

AFRILEX^{-reeks} series 10:2000

LEXIKOS 10

BURO VAN DIE WAT

AFRILEX

Lexikos 10

Redakteur

Editor

J.C.M.D. du Plessis

Resensieredakteur

Review Editor

T. Harteveld



African Association for Lexicography

AFRILEX-REEKS 10:2000

AFRILEX SERIES 10:2000



BURO VAN DIE WAT

STELLENBOSCH

Uitgewer Publisher

**BURO VAN DIE WAT
Posbus 245
7599 STELLENBOSCH**

**Kopiereg © 2000 deur die uitgewer
Alle regte streng voorbehou
Eerste uitgawe 2000**

**Tipografie en uitleg
deur Tanja Harteveld
Bandontwerp deur Piet Grobler**

Geset in 10 op 12 pt Palatino

**Gedruk en gebind deur Onyx Press
Stewartstraat 61 Goodwood**

ISBN 0 9584120 5 7

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

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

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

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

Adviesraad / Advisory Board

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

Redaksiekomitee / Editorial Committee

- Prof. A. Carstens (RSA)
Prof. W.A.M. Carstens (RSA)
Mnr. E. Chabata (Zimbabwe)
Dr. A.R. Chuwa (Tanzanië/Tanzania)
Prof. C.J. Conradie (RSA)
Prof. L.G. de Stadler (RSA)
Dr. M.M.R. Dube (RSA)
Prof. J.A. du Plessis (RSA)
Dr. A.E. Feinauer (RSA)
Prof. E.F. Kotzé (RSA)
Me. K.P. Kavanagh (RSA)
Dr. M. Lisimba (Gaboen/Gabon)
Mnr. K.J. Mashamaite (RSA)
Dr. J.S. Mdee (Tanzanië/Tanzania)
Mnr. T.X. Mfafa (RSA)
Mnr. M.H. Mpungose (RSA)
Prof. F.A. Ponelis (RSA)
Prof. D.J. Prinsloo (RSA)
Dr. R. Sitaram (RSA)
Prof. P.H. Swanepoel (RSA)
Dr. J. Tsonope (Botswana)

Inhoud / Contents

Voorwoord	ix
Foreword	xi
<i>J.C.M.D. du Plessis</i>	
'n Woord van AFRILEX	xiii
A Few Words from AFRILEX	xiv
<i>Danie Prinsloo</i>	
Redaksionele doelstellings	xv
Editorial Objectives	xvi
Redaktionelle Ziele	xvii
Navorsingsartikels / Research Articles	
The Concept of "Simultaneous Feedback": Towards a New Methodology for Compiling Dictionaries	1
<i>Gilles-Maurice de Schryver and D.J. Prinsloo</i>	
Adressierung in der ein- und zweisprachigen Lexikographie. Eine einführende Übersicht über die Forschungs- und Problemlage	32
<i>Herbert Ernst Wiegand</i>	
Beskouende artikels / Contemplative Articles	
The Shona Corpus and the Problem of Tagging	75
<i>Emmanuel Chabata</i>	
Should LSP Dictionaries also Include Professional Jargon and Slang?	86
<i>Rosemarie Gläser</i>	
Words in a Cultural Context	99
<i>Kathryn Kavanagh</i>	

An Integrated Semasiological and Onomasiological Presentation of Semantic Information in General Monolingual Dictionaries as Proposed in H.E. Wiegand's <i>Semantics in Lexicography</i> <i>Phillip Louw</i>	119
The Use of Examples in Polyfunctional Dictionaries <i>D.J. Prinsloo and R.H. Gouws</i>	138
Standards retrodigitaler Wörterbücher <i>Michael Schlaefer</i>	157
Using Frames in Special-Field Lexicography: An Ethnomusical Case Study <i>Maria Smit</i>	173
Theoretical Challenges to LSP Lexicography <i>Sven Tarp</i>	189
Leksikonotas / Lexiconotes	
Challenges Encountered in the Compilation of an Advanced Shona Dictionary <i>Moreblessings Busi Chitauro-Mawema</i>	209
Grammatical Terminology for the Teaching of African Languages at Tertiary Level <i>T.C. Msimang</i>	225
Leksikovaria / Lexicovaria	
Terminology Management at the National Language Service <i>Mariëtta Alberts</i>	234
Lexicography in Gabon: A Survey <i>B. Nyangone Assam and P.A. Mavoungou</i>	252
Resensieartikels / Review Articles	
Nuwe dinge in <i>Nuwe Woorde</i> — 'n Oorsig oor buitemikrostrukturele leksikografiese vernuwings <i>Herman Beyer</i>	275

Guide Structures in CD-ROM Dictionaries, with Specific Reference to the EWED and the ELHAT <i>Phillip Louw</i>	291
Juta's Concise Dictionary of Accounting Terms with their Afrikaans Equivalents <i>N. Stegmann</i>	306
Resensies / Reviews	
Igor Burkhanov. <i>Lexicography: A Dictionary of Basic Terminology</i> <i>Mohammed H. Heliel</i>	320
Latha Menon (Executive Editor). <i>Microsoft Encarta Encyclopedia Deluxe 2000</i> <i>Phillip Louw</i>	326
T.J. Kriel (original compiler), D.J. Prinsloo and B.P. Sathekge (compilers — revised edition). <i>Popular Northern Sotho Dictionary</i> <i>Kwena J. Mashamaite</i>	329
Wilfrid H.G. Haacke and Eliphaz Eiseb. <i>Khoekhoegowab-English / English-Khoekhoegowab Glossary / Mîdi Saogub</i> <i>J.W. Snyman</i>	331
Publikasieaankondigings / Publication Announcements	334
Voorskrifte aan Skrywers	336
Instructions to Authors	337
Richtlinies für Autoren	338

Voorwoord

In hierdie uitgawe van *Lexikos* publiseer ons verskeie lesings wat by AFRILEX 2000 aangebied is. Die tema was Woordeboeke van Tale vir Spesiale Doeleindes. Ongelukkig was dit vanweë 'n gebrek aan ruimte nie moontlik om meer van hierdie lesings in te sluit nie. Sommige daarvan sal volgende jaar gepubliseer moet word.

Ons voeg ook 'n nuwe kategorie in waaronder artikels geklassifiseer kan word: *Lexikonotas*. Hierdie kategorie is bedoel vir artikels wat voorstelle maak, probleme stel, vrae vra, inligting verskaf, kommentaar lewer, en oplossings voorstel, maar waarin die menings nie so uitvoerig beredeneer en ontwikkel word soos in die kategorie *Beskouende artikels* nie. Daar word selfs gehoop dat lesers sal reageer op punte wat in hierdie artikels aangeraak word.

Daar was veranderinge in die Adviesraad en die Redaksiekomitee. Die funksie van die Adviesraad is om toesig te hou oor die gehalte van *Lexikos* in die algemeen. In die plek van prof. V. February wat uittree, het dr. H. Chimhundu (voorheen lid van die Redaksiekomitee), dr. A.F.M.J. Moerdijk en dr. L.S. Vikør die uitnodiging aanvaar om vanaf hierdie uitgawe op die Adviesraad te dien.

Die taak van lede van die Redaksiekomitee is om te help met die beoordeling van artikels vir insluiting in *Lexikos*. Verskeie redaksielede wat vir 'n termyn van drie of meer jare gedien het, het uitgetree. Hulle is prof. W.R.G. Branford, mev. M. Matthews, prof. B.M. Mini, prof. E.H. Raidt, mev. P.M. Silva, prof. G.J. van Jaarsveld en prof. E.B. van Wyk. Vanaf hierdie uitgawe is die nuwe lede van die Redaksiekomitee dr. M.M.R. Dube, me. K.P. Kavanagh, mnr. K.J. Mashamaite en mnr. T.X. Mfafa.

Ons wil die nuwe lede van die Adviesraad en Redaksiekomitee verwelkom. Ons hoop hul deelname aan die produksie van *Lexikos* sal 'n uitdagende en vrugbare ervaring wees. Al die uittredende lede moet bedank word vir hul bydrae. Sonder hul raad sou die standaard van *Lexikos* nie volgehou kon word nie en sonder hul oordeel sou die gehalte van die artikels nie gewaarborg kon word nie.

Dankbetuiging

Vanjaar moes die keurders baie artikels op baie kort kennisgewing beoordeel. Ten spyte daarvan, kon hulle nog die tyd vind om dikwels uitvoerig op die gehalte en moontlike verbetering van die artikels kommentaar te lewer. Die

Voorwoord

aansienlike moeite wat hulle met die beoordeling van die artikels gedoen het, word hoog gewaardeer.

Me. Tanja Harteveld het in 'n drieledige hoedanigheid opgetree: behalwe dat sy die resensieredakteur was, het sy ook al die administrasie en korrespondensie behartig en al die setwerk van die manuskrip gedoen. Mnr. Etienne Botha het al die Engelse artikels gelees en waardevolle opmerkings gemaak vir hul verbetering. Sonder me. Harteveld en mnr. Botha se aansienlike ervaring en geduldige hulp sou hierdie uitgawe van *Lexikos* nie gepubliseer kon word nie. Ek wil hulle altwee bedank.

J.C.M.D. du Plessis

Buro van die Woordeboek van die Afrikaanse Taal

Foreword

In this issue of *Lexikos*, we publish several papers presented at AFRILEX 2000. The theme was Dictionaries of Languages for Special Purposes. Unfortunately it was not possible to include more of these papers due to a lack of space. Some have to be published next year.

We also introduce a new category under which articles can be classified: *Lexiconotes*. This category is meant for articles making suggestions, stating problems, asking questions, giving information, offering comments and suggesting solutions, but in which the ideas are not so extensively argued and developed as those in the category *Contemplative Articles*. It is even hoped that readers will react to issues raised in these articles.

There have been changes to the Advisory Board and the Editorial Committee. The function of the Advisory Board is to supervise the standard of *Lexikos* in general. In the place of Prof. V. February who retires, Dr H. Chimhundu (previously member of the Editorial Committee), Dr A.F.M.J. Moerdijk and Dr L.S. Vikør have accepted the invitation to serve on the Advisory Board from this issue onwards.

The task of members of the Editorial Committee is to help judge articles for inclusion in *Lexikos*. Several committee members who have served a term of three or more years, are now retiring. They are Prof. W.R.G. Branford, Mrs M. Matthews, Prof. B.M. Mini, Prof. E.H. Raidt, Mrs P.M. Silva, Prof. G.J. van Jaarsveld and Prof. E.B. van Wyk. As from this issue, the new Editorial Committee members are Dr M.M.R. Dube, Ms K.P. Kavanagh, Mr K.J. Mashamaite and Mr T.X. Mfana.

We wish to welcome the new Advisory Board and Editorial Committee members. We hope their participation in the production of *Lexikos* will be a challenging and fruitful experience. All the retiring members must be thanked for their contribution. Without their advice the standard of *Lexikos* could not have been maintained and without their judgement the quality of the articles could not have been guaranteed.

Acknowledgements

This year the judges had to evaluate many articles at very short notice. In spite of this, they could still find time to often comment extensively on the quality

Foreword

and possible improvement of the papers. The considerable effort they put into the judging of the papers is much appreciated.

Ms Tanja Hartevelde acted in a triple capacity. Apart from being the review editor, she handled all the administration and correspondence and did all the typesetting of the manuscript. Mr Etienne Botha read all the English articles and made valuable suggestions for their emendation. Without Ms Hartevelde's and Mr Botha's considerable experience and patient assistance, this issue of *Lexikos* would not have been published. I wish to thank them both.

J.C.M.D. du Plessis

Bureau of the Woordeboek van die Afrikaanse Taal

'n Woord van AFRILEX

Die leksikografiese proses in Suid-Afrika het momentum in 1999-2000 verkry met 'n reeks aktiwiteite wat toegespits is op die stigting van woordeboekeenhede vir die amptelike tale in Suid-Afrika. Daar lê egter 'n lang pad voor alvorens die samestelling van woordeboeke vir die elf amptelike tale net so gesofistikeerd sal wees as byvoorbeeld moderne woordeboeke vir Afrikaans, Engels, Duits, Frans, ens. Heelwat struikelblokke moet nog oorkom word veral ten opsigte van opleiding, rekenaarprogramme, gebruikersbehoefte en burokrasie. Die African Association for Lexicography (Afrilex) loof hierdie inisiatiewe en neem aktief deel aan die proses ter ondersteuning van die inisiatiewe van die Pan Suid-Afrikaanse Taalraad en sal voortgaan met sodanige aktiewe betrokkenheid ten einde hierdie doelstellings te bereik.

Afrilex het vanjaar in Stellenbosch sy Vyfde Internasionale Kongres gehou en kyk dan ook met trots en dankbaarheid terug na vyf jaar waarin die vereniging gegroei en sy regmatige plek ingeneem het tussen die ander leksikografieverenigings in die wêreld. Dit is grootliks te danke aan harde werk deur lede van die bestuur asook entoesiastiese ondersteuning deur lede van die vereniging. Die Sesde Internasionale Konferensie is geskeduleer vir 2-4 Julie 2001, en word georganiseer deur die Sepedi Woordeboekeenhede aan die Universiteit van die Noorde in Pietersburg.

Afrilex is trots daarop om 'n tydskrif soos *Lexikos* as mondstuk te hê wat plaaslik en internasionaal erken word en moedig graag ons lede aan om artikels voor te lê vir publikasie. Afrilex kon tot dusver ook daarin slaag om jaarliks 'n eksemplaar van *Lexikos* gratis aan lede te versend en die bereidwilligheid van die Buro van die WAT om die kostes so laag as moontlik te hou word opreg waardeer.

Daar word gepoog om lede deur middel van omsendbriewe en nuusbriewe voortdurend op hoogte van die vereniging se aktiwiteite te hou asook deur middel van die tuisblad:

www.up.ac.za/academic/libarts/afrilang/homelex.html

Hierdie uitgawe van *Lexikos* bied weer eens puik leesstof aan ons lede asook aan andere wat belang het by woordeboekmaak en woordeboekgebruik.

Danie Prinsloo
Voorsitter: AFRILEX

A Few Words from AFRILEX

The lexicographic process in South Africa moved into high gear in the years 1999-2000 with a series of activities aimed at the establishment of lexicographic units for the official South African languages. There is, however, a long way to go before dictionary compilation for all of the eleven official languages will reach the level of sophistication of, say, modern dictionaries for Afrikaans, English, German, French etc. Many hurdles in respect of training, computer programs, users' needs and bureaucracy have to be overcome. The African Association for Lexicography (Afrilex) applauds these initiatives and actively participates in this process in support of the initiatives undertaken by the Pan South African Language Board and will continue to play an active role in reaching of these objectives.

Afrilex held its Fifth International Conference in Stellenbosch this year and looks back with pride and gratitude to five years during which the association has grown and taken its rightful place among the other lexicography associations of the world. This is a result of the hard work done by members of the management committee, as well as the enthusiastic support of the members of the association. The Sixth International Conference will be held from 2-4 July 2001, organised by the Sepedi Dictionary Unit based at the University of the North in Pietersburg.

Afrilex furthermore takes pride in having as its mouthpiece *Lexikos*, a journal which is recognised both locally and internationally. We would like to encourage our members to submit articles for publication in this journal. To date Afrilex has managed to send members a copy of *Lexikos* free of charge annually and the willingness of the Bureau of the WAT to keep costs at a minimum level is sincerely appreciated.

The association attempts to keep members informed about activities by means of circulars and newsletters as well as through our web site:

www.up.ac.za/academic/libarts/afrilang/homelex.html

This issue of *Lexikos*, once again, offers excellent reading matter to our members and those interested in dictionary compilation and dictionary use.

Danie Prinsloo
Chairperson: AFRILEX

Redaksionele doelstellings

Lexikos is 'n tydskrif vir die leksikografiese vakspesialis en word in die AFRILEX-reeks uitgegee. "AFRILEX" is 'n akroniem vir "leksikografie in en vir Afrika". Van die sesde uitgawe af dien *Lexikos* as die amptelike mondstuk van die *African Association for Lexicography* (AFRILEX), onder meer omdat die Buro van die WAT juis die uitgesproke doel met die 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 leksikografiese gesprek te skep, en in die besonder die leksikografie in Afrika met sy ryk taleverskeidenheid te dien;
- (2) om die gesprek tussen leksikograwe onderling en tussen leksikograwe en taalkundiges te stimuleer;
- (3) om kontak met plaaslike en buitelandse leksikografiese projekte te bewerkstellig en te bevorder;
- (4) om die interdisiplinêre aard van die leksikografie, wat ook terreine soos die taalkunde, algemene taalwetenskap, leksikologie, rekenaarwetenskap, bestuurskunde, e.d. betrek, onder die algemene aandag te bring;
- (5) om beter samewerking op alle terreine van die leksikografie moontlik te maak en te koördineer, en
- (6) om die doelstellings van die *African Association for Lexicography* (AFRILEX) te bevorder.

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

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

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

Editorial Objectives

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

The objectives of the AFRILEX Series are:

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

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

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

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

Redaktionelle Ziele

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

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

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

Gemäß den Zielsetzungen der AFRILEX-Serie werden:

- (1) Beiträge zum lexikographischen Gespräch 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.

The Concept of "Simultaneous Feedback": Towards a New Methodology for Compiling Dictionaries*

Gilles-Maurice de Schryver, *Department of African Languages and Cultures, University of Ghent, Belgium, and*

D.J. Prinsloo, *Department of African Languages, University of Pretoria, South Africa*

Abstract: Good lexicographers are constantly striving to enhance the quality of their dictionaries. Since dictionaries are ultimately judged by their target users, there is an urgency to provide for the target users' needs. In order to determine such needs more accurately, it has become common practice to submit users of a dictionary to a series of tests to monitor their success in information retrieval. In most cases such feedback unfortunately comes too late so that it can at best be considered for implementation in the next or revised edition of the dictionary. In this article it is argued that feedback from the target users should be obtained *while the compilation of the dictionary is still in progress*, a process referred to as "simultaneous feedback". This concept, which offers a new methodology for compiling dictionaries, overcomes the major problem of creating and publishing entire dictionaries before feedback from target users can be obtained. By this new methodology, the release of several small-scale parallel dictionaries triggers feedback that is immediately channelled to the compilation process of a main dictionary. As such, the target users constantly guide the compilers during the entire compilation process. After a theoretical presentation of the new concept, the feasibility of simultaneous feedback is illustrated with reference to the creation of a bilingual Cilubà-Dutch learner's dictionary. It is shown how this main project has been successfully complemented by three parallel projects.

Keywords: SIMULTANEOUS FEEDBACK, NEW METHODOLOGY, MAIN DICTIONARY, PARALLEL DICTIONARIES, TARGET USERS' DESIRES, QUESTIONNAIRES, ELECTRONIC CORPORA, WORD-FREQUENCY STUDIES, CONCORDANCES, AFRICAN LANGUAGES, CILUBÀ

Opsomming: Die konsep van "gelyktydige terugvoering": Onderweg na 'n nuwe metodologie vir die samestelling van woordeboeke. Goeie leksikograwe streef voortdurend daarna om die gehalte van hul woordeboeke te verbeter. Aangesien woorde-

* An earlier version of this article was presented by G.-M. de Schryver as keynote lecture at the Fourth International Conference of the African Association for Lexicography, held at the University of Pretoria, 5-6 July 1999. This lecture was an extract from G.-M. de Schryver's M.A. dissertation (De Schryver 1999a), written under the guidance of D.J. Prinsloo and R.H. Gouws. For this version D.J. Prinsloo joined in the writing.

boeke uiteindelik deur hul teikengebruikers beoordeel word, is dit uiters noodsaaklik dat aan gebruikersbehoefes voldoen moet word. Ten einde sodanige behoeftes noukeuriger te kan bepaal, is dit algemeen gebruiklik dat woordeboekgebruikers aan 'n reeks toetse onderwerp word ten einde hulle sukses ten opsigte van inligtingsinwinning te kan bepaal. Ongelukkig kom sodanige terugvoer in die meeste gevalle te laat deurdadig dat slegs in die volgende of hersiene uitgawe van die woordeboek in berekening gebring kan word. In hierdie artikel word betoog dat terugvoering vanaf die teikengebruikers verkry moet word *terwyl die samestelling van die woordeboek nog aan die gang is*, 'n proses waarna verwys word as "gelyktydige terugvoering". Hierdie konsep wat 'n nuwe metodologie vir woordeboekmaak bied, omseil die kernprobleem van die voltooiing en publikasie van volledige woordeboeke alvorens terugvoer vanaf die teikengebruikers verkry kan word. Deur hierdie nuwe metodologie lei die vrystelling van verskillende kleiner parallelle woordeboeke tot terugvoer wat onmiddellik in die samestellingsproses van die hoofwoordeboek gekanaliseer word. Sodoende ontvang die samestellers deurlopend riglyne vanaf die teikengebruikers gedurende die volle duur van die samestellingsproses. Na 'n teoretiese uiteensetting van hierdie nuwe konsep, word die uitvoerbaarheid van gelyktydige terugvoering geïllustreer met verwysing na die voltooiing van 'n Cilubà-Nederlandse aanleerderswoordeboek. Daar word aangetoon hoe hierdie hoofprojek deur drie parallelle projekte aangevul word.

Sleutelwoorde: GELYKTYDIGE TERUGVOERING, NUWE METODOLOGIE, HOOFWOORDEBOEK, PARALLELE WOORDEBOEKE, TEIKENGEBRUIKERS SE WENSE, VRAELYSTE, ELEKTRONIESE KORPUSSE, WOORDFREKWENSIESTUDIES, KONKORDANSIES, AFRIKATALE, CILUBÀ

1. Introduction

In modern lexicography the so-called user-perspective has emerged as an all-important criterion in the selection and lexicographical treatment of lexical items.

Dictionaries first came into being in response to very practical needs. (Osselton 1983: 13)

It is the function of a popular dictionary to answer the questions that the user of the dictionary asks, and dictionaries on the commercial market will be successful in proportion to the extent to which they answer these questions of the buyer. This is the basis on which the editor must determine the type of information to include. (Barnhart 1967² [1962]: 161)

The user-perspective, so prevalent in modern-day metalexicography, compels lexicographers to compile their dictionaries according to the needs and research skills of well-defined target user groups. The dominant role of the user has had a definite effect on the compilation of dictionaries as well as on the evaluation of their quality. Good dictionaries do not only display a linguistically sound treatment of a specific selec-

tion of lexical items. They are also products that can be used as linguistic instruments by their respective target user groups. The better they can be used, the better dictionaries they are. (Gouws and Prinsloo 1998: 18)

Since the lexicographer is constantly striving to enhance the quality of a dictionary, there is an urgency to take the target users' needs, their expectations and their reference skills into account (cf. also Atkins 1998, Wiegand 1998, Van der Merwe 1999). In addition to these aspects, target users' *desires* are introduced in the present article. These desires denote a set of requests which go beyond the linguistic information required to satisfy the users' *needs*, beyond the users' *expectations* based on their perception of how dictionaries ought to be, and beyond the users' *reference skills* developed in the course of consulting so-called "standard dictionaries". Indeed, by referring to target users' desires we want to go to the core of *what* the target users would like to find in a dictionary, and *how* they would like to find it — it is the users who decide on the types of data to include in or omit from a dictionary, and it is the users who decide on the way this data is presented. For example, users might indicate that they desire an indication of frequency of use but not etymological information, or users might indicate that they desire nouns to be lemmatised in one format but adjectives in another. Information on such desires must be collected by both informal and formal means. From the moment dictionary compilers take target users' desires into account, they acknowledge that today's target users are the *privileged critics* of a dictionary.

It is however not easy for the compiler to determine exactly what these needs, expectations, reference skills and desires of the target users are. Many compilers simply rely on their intuition as far as these are concerned. In order to determine these needs, expectations, reference skills and desires more accurately, it has become common practice to submit users of a dictionary to a series of tests to monitor their success in information retrieval. Atkins (1998) is an excellent example of current efforts to research dictionary use. However, such research is carried out at a stage when the dictionaries are already completed and published. Feedback from target users can only be implemented in forthcoming editions of these dictionaries. Formulated differently, it means that the feedback obtained in this way comes too late. Tono (1992: 232) is correct when he remarks:

Probably the most widely used research method with regard to the dictionary user is the survey ... All variables are studied *ex post facto* that is, as they exist in the situation. No experimental variables are manipulated. Furthermore, most of the surveys on dictionary-users' study are limited to describing the *status quo*.

In this article it is argued that feedback from the target users should be obtained while the compilation of the dictionary is *still in progress*, a process for

this reason referred to as "simultaneous feedback". This concept which offers a new methodology for compiling dictionaries, overcomes the major problem of having to create and publish entire dictionaries first before one can start to evaluate them.

In a nutshell, this new methodology entails the release of several small-scale parallel dictionaries which trigger feedback that is channelled back to the compilation process of a main dictionary. Within the framework of simultaneous feedback, the variables are thus not studied *ex post facto*, as Tono notes. They are studied while they are being implemented, since the target users guide the compilers during the entire compilation process. As such, it is not a description of the *status quo*.

The concept of simultaneous feedback grew out of a dire necessity to compile a bilingual Cilubà-Dutch learner's dictionary. It quickly became evident that the main problems encountered in the lexicographical treatment of Cilubà are shared by the other African languages.¹ The solutions that will be presented therefore have implications that go beyond the Lubà language. The theoretical concept of simultaneous feedback itself goes far beyond the African languages — it applies to all languages. This is why we claim that it is a new methodology for compiling dictionaries.

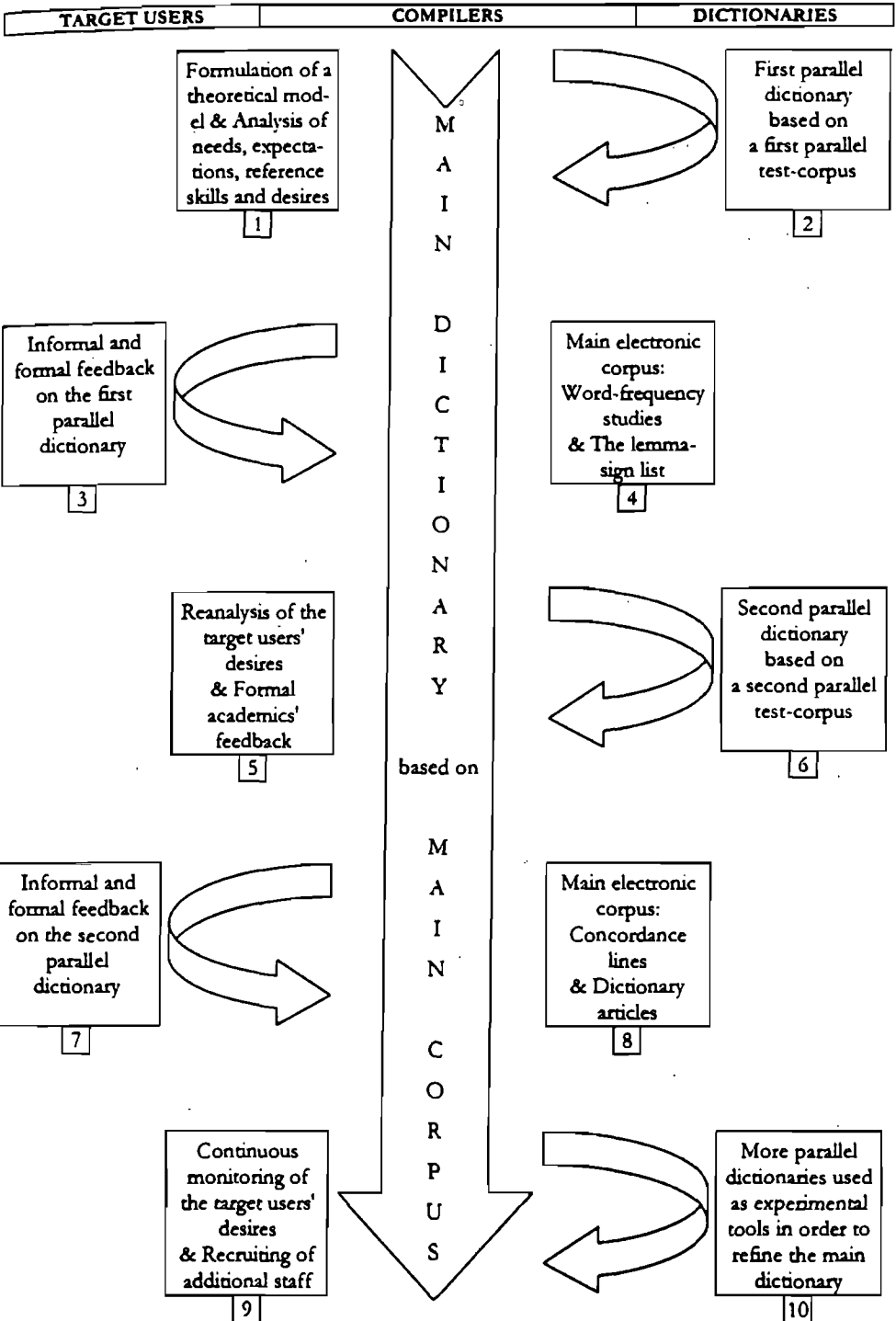
2. The theoretical framework of simultaneous feedback

Compilers write dictionaries to be used by target users. Target users consult the dictionaries and in this process they come to judge those very dictionaries. Dictionaries' privileged critics are therefore the target users who were the original focus of the compilers. We have come full circle, from compilers to target users, to dictionaries and back to compilers. This full circle is not just a linguistic loop, because the loop represents indispensable feedback. However, for most dictionary projects this feedback comes too late. In fact, one would like to have feedback from the privileged critics (the target users) both from the very start and during the entire compilation process of a dictionary — hence, the chimera would be to listen to simultaneous feedback from the target users to the compilers. But is it really a chimera? With the concept of simultaneous feedback, it is *not*. A schematic representation of the theoretical framework of simultaneous feedback is shown in (1).

In (1) one recognises the three primary constituents of any dictionary compilation process, viz. target users, compilers and dictionaries. The framework itself should roughly be read from left to right, and from top to bottom — [1] through to [10]. However, as will be seen, every single component is actually interlinked with every other single component.

The compilers' central task is the compilation of a main dictionary, based on a main electronic corpus — depicted by the central arrow in (1). The first component in (1) is also the first task. That is, the formulation of a theoretically motivated model for the structure and contents of the dictionary to be com-

(1) The theoretical framework of "simultaneous feedback"



piled must be followed by a preliminary analysis of the needs, expectations and reference skills of the potential target users of the dictionary — [1]. It is of paramount importance that from this initial stage onwards, information concerning the target users' desires be gathered through informal and formal consultation with the future target users — [1]. *As such, feedback is simultaneously introduced right from the very start.*

Since the main dictionary is to derive its data from a main corpus, the compilers have to build an electronic corpus for the specific language. As a result, the compilers cannot start the compilation of the main dictionary right away and are moreover confronted with the prospect of an extremely time-consuming undertaking. In order to overcome this deadlock the main dictionary project is *complemented* by a series of small and inexpensive parallel dictionary projects. These parallel projects have the same basic structure, contents and target users as the main project — [2], [6], [10], etc. These parallel projects are to derive their data from small-scale parallel test-corpora and are to be completed in short periods of time. From the release of the first parallel dictionary onwards, informal and formal feedback is received from the parallel projects and channelled back into the time-consuming main project — [3], [7], etc. *From this instant, the compilation of the main dictionary becomes a true "work in progress" with simultaneous feedback from the target users to the compilers.*

The parallel projects are thus used as experimental tools to test a plethora of strategies in order to refine the presentation of the information in the main project under construction. Once a structured main corpus has been built, word-frequency studies can be done to assist the lexicographers in the compilation of the lemma-sign list of the main dictionary — [4]. Subsequently, concordance lines, also derived from the main corpus, supplement and confirm the compilers' intuition in the compilation of the main dictionary articles — [8]. Until completion of the main project the parallel projects continue to elicit feedback — [5], [9], etc. All these instances of simultaneous feedback ultimately enable the compilers to select the most appropriate blend of lexicographical procedures to ensure the most effective retrieval of information by the target users in the main dictionary.

Consequently, within the framework of simultaneous feedback, we cannot agree with the claim that "lexicography is not a terrain in which you can experiment from one day to the next" (Lombard 1994: 211). On the contrary, complying with the target users' desires means doing just that!

It goes without saying that this theoretical framework should be considered the backbone of any potential dictionary project. In the remainder of this article, the concept of simultaneous feedback will be applied to the compilation of a functional Cilubà-Dutch source of reference. For this project, the components [1] through to [10] shown in (1) will be analysed, with the main focus on the first few components.

- [1] **Formulation of a theoretical model, and Analysis of needs, expectations, reference skills and desires**
- (a) **Formulation of a theoretically motivated model for the structure and contents of the main dictionary project**

The ultimate goal of a *Woordenboek Cilubà-Nederlands* (Cilubà-Dutch Dictionary), WCN for short, is "a general and alphabetically ordered, pocket-sized unidirectional decoding learner's paper dictionary, based on frequency of usage and meant for university students". It is important to stress the fact that the promptness with which the precise target and target group were determined should not be considered as mere catch-as-catch-can choices. Indeed, before embarking on a dictionary project one must take the nature and comprehensiveness of "currently available lexica" seriously into account in order not to compile "just another general dictionary" to be used by "unspecified target users". Prior to the project being discussed here, only one general dictionary with Cilubà as the source language was readily available, viz. the *Dictionnaire Tshilubà-Français* (De Clercq and Willems 1960³). In this dictionary, Cilubà, French and (some "keywords" in) Dutch are covered in a one-way configuration. Unfortunately, a speaker of Dutch who is not acquainted with French is unable to use this dictionary satisfactorily. A bilingual Cilubà-Dutch dictionary is thus certainly not a redundant product.

- (b) **Preliminary analysis of the needs, expectations and reference skills of the target users**

From the start we had a very specific target group in mind, viz. university students. Their general needs and expectations are reflected in the project's ultimate target: "a general and alphabetically ordered, pocket-sized unidirectional decoding learner's paper Cilubà-Dutch dictionary". It should be noted that these needs and expectations were only finalised through informal conversations with the future target group.

As far as the target users' reference skills are concerned, suffice it to say that as they are university students, we could presuppose that these skills would be quite well developed.

- (c) **Informal and formal consultations with the target users in order to determine their desires**

The target users' desires were recorded in two ways. Firstly, the actual use of the *Dictionnaire Tshilubà-Français* by university students was monitored in an informal way. The gathering of informal data from actual dictionary use might seem trivial, but it definitely constituted "a battery of learners' desires" — col-

lected, it should be stressed, through natural participant observation. Secondly, information was also gathered in a more formal way via direct questions like "Imagine someone would compile a Cilubà-Dutch Dictionary. Ideally, what would *you* like to find in such a dictionary? And how would *you* like to find it?"

[2] First parallel dictionary based on a first parallel test-corpus

In order to get a dictionary project going within the framework of simultaneous feedback, one quickly needs to determine a subtarget and a subtarget group for a first pocket-sized parallel project which can be compiled in a short period of time and which is based on a first parallel test-corpus which can be built promptly — and all this within a limited budget. We projected that the first parallel project would have to result in a small unidirectional decoding learner's lexicon for Dutch-speaking learners of Cilubà at the Department of African Languages and Cultures of the University of Ghent, Belgium. For that specific subtarget and subtarget group, it was easy to foresee the series of potential Lubà lemma-signs for which the future users would be likely to take to the lexicon, namely those "heard" during the theoretical courses and practical exercises, and those "read" while translating texts from Cilubà into Dutch and studying the different syllabi. It was therefore decided to put together a first parallel test-corpus consisting of all elements encountered during the first two years of the Cilubà courses. Hence, the orthographic, semantic and grammatical information required to construct the lexicon were to be extracted primarily from those courses. The corpus itself was called the "Course Corpus", while the resulting *Lexicon Cilubà-Nederlands* (De Schryver and Kabuta 1997), henceforth LCN, was compiled in just two months and contained roughly 2 500 lemma-signs. From the sample pages shown in Appendix A, it is clear that the aim was to deviate as little as possible from so-called "standard dictionary conventions".

Instead of thoroughly discussing all the aspects of this first parallel dictionary, we will focus on two topics, nouns and verbs. On the lexicographical treatment of nouns in African languages, the dilemma has always been "How does one lemmatise nouns successfully?" (see e.g. Prinsloo and De Schryver 1999). We chose to enter nouns as singulars, followