodig, selfs van etikette voorsien, bv. *manlik (masc.)/(ml.), vroulik (fem.)/(vr.)* en *verkleinwoord (dim.)/(vkw.).* As daar na twee soortgelyke verkleinwoordlemmas soos **chuckie** *(vkw.)* hoender; liefste, skat, en **kitty** katjie, kietsie gekyk word, kan die konsekwentheid van die verkleinwoord se etiketaanduiding in die woordeboek egter in 'n mate bevraagteken word weens die afwesigheid van die etiket *(vkw.)* in laasgenoemde geval.

Alle grammatikale inligting word tussen hakies in kursiewe druk aangetoon. Dit sluit onder andere in naamwoorde se enkelvoud- en meervoudstatus, die attributiewe en predikatiewe gebruik van adjektiewe, die verledetyds- en verlededeelwoordsvorm van sommige werkwoorde, werkwoorde se oorganklikheid en onoorganklikheid, ensovoorts.

In gevalle van naamwoorde met reëlmatige en onreëlmatige meervouds-vorme, word beide opsies aangetoon, byvoorbeeld: **akademikus** *=mikusse, =mici.* Adjektiewe se trappe van vergelyking *(=er* en *=ste* en soms *meer* en *mees(te))* word direk na die lemma gegee. Uitsonderlike trappe van vergelyking word volledig vermeld, byvoorbeeld: **baie** *meer mees(te)*. Onverboë vorme van die attributiewe byvoeglike naamwoord word deur middel van 'n tilde ná die trefwoord aan-gedui.

In die Afrikaans–Engelse gedeelte van die woordeboek word etikette in Engels aangegee, teenoor Afrikaanse etikette in die Engels–Afrikaanse helfte van die woordeboek. In gevalle waar woordeboekgebruikers onseker is oor die afkortings van etikette, kan die volledige lys afgekorte redaksionele etikette net na die Gebruiksriglyne vanaf p. xvii tot p. xx opgespoor word. Etiketgebruik in 'n tweetalige woordeboek soos *Pharos 1* voorsien die woordeboekgebruiker van nuttige gebruiksleiding, veral in die geval van die fynere nuanses van die taal waarmee anderstalige woordeboekgebruikers minder vertroud is. Nuttige eti-ketgebruik in die woordeboek dui byvoorbeeld aan of 'n betrokke lemma *verou-derd, sleng, vulgêr* of *taboe* is. Rassistiese benamings word steeds opgeneem, maar is voorsien van duidelike gebruiksleiding aan die hand van etikette, by-voorbeeld:

> **rooie** *rooies* red one; *(R~, infml., chiefly hist., also derog.)* Englishman, Brit; jingo; Red, Communist; ...

Die volgende kategorieë etikette kan geïdentifiseer word: stilistiese, temporele, vaktaal- en geografiese etikette. Stilistiese etikette dui op die sosiolinguistiese waarde wat aan 'n bepaalde gebruiksmilieu gekoppel kan word. Etikette wat die stilistiese gebruiksfeer kan tipeer, sluit in *formeel (fml.), informeel (infml.), sleng (sl.), neerhalend (derog.), plat (coarse)* en *vulgêr (vulg.)*. Temporele etikette dui op tydsaspekte wat spesifiek betrekking het op bepaalde lemmas wat nie meer as deel van die alledaagse standaardtaal aangevoel word nie, byvoor-beeld *argaïes (arch.)* en *verouderd (obs.)*. Geografiese etikette dui gewoonlik op streek- of gebiedspesifieke woorde sonder standaardtaalstatus en word dan soos volg aangedui: *(Du.) = Dutch*. Etikette met Bybelse verwysings kom ook voor en daar word selfs gespesifiseer of 'n verwysing uit die Ou of Nuwe Tes-tament kom, byvoorbeeld: **Mispa** *(OT, geog.)* Mizpah. Met die toenemende rekenarisering van woordeboeke, die Bybel en ander gesaghebbende werke, sou etikette moontlik voortaan volledigheidshalwe die spesifieke Bybelboek, hoofstuk en vers kon aandui. Ook inheemse tale word etiketsgewys verreken, byvoorbeeld: **kaia** *=as, (<Ngu.)* kaya, shack. Vaktaaletikette toon lemmas aan wat in vaktaalverband gebruik word, byvoorbeeld: *(psych.) = psychology*. 'n Voorbeeld op mediese terrein is: **leptospirose** *(med.)* leptospirosis, Weil's dis-ease, en op ekonomiese gebied: **inflasie** *(econ.)* inflation; ~ *veroorsaak* cause inflation, be inflationary; *voorstander van* ~ inflationist…

Weens die aard van 'n tweetalige woordeboek soos *Pharos 1*, is die verre-kening van semantiese en pragmatiese aspekte met betrekking tot (veral kul-

tuurspesifieke) verskille tussen die bron- en doeltaal, van groot belang — meer bepaald vir vertalers. Dit is belangrik dat leksikograwe kultuuraspekte wat inherent aan 'n taal is, moet aandui. Hierin speel die gebruik van etikette, soos hierbo beskryf, 'n groot rol. Deur die hantering van een konsep wat in die Afrikaans–Engelse gedeelte voorkom, te vergelyk met die hantering van dieselfde konsep in die Engels–Afrikaanse gedeelte en ook deur die vroeëre en huidige woordeboeke te vergelyk, kan heelwat kultuurspesifieke inligting (in *Pharos 1* meer volledig as in die 1984-uitgawe) bekom word.

Eerstens word die hantering van 'n konsep in die *Tweetalige Woordeboek/ Bilingual Dictionary* vergelyk met *Pharos 1* se behandeling daarvan na verreikende veranderinge in die politieke milieu die afgelope dekade in Suid-Afrika. In die Afrikaans–Engelse gedeelte van die *Tweetalige Woordeboek/Bilingual Dictionary* is die hantering van 'n lemma met kultuurspesifieke elemente, soos *gelofte*, soos volg:

> **gelofte** =s vow, covenant, solemn promise. **G~dag** Day of the Covenant. **G~fees** Festival of the Covenant. **G~kerk** Church of the Vow (Covenant).

Dit verskil in 'n groot mate van *Pharos 1* wat meer sintaktiese leiding verskaf deur kollokasies met die lemma te voorsien, groter morfologiese leiding deur meer samestellings te gee (soos **~aflegging**) en meer kultuurspesifieke inligting (soos temporele en geografiese verwysings) deur middel van etikette te voorsien, soos blyk uit die volgende artikel:

> **gelofte** *=tes* vow, covenant, solemn promise. *'n ~ aflê/doen om iets te doen* make/ take a vow to do s.t.; *die ~ aflê* take one's vows *(as a member of a religious order); deur 'n ~ gebind word om iets te doen* be under a vow to do s.t. **~aflegging** vow-taking, pledge-taking. **G~dag** *(SA hist.: 16 Dec.)* Day of the Covenant. **G~kerk** Church of the Vow/Covenant *(in Pietermaritzburg).*

In die Afrikaans–Engelse gedeelte van *Pharos 1* is daar onder *Dingaansdag* 'n etiket *(hist.)*, met 'n kruisverwysing na *Geloftedag*. By *Geloftedag* is daar egter nie 'n kruisverwysing na *Dingaansdag* nie. Interessant genoeg ontbreek die lemma, *Dingaan's Day* in *Pharos 1* se Engels–Afrikaanse gedeelte geheel en al. Of die weglating in die Engels–Afrikaanse gedeelte iets van die kultuurspesifieke sentimente of belangeverskille tussen Afrikaners en Engelse sê, is oop vir debattering. Ook kan die hantering van 'n lemma soos *Versoeningsdag* nog verder bestudeer word. Dalk is *kultuurbelangegroep* 'n verdere kwalifikasie wat volledigheidshalwe vir 'n kultuurspesifieke lemma soos hierdie onder etikette opgeneem behoort te word.

Vervolgens word die Engels–Afrikaanse gedeelte se hantering van dieselfde lemma in die *Tweetalige Woordeboek/Bilingual Dictionary* vergelyk met *Pharos 1*. By die lemma *vow* was daar min kultuurspesifieke elemente; daar is dus gekonsentreer op die lemma *covenant.* Die vroeëre en huidige woordeboeke is identies, behalwe dat daar 'n kultuurspesifieke, temporele, geografiese etiket by *Day of the Covenant* bygekom het, naamlik: *(SA, hist.).*

> **covenant** *n.* verbond, verdrag, ooreenkoms, akkoord; handves; gelofte; *Ark of the*
> *C~* Verbondsark; *Church of the C~* Geloftekerk; *Day of the C~ (SA, hist.)*
> Geloftedag; *Festival of the C~* Geloftefees; *God of the C~* Verbondsgod; *~ of*
> *grace* genadeverbond; *people of the ~* (ver)bondsvolk. **covenant** *v.* ooreen-
> kom, 'n verbond *of* verdrag sluit. **~theology** verbondsteologie.

Uit bogenoemde bespreking is dit duidelik dat nie net semantiese en pragma-
tiese aspekte meer aandag (deur middel van etikette) in *Pharos 1* gekry het nie,
maar ook sintaktiese leiding (deur middel van meer kollokasies met die lem-
ma).

Met pleknaamsveranderinge wat in die nuwe politieke bedeling 'n rege-
ringsprioriteit is (en wat vanselfsprekend die verandering van vele Afrikaanse
pleknaamlemmas en geografiese verwysings kan teweeg bring), kan die baie
pleknaamlemmas en geografiese verwysings in etikette in *Pharos 1* veroorsaak
dat die nuwe woordeboek moontlik voortydig verouderd kan voorkom soos
pleknaamsveranderinge gestandaardiseer en gewettig word. Sien byvoorbeeld
bogenoemde artikel **gelofte** waarin daar in 'n etiket na *Pietermaritzburg* verwys
word. So ook selfstandige lemmas soos *Pretoria, Pretorianer*, ensovoorts. Etikette
in latere hersienings van *Pharos 1* sou moontlik kon tred hou met pleknaams-
veranderinge deur die vorige pleknaamlemma (met dateringe tot watter jaartal
die naam in Suid-Afrika relevant gebly het) in die woordeboek te behou en ook
die nuwe pleknaam op te neem.

Vanuit makrostrukturele perspektief beskou, kan drie stappe, wat deur
Pharos 1 se redaksie gevolg is, geïdentifiseer word. Eerstens moes heelwat ar-
gaïsmes, verstokte uitdrukkings, Nederlandismes en voornaamlemmas uiter-
aard weens onbruik sedert die vorige woordeboek geskrap word. Soms is ouer
woorde nie weggelaat nie, maar net van 'n temporele etiket (byvoorbeeld *hist.*)
voorsien, soos in:

> **babbeljoentjie** *=tjies, (hist.)* bed canopy; four-poster (bed).

'n Ander argaïese woord soos die volgende is egter nie van 'n (temporele) etiket
voorsien nie:

> **kwispedoor** *=doors, =dore* spittoon, spit box, cuspidor(e).

Die tweede stap was die opname van nuwe woorde en uitdrukkings. Die
woordeboek bevat talle woorde wat vir die eerste maal in 'n tweetalige woor-
deboek opgeneem is, ten spyte van die feit dat hulle al vir 'n aansienlike tyd in
alledaagse gebruik is, soos: *aardverwarming/global warming, pryslied/praise poem,
isibongo* (Nguni), *bosberaad/lekgotla* (Sotho/Tswana), *toi-toi/toyi-toyi, rakmelk/
long-life milk, satellietskottel/satellite dish, kuberkraker/hacker, fopdosser/drag queen,
minimaletoegangchirurgie/laparoscopic (minimally invasive) surgery, keyhole surgery*
(informeel), *skeidingspakket/redundancy payment, aardverwarming/global warm-
ing, CD-speler/CD player, telefoonkaart/phonecard, tsoenamie/tsunami*, ensovoorts.
Let veral op die verwysings na ander inheemse tale wat aan *Pharos 1* 'n ryker

geskakeerdheid gee wat die veeltalige taalwerklikheid in Suid-Afrika en uiteraard ook elke Suid-Afrikaner se noodwendig gedeelde veeltalige algemene kennis goed weerspieël.

Laastens is die inhoud aangepas by die jongste skryf- en spelreëls, soos onder andere vervat in die nuutste *Afrikaanse Woordelys en Spelreëls* wat in 2002 verskyn het. Voorts word *Pharos 1* deur omvangryke papier- en elektroniese teksbronne (wat die *Pharos-Media 24*-korpus koerant- en tydskrifargiewe, die *Pharos-NB*-korpus en sowat tien gesaghebbende Engelse woordeboeke, asook die *British National Corpus* insluit) gerugsteun. Uit bogenoemde stappe wat gevolg is, is dit duidelik dat *Pharos 1* se redaksie die hedendaagse teikenmark goed opgesom en voor oë gehou het tydens die saamstel van die woordeboek.

Tradisionele leksikografie moes vroeër op handgeskrewe indekskaartjies steun. Nuwe tegnologie stel leksikograwe egter deesdae in staat om inskrywings te rekenariseer en selfs te kodeer — 'n praktyk wat ook in *Pharos 1* toegepas is. Deur verlangde data (soos retrograde, alfabetiese, frekwensie- en woord-in-kontekslyste) rekenaarmatig uit *Pharos 1* se databasis te onttrek, kan kleiner satellietwoordeboeke (soos sakwoordeboeke) dus voortaan moeiteloos deur Pharos saamgestel word.

Pharos 1 se teksontwerp word gekenmerk deur 'n oopheid van uitleg en 'n gebruikersvriendelike tipografie wat die maklike naslaan van trefwoorde verseker. 'n Verdere pluspunt lê op die praktiese vlak — bladsye wat oopbly waar die woordeboekgebruiker besig is. *Pharos 1* is 'n welkome ligstraal in 'n talige duisternis — 'n goeie voorbeeld van dié talige duisternis was die Babelse verwarring wat vir effek in die radioadvertensies vir die bekendstelling van *Pharos 1* gebruik is. Dit het ontstellend baie na hedendaagse (sosiale) taalgebruik (in Suid-Afrika se veeltalige opset) geklink.

Dié noodsaaklike publikasie wat in 'n groot mate finansieel gesteun is deur die L.W. Hiemstra Trust, dek die veelvuldige spektrum terreine wat die moderne mens se leefwêreld kenmerk. Hierdie tweetalige *tafel*woordeboek (omvangryker en swaarder as 'n *hand*woordeboek) is 'n gesaghebbende naslaanwerk en 'n onontbeerlike hulpmiddel vir elke vertaler en tolk, opvoedkundige instansie, huishouding en werkplek om suksesvolle en doeltreffende kommunikasie in Afrikaans en Engels te verseker.

Corlietha Swart
Afrikaanse Taalkunde
Departement Afrikaans en Nederlands, Duits en Frans
Universiteit van die Vrystaat
Bloemfontein
Republiek van Suid-Afrika
(SwartC.blms@mail.uovs.ac.za)

Geoffrey Leech, Paul Rayson and Andrew Wilson. *Word Frequencies in Written and Spoken English.* 2001, xv + 304 pp. ISBN: 0 582 32007 0. Harlow: Pearson Education. Price: £45.99

Geoffrey Leech, Emeritus Professor in the Department of Linguistics and Modern English Language at Lancaster University, has been the co-editor and co-author of much research on English grammar, and computational and corpus linguistics. During the past thirty years, his major focus has been the analysis and processing of English with the building of the 100-million word British National Corpus, composed of modern English written texts as well as spoken transcriptions. Derived from a new version of this large electronic corpus, Leech *et al*'s word-frequency book is a very comprehensive database for the linguist and lexicographer alike. In the past, frequency lists tended to focus on written language only, but this book has the merit of covering not only present-day written English, but also of encapsulating lexical frequency in modern spoken language. Spoken data only makes up 10% of the British National Corpus, however, as the transcription and digitisation of spoken language is extremely time-consuming. The spoken data in this book has been divided into a conversational part which includes informal dialogues (recordings of everyday spontaneous interaction) and a task-oriented part that represents 60% of the spoken data of the corpus and includes more formal audio material (lectures, TV and radio broadcasting).

The book begins with a 20-page introduction, which is sometimes quite technical, but which is indispensable for familiarising oneself with the data. It also gives guidelines on how to read the frequency lists. This introduction further discusses problems of data collection, transcription and spelling conventions. Each frequency list appears with a frequency score of tokens per million words. The book is structured into six main chapters that collect together the different word-frequency lists. Chapter 1 presents the whole corpus. The words are ranked without any distinction between speech and writing. Chapter 2 looks at word frequency across speech and writing. In chapter 3, the focus is on the spoken part of the BNC corpus with comparison between the conversational and task-oriented subcorpora. The written data of the corpus is the subject of chapter 4, comparing imaginative and informative writing. Thanks to a meticulous word-class tagging of nouns, adjectives, verbs etc., this book is not limited to word forms alone, but classifies lexical items according to their part of speech. In chapter 5, frequency tables give the most frequent words within each grammatical word class and the frequency of each grammatical class is presented in the final chapter.

The corpus used is varied enough, drawing from multiple sources, and is far more up-to-date than previous word lists representing tendencies of spoken as well as contemporary written English. For the lexicographer, this is an essential tool which throws light on the trends in modern English: it foregrounds words that belong to the core vocabulary of the language such as tool-

words (articles, prepositions, conjunctions etc.), modal verbs and auxiliaries, verbs of movement and thinking, and draws a contrast with words of lower frequency that would be worth including in a compact dictionary or a lexicographic collection, or on the contrary should be excluded in favour of more frequently-used forms or variants. The frequency lists allow the lexicographer to view the data sorted and presented according to different criteria (alphabetically, by descending order of frequency and by construction). The corpus is large enough for him/her to analyse trends and track changes diachronically or synchronically, establish the vitality of rare words, and is also available on the Internet in electronic form. For corpus linguists, this book is of paramount importance for analysing and establishing the saliency of lexical items with a comparative corpus, such as journalistic or scientific English.

Similar word-frequency lists have been compiled for Romance languages such as Spanish, Rumanian and French (Juilland *et al* 1964, 1965, 1970, Baudot 1992), but these studies all need updating. There also exist word lists in each respective language that have been devised using concordancers, but they are usually quite limited. It is hoped that Leech *et al*'s book will lead the way for the compilation of comprehensive word lists in other foreign languages.

References

Baudot, J. 1992. *Fréquences d'utilisation des mots en français écrit contemporain.* Montreal: Les Presses Universitaires de Montréal.

Juilland, A., D. Brodin and C. Davidovitch. 1970. *Frequency Dictionary of French Words*. The Hague/Paris: Mouton.

Juilland, A. and E. Chang-Rodrigues. 1964. *Frequency Dictionary of Spanish Words*. London/The Hague/Paris: Mouton.

Juilland, A., P.P. Edwards and I. Juilland. 1965. *Frequency Dictionary of Rumanian Words*. London/The Hague/Paris: Mouton.

Michaël Abecassis
Department of Modern Languages
University of Oxford
Oxford
United Kingdom
(michael.abecassis@modern-languages.oxford.ac.uk)

Ingrid Lemberg, Bernhard Schröder and Angelika Storrer (Editors).
Chancen und Perspektiven computergestützter Lexikographie. Lexicographica. Series Maior 107. 2001, 270 pp. ISBN 3-484-39107-3. Tübingen: Max Niemeyer. Price: € 72.

The present volume is a collection of thirteen papers resulting from a symposium on the computer-aided production and publication of dictionaries held in Heidelberg in 1988 (p. 4). It is subdivided into two sections, Basics (*Grundlagen*), which discusses the impact of computational methods on lexicography in general, and Applications (*Anwendungen*), which contains descriptions of various existing computer-aided lexicographical projects. All of these projects deal with German as the object of lexicographical description and many of them are concerned with specialized scholarly dictionaries, often of historical stages of German. Thus two articles (Burch and Fournier; Plate and Recker) deal with Middle High German and another one (Gloning and Welter) with the language of Goethe. Some articles, however, also describe projects dealing with present-day German and the design of general purpose dictionaries. Thus Petelenz discusses information design in an electronic Polish–German dictionary and Haß-Zumkehr focuses on the organisation of the microstructure in the LEKSIS-database of the Institut für Deutsche Sprache in Mannheim.

There are several themes of computational lexicography which resurface time and again in many of the articles. One such theme is the question of the encoding of lexicographical data. Generally, the authors recommend SGML or XML for this purpose because both markup languages are sufficiently powerful as well as platform and software independent. SGML and XML feature prominently in several articles. Thus Büchel and Schröder give an introduction to the basic concepts of data markup as well as the storage of data in database systems; Schmidt and Müller set out to develop a new model of lexicographical data structuring, which clearly separates the level of the storage of structured data from the level of data presentation of the user; and Burch and Fournier discuss the application of the guidelines of the Text Encoding Initiative (TEI) in digitising the entries of traditional print dictionaries of Middle High German, arguing that the difficulties they encountered by and large do not result from weaknesses in SGML but rather from the print dictionaries themselves, which are not consistent in their presentation of information.

Another dominant theme is the "added value" which authors see in computer-aided lexicography both for lexicographers as well as users. With respect to the compilation process, Plate and Recker discuss the creation and use of a digitised text archive as an empirical basis for the creation of a new Middle High German dictionary. In contrast to the time-honoured process of collecting citations first and then writing up the lexicographical entry on the basis of the collected citations, lexicographers can now directly access and electronically search through the full text of the archive, which will result in a more balanced view of the properties of the lemma. As Lemnitzer shows, electronic online dictionaries can also help the lexicographer to find out more about users' needs by registering the frequency and kind of users' queries as well as their success

rate. Also of interest for lexicographers is computer-aided research in the neighbouring field of lexicology; in this vein, Kunze and Wagner present the GermaNet project and compare it to its American cousin WordNet.

From the user perspective the added value of electronic dictionaries includes the multimedial possibilities of the electronic platform as well as the essential flexibility in data presentation. Thus Richter explains how a dictionary of *Flurnamen* (i.e. names containing topographical elements) can be enriched by material such as aerial photographs or pronunciations of local speakers, which in a printed dictionary is possible only to a very limited extent (pictures) or not at all (sounds). More generally, this point is also argued by Storrer, who explores the possibilities of the hypertext concept in lexicography. The essential flexibility in data presentation materialises in the possibility to access the dictionary in various ways in complex searches as well as the possibility to adapt the selection and display of data to the specific needs of the user as is argued for example by Klosa in her article presenting assessment criteria for the quality of CD-ROM dictionaries.

In sum, the anthology gives a well-rounded picture of aspects of electronic lexicography. However, since electronic lexicography has been developing at such a fast pace since the book was published, not everything rings quite so new (and newsworthy) as it might have rung five years ago. For example, from today's perspective there is little need to argue in favour of SGML/XML, the latter of which is nowadays generally and uncontroversially regarded as a markup standard.

Other aspects of the book are clearly outdated; for example, Storrer's complaint about electronic dictionaries more or less directly copying their corresponding print dictionary (p. 54) was certainly justified at the time of writing but is much less valid today. A good up-to-date CD-ROM dictionary such as the electronic version of the *Longman Dictionary of Contemporary English* (4th ed.) shows a host of advanced features including additional material, field-based and spelling-tolerant searches, and multimedia features, i.e. exactly those desiderata called for by authors in the present volume. Not surprisingly, this outdatedness also concerns internet links. Thus the link given for Burch and Fournier's Middle High German project <gaer27.uni-trier.de/mwv-online/mwv-online.html> (p. 135) as well as the link given by Lemnitzer (p. 248) for the dictionary publisher Zeres <www.zeres.de/dict> is broken and gives an error message.

Nevertheless, readers who are new to the field of electronic lexicography and have a special interest in German lexicography will find the anthology a rewarding read.

Michael Klotz
Institut für Anglistik und Amerikanistik
Universität Erlangen-Nürnberg
Erlangen, Germany
(mklotz@phil.uni-erlangen.de)

Sentrum vir Tekstegnologie. *Afrikaanse Speltoetser 2.1 en Woordafbreker vir Microsoft Office®*. 2004. Potchefstroom: Noordwes-Universiteit.

Afrikaanse Speltoetser 2.1 en Woordafbreker vir Afrikaans is tans die nuutste in sy soort op die mark en kom van die Potchefstroomkampus van die Noordwes-Universiteit. Dit is 'n uitbreiding en verbetering op die vorige weergawe 2.0 en doen die beloftes wat op die omslag van die CD gemaak word, meestal gestand.

Dit is 'n vinniger en omvattender produk as die vorige weergawe. Dit installeer baie maklik, waarby gevoeg moet word dat die gebruikershandleiding baie deeglik en verstaanbaar opgestel is, en dit dus vir enige rekenaargebruiker maklik behoort te wees om dié speltoetser te installeer én te gebruik. Die woordafbreker werk op dieselfde wyse as dié van ander tale en is goed geïntegreer met Microsoft Office.

Die speltoetser vergelyk goed met die verstekspeltoetsers van Microsoft Office, en gebruikers wat bekend is met byvoorbeeld die Engelse speltoetser van Microsoft Office sal *Afrikaanse Speltoetser 2.1* maklik vind om te gebruik. Die gebruiker sal intuïtief sy pad vind. Een so 'n voorbeeld is die lys van alternatiewe wat vir die verkeerd gespelde woord (in rooi onderstreep) gegee word. Daarmee saam is daar ook die opsie om die onbekende woord by die verstekwoordeboek te voeg deur op "Add to dictionary" te klik. Dit is 'n baie nuttige funksie vir gebruikers wat gereeld met vaktaalterme, samestellings en ander minder bekende leksikale items te make het.

Nog 'n funksie wat inderdaad waarde toevoeg tot die speltoetser, is die regsklikfunksie. Dit bied vinnige toegang tot alternatiewe vir die woord wat verkeerd gespel is, en die gebruiker kan deur regs te klik, hierdie alternatiewe gebruik, ignoreer of tot die verstekwoordeboek toevoeg. Die outomatiese regstellingfunksie (Auto correct) word ook hier as 'n opsie aangebied, wat die gebruiker help om woorde wat dikwels of uit gewoonte verkeerd gespel word, onmiddellik outomaties te korrigeer. Onder die regsklikfunksie is daar ook 'n opsie om die woord wat as foutief gespel aangedui word, verderaan te ignoreer of , indien die woord 'n leenwoord is, die taal te kies waarin dit voorkom, sodat dit nie langer as 'n fout aangedui word nie.

Volgens die omslag van die CD-ROM, herken hierdie *Speltoetser* 'n "groot aantal nuwe samestellings en verbuigings van woorde". Dit is egter nie duidelik op grond waarvan sekere samestellings herken of nie herken word nie. *Springbokhaker* word byvoorbeeld herken, maar nie *Springboksenter* nie. *Topgehalterugby* word ook nie herken nie, maar wel *hoëgehaltestaal, topklasrugby* en *topgehalteproduk*. Die opstellers sal miskien weer moet aandag gee aan die herkenning van komposisionaliteit al dan nie, en vir die gebruiker leiding gee in hierdie verband. Intussen is die "Add"-funksie wel baie nuttig, waar bestaande woorde wat nie deur die speltoetser herken word nie, tot die verstekwoordeboek toegevoeg kan word.

'n Ander probleem wat ook na vore tree, is die aanduiding van die ver-
dubbeling van die persoonlike voornaamwoord as 'n fout. In 'n sin soos *As 'n*
mens na die syfers kyk, is dit onduidelik hoekom **hulle hulle** *nie tot die bank gewend*
het nie, is dit onduidelik waarom *hulle* in die tweede posisie as 'n fout aangegee
word. Daar word ook nie 'n alternatief vir die oënskynlike fout aangebied nie.
Dieselfde gebeur ook met *julle julle*.

Afrikaanse Speltoetser 2.1 is inderdaad 'n welkome verskyning op die rak
van Afrikaanse speltoetsers en 'n groot aanwins vir vertalers, taalpraktisyns en
ander persone wat in Afrikaans skryf. Indien die genoemde probleme in die
volgende weergawe uitgestryk kan word, behoort hierdie produk met waar-
digheid sy plek langs ander internasionale speltoetsers in te neem — veral wat
betref Microsoft Office se verstekspeltoetsers vir ander tale.

Gerda de Wet
Stellenbosch
(dewets@ananzi.co.za)

W.F.J. Steenkamp en/and H.J.J. Reynders (Hoofredakteurs/Chief Editors. *Ekonomiese en Bedryfswoordeboek: Engels–Afrikaans/Dictionary of Economics and Business: English–Afrikaans*. Tweede, hersiene en vermeerderde uitgawe/Second, revised and enlarged edition. 2004, xii + 1199 pp. ISBN 1868900312. Kaapstad/Cape Town: Pharos Woordeboeke/Dictionaries. Prys/Price: R400.

Die *Ekonomiese en Bedryfswoordeboek/Dictionary of Economics and Business* (voortaan afgekort as EBW) is 'n hersiene en vermeerderde uitgawe van die *Ekonomiewoordeboek/Economics Dictionary* (EW) van W.F.J. Steenkamp wat reeds in 1976 verskyn het. In die voorberig van EW sê Steenkamp: "Die boek was ruim 'n kwarteeu in die maak" en "As hierdie laat vrug van ons taalbewegings egter daartoe bydra om die Afrikaans van die ekonomiese wetenskappe en die bedryfslewe te suiwer en te versterk, sal 'n mens jou met hierdie nalopery maklik genoeg kan versoen". En nou, na nog 28 jaar, verskyn die EBW onder die redaksiekomitee van wyle prof. W.F.J. Steenkamp en drr. C. de Jong, S.J. Kleu en H.J.J. Reynders. Hierdie EBW is dus die resultaat van meer as 50 jaar se entoesiasme vir Afrikaans en skerp taalkundige insig van al die instansies en persone wat op een of ander wyse 'n bydrae gelewer het tot die inhoud en totstandkoming daarvan. Met Engels as brontaal het die getal lemmas meer as verdubbel — van 39 000 tot oor die 90 000.

Die EBW lê in die kader van die gespesialiseerde leksikografie. Die omvattende behoefte deur dosente en studente, onderwysers en skoliere, skrywers van handboeke en vertalers, joernaliste en ander mediawerkers, staatsamptenare en sakelui aan so 'n gespesialiseerde woordeboek is veral gestimuleer deur 'n sterk groei in die ekonomiese literatuur en in dié van verwante wetenskappe wat, saam met die ontwikkelings in die bedryfslewe, om Afrikaanse terme gevra het. Sonder om aanspraak te maak op volmaaktheid voorsien die EBW dus minstens omvattend in die behoefte aan 'n Afrikaanse vakwoordeboek in die Ekonomiese en Bestuurswetenskappe met Engels as brontaal.

Die EBW as gespesialiseerde leksikografie bestryk die terrein van die makro- en die mikro-ekonomie. By die mikro-ekonomie is hoofsaaklik gefokus op die terme wat die meeste gebruik word in al die deelgebiede van die bedryfsekonomie. Die terme wat eintlik net deur die vakwetenskaplike en -kenner in 'n bepaalde deelgebied gebruik word, is dus uitgesluit. Voorts is sekere geselekteerde terme uit die onderskeie produksie- en bedryfstakke en verwante vakgebiede soos die statistiek en die versekeringswese, en wat dikwels deur ekonome en sakelui gebruik word, ook by die EBW ingesluit.

Die behandeling van die terme wat in die EBW opgeneem is, is besonder gebruikersvriendelik. Hierdie stelling word bevestig deur slegs in breë trekke te verwys na enkele hoofelemente van die behandeling.

— Die opname van trefwoorde is streng alfabeties maar met die gerief dat, indien 'n trefwoord meer as een betekenis het, die verskillende beteke-

nisse onder mekaar geskryf is. Vir die gebruikers wat die woordeboek raadpleeg ten einde 'n beter begrip van 'n gespesialiseerde konsep binne 'n bepaalde studieveld te kry, is hierdie sistematiese ordening besonder nuttig. In hierdie sistematies geordende woordeboek word konsepte gerangskik volgens die verwantskap wat hulle met mekaar het. Die resultaat hiervan is 'n beter geheelbeeld van die bepaalde konsep.

— Baie moeite en noukeurige sorg is aan sinonieme en wisselvorme in die EBW gegee. So verskyn sinonieme en wisselvorme aan die Engelse kant alfabeties tussen hakies en word dit aan die Afrikaanse kant met kommas geskei. As die grondwoord meerdere sinonieme het, is 'n volle inskrywing van sinonieme by een grondwoord gedoen. 'n Lys van sinoniemgroepe in Afrikaans en Engels is ook saamgestel met, waar nodig, 'n aantekening oor die gebruik van die gegewe terme. Die voorkeurterme by sinonieme is in die woordelys met die boteken ' gemerk. Al hierdie handelinge rondom sinonieme het besondere voordele vir die gebruiker wat onmiddellik 'n duideliker beeld van die term wil kry.

— By baie terme is verbandswoorde of termtoeligting gevoeg. Verbandswoorde verskyn kursief gedruk tussen hakies aan die Engelse asook aan die Afrikaanse kant. Dit is gedoen om die gebruik en betekenis hier verder toe te lig — nog 'n voorbeeld van gebruikersvriendelikheid!

Wat die bylaes in die EBW betref, het die samestellers veral aan drie sake aandag gegee.

— Afkortings van uitgesoekte ekonomiese en bedryfsterme tans in gebruik in die sakelewe van Suid-Afrika word in Bylae A net in Engels gegee. Soortgelyke afkortings in Afrikaans ontbreek egter en sou besonder nuttig gewees het vir die Afrikaanssprekende woordeboekgebruiker.

— 'n Lys van die belangrikste Afrikaanse en Engelse sinoniemgroepe verskyn in Bylae B. Daar word volstaan met die opmerkings wat vroeër in hierdie resensie oor sinonieme en wisselvorme gemaak is.

— Bylae C wat handel oor die verantwoording vir die gebruik van sekere Afrikaanse en Engelse terme in die teks verdien spesiale vermelding. Dit is so dat heelwat debat deur belanghebbendes gevoer word oor die korrekte gebruik van sekere Afrikaanse en Engelse terme. Slegs enkele voorbeelde word genoem:

   – *advertering* versus *promotion* versus *publicity*;
   – *afskrywing* versus *waardevermindering* versus *depreciation*;
   – *bedryf* en sy verskillende woordverbindings;

– *belegging* versus *investering*; en

– *doelmatig* versus *doeltreffend* versus *effektief*.

Die samestellers slaag doeltreffend (of is dit nou *doelmatig* of selfs *effektief*) daarin om bogenoemde en nog talle ander terme in perspektief te plaas en so die "regte" gebruik aan te dui.

Dit is egter opvallend dat die terme *onderneming/firm/enterprise/organiza-tion/organisasie/business/undertaking/company* wat dikwels as sinonieme ge-bruik word, nie ook in Bylae C behandel word nie. Die gebruik van die term *organisasie* in Afrikaans skep nogal heelwat verwarring in hierdie verband, veral as onthou word dat organisasie een van die elemente van die bestuurs-taak is.

Voorgaande is slegs 'n baie hoë voëlvlug oor enkele aspekte van die for-midabele *Ekonomiese en Bedryfswoordeboek*. Hierdie woordeboek kan slegs ten volle waardeer en na waarde geskat word as dit deur die onderskeie belang-hebbendes ter hand geneem en bestudeer word.

Daar is werklik nou nie meer 'n verskoning vir veral nagraadse studente en skrywers van vakliteratuur in die Ekonomiese en Bestuurswetenskappe om nie hulle verhandelings, proefskrifte, handboeke, artikels en ander wetenskap-like geskrifte ook in Afrikaans te laat verskyn nie.

Trouens, die EBW behoort ook 'n einde te bring aan die dikwels slordige, anglisistiese, verkeerde en verwarrende woord- en termgebruike deur veral joernaliste en kommentators oor sake- en ekonomiese aangeleenthede in Afri-kaans by die radio, televisie, koerante en ander kommunikasiemedia. As dit nie gaan gebeur nie, is die samestelling van hierdie monumentale EBW heelwaar-skynlik tevergeefs!

Sieg Marx
Voormalige Dekaan van die Fakulteit Ekonomiese en Bestuurswetenskappe
en Viserektor
Universiteit van Pretoria
Pretoria
Republiek van Suid-Afrika

Gabriele Stein. *Developing Your English Vocabulary: A Systematic New Approach.* 2002, VIII + 272 pp. Stauffenburg Linguistik 26. ISBN 3-86057-727-1. Tübingen: Stauffenburg Verlag. Price: €33.

Gabriele Stein is professor of English linguistics at the University of Heidelberg in Germany and has published widely on lexicography and lexicology. The objective of this book is twofold: to compile a lexical core and to maximise the skills of language students by developing ways of expanding this core. It is intended to function as a teaching aid for teachers of English as well as a self-study book for learners of English as a second language. Lexical knowledge is a crucial part of language acquisition and depends on different external factors such as the age and profession of the learner, his/her goals, expectations and needs in learning a language. Beck *et al.* (2002) have demonstrated the small extent of the emphasis on the acquisition vocabulary in school curricula. After a brief introduction, the author looks in chapter 2 at major books which in the 20th century worked on a controlled vocabulary for foreign-language learners in Europe, Asia and America. This section provides the background for the elaboration of a core lexicon and shows how extensive studies have gradually developed in the area of second-language acquisition. In chapter 3, the author elaborates a *Common Core Vocabulary* and in chapter 4 as a continuation of chapter 3, she offers twelve major ways of expanding this core-word list and differentiating lexical items to create a learning framework in the mind of the learner. For foreign-language teachers, the introduction of a relational system of synonyms, antonyms and elementary grammatical structures will gradually build up a structural network for nouns, verbs, adjectives and adverbs that could then be put into practice by the user. The project of a core lexicon (further expanded in chapter 5) targets a more advanced and specialised level of study and acquisition, providing the learner with an *Advanced Common Core Vocabulary*.

The development of core vocabularies in the 1920s and 1930s have been paramount for modern lexicography and the understanding of what was required in English dictionaries for foreign users (Cowie 1998, 1999). They also helped in creating new, user-friendly ways of presenting the information. In the light of previous studies, Stein establishes a system of construction patterns that would prepare lexicographers for the compilation of successful monolingual and bilingual dictionaries, dictionaries of synonyms and thesauri, making them aware of the core concepts and utility words of the language required at various stages of the learning process. The success of such works would depend on the use of both spoken and written data.

Vocabulary expansion develops by associating strings of words. This can be realised, as Stein shows, by a variety of word-building techniques: exploring thematic associations, lexical fields, word-formation patterns, etymology, stylistic variation and phonetic relations, all build word awareness. Such exercises on enriching one's vocabulary increase the learner's understanding of these

items, help him/her in memorising their meaning and in finally putting them into practice.

In the wake of Stein's study, various scholars have recently compiled academic core English vocabularies that serve as the basis for multilingual dictionaries. In South Africa, for instance, where English is the main language for tertiary education, Read *et al.*'s dictionary covers English, Afrikaans, Zulu and Sepedi. Such dictionaries with English as the main language 'will participate in much-needed corpus development of the African languages' (Read *et al.* 1998).

Despite the large amount of research conducted in second-language vocabulary acquisition (Templeton 1989, Schmitt *et al.* 1997, Coady *et al.* 2002), this is a very useful book for enhancing vocabulary development and expanding writing and thinking skills. Although the book is quite accessible to students lacking a background in linguistics, one might regret the general layout which makes it a bit dry. However, this is a valuable resource that should help lexicographers in compiling the most appropriate learner's dictionaries.

References

Beck, I.L., M.G. McKeown and L. Kucan. 2002. *Bringing Words to Life: Robust Vocabulary Instruction*. New York: Guilford.

Coady, J. and T. Huckin (Eds.). 2002. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. Cambridge: Cambridge University Press.

Cowie, A.P. 1998. *Phraseology: Theory, Analysis, Applications*. Oxford: The Clarendon Press.

Cowie, A.P. 1999. *English Dictionaries for Foreign Learners*. Oxford: The Clarendon Press.

Read, J. and M. Ambrose, M. 1998. Towards a Multilingual Dictionary of Academic Words. *Lexikos* 9: 172-187.

Schmitt, N. and M. McCarthy. 1997. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

Templeton, S. 1989. Tacit and Explicit Knowledge of Derivational Morphology: Foundations for a Unified Approach to Spelling and Vocabulary Development in the Intermediate Grades and Beyond. *Reading Psychology* 10: 233-253.

Michaël Abecassis
Department of Modern Languages
University of Oxford
Oxford, United Kingdom
(michael.abecassis@modern-languages.oxford.ac.uk)

Publikasieaankondigings /
Publication Announcements

W.F. Botha (Hoofredakteur). *Woordeboek van die Afrikaanse Taal, Deel XII.* 2005, xxiii + 760 pp. ISBN 0 9584401 8 2 (plastiekband), 0 9584401 9 0 (leerband). Stellenbosch: Buro van die WAT. Prys: R290 (plastiekband), R580 (leerband).

Albert Busch. *Diskurslexikologie und Sprachgeschichte der Computertechnologie.* 2004, XIV + 478 pp. Reihe Germanistische Linguistik 252. ISBN 3-484-31252-1 (Pb.). Tübingen: Max Niemeyer. Price: €138.

Sin-wai Chan. *Translation and Bilingual Dictionaries.* 2004, iv + 190 pp. Lexicographica. Series Maior 119. ISBN 3-484-39119-7 (Pb.). Tübingen: Max Niemeyer. Price: €56.

Julie Coleman. *A History of Cant and Slang Dictionaries. Volume II: 1785–1858.* 2004, xiv + 338 pp. ISBN 0 19 925470 2 (Hb.). Oxford: Oxford University Press. Price: £45.

Julie Coleman and Anne McDermott (Editors). *Historical Dictionaries and Historical Dictionary Research: Papers from the International Conference on Historical Lexicography and Lexicology at the University of Leicester, 2002.* 2004, vii + 224 pp. Lexicographica. Series Maior 123. ISBN 3-484-39123-5 (Pb.). Tübingen: Max Niemeyer. Price: €68.

A.P. Cowie. *English Dictionaries for Foreign Learners: A History.* 1999, xiii + 232 pp. ISBN 0-19-925084-7 (Pb.), 0-19-823506-2 (Hb.). Oxford: Oxford University Press. Price: £16.99 (Pb.), £65 (Hb.). (Review in this issue.)

Madaleine du Plessis (Hoofredakteur/Chief Editor). *Pharos 1, Afrikaans–Engels/ English–Afrikaans Woordeboek/Dictionary.* Eerste uitgawe/First edition. 2005, xx + 1552 pp. ISBN 1868900444. Kaapstad/Cape Town: Pharos Woordeboeke/Pharos Dictionaries. Prys: R450. (Resensie in hierdie uitgawe.)

Henrik Gottlieb, Jens Erik Mogensen and Arne Zettersten. *Symposium on Lexicography XI: Proceedings of the Eleventh International Symposium on Lexicography May 2–4, 2002 at the University of Copenhagen.* 2005, XXI + 555 pp. Lexicographica. Series Maior 115. ISBN 3-484-39115-4 (Pb.). Tübingen: Max Niemeyer. Price €156.

Geoffrey Leech, Paul Rayson and Andrew Wilson. *Word Frequencies in Written and Spoken English*. 2001, xv + 304 pp. ISBN 0 582 32007 0. Harlow: Pearson Education. Price: £45.99. (Review in this issue.)

Ingrid Lemberg, Bernhard Schröder and Angelika Storrer (Editors). *Chancen und Perspektiven computergestützter Lexikographie.* Lexicographica. Series Maior 107. 2001, 270 pp. ISBN 3-484-39107-3. Tübingen: Max Niemeyer. Price: €72. (Review in this issue.)

Sentrum vir Tekstegnologie. *Afrikaanse Speltoetser 2.1 en Woordafbreker vir Microsoft Office®*. 2004. Potchefstroom: Noordwes-Universiteit. (Resensie in hierdie uitgawe.)

W.F.J. Steenkamp en/and H.J.J. Reynders (Hoofredakteurs/Chief Editors). *Ekonomiese en Bedryfswoordeboek: Engels–Afrikaans/Dictionary of Economics and Business: English–Afrikaans.* Tweede, hersiene en vermeerderde uitgawe/Second, revised and enlarged edition. 2004, xii + 1199 pp. ISBN 1868900312. Kaapstad/Cape Town: Pharos Woordeboeke/Dictionaries. Prys/Price: R400. (Resensie in hierdie uitgawe.)

Gabriele Stein. *Developing Your English Vocabulary: A Systematic New Approach*. 2002, VIII + 272 pp. Stauffenburg Linguistik 26. ISBN 3-86057-727-1. Tübingen: Stauffenburg Verlag. Price: €33. (Review in this issue.)

VOORSKRIFTE AAN SKRYWERS

(Tree asseblief met die Buro van die WAT in verbinding (wat@wat.sun.ac.za) vir 'n uitvoeriger weergawe van hierdie instruksies of besoek ons webblad: <http://www.sun.ac.za/wat>)

A. REDAKSIONELE BELEID

1. Aard en inhoud van artikels
Artikels kan handel oor die suiwer leksikografie of oor implikasies wat aanverwante terreine, bv. linguistiek, 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 evaluerend 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 Beheerraad 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 evaluerende 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@wat.sun.ac.za)
or refer to our website: <http://www.sun.ac.za/wat>)

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 Control 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 typewritten 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@wat.sun.ac.za) für eine ausführlichere Wiedergabe dieser Hinweise oder besuchen Sie usere Webseite: <http://www.sun.ac.za/wat>)

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 Aufsichtsrat 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 Freiexemplare

Autoren erhalten fünf Sonderdrucke ihrer veröffentlichten Artikel oder Rezensionsartikel gratis sowie ein Freiexemplar 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 Freiexemplar 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.sun.ac.za/wat>)

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 Board of Control du *Woordeboek 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*.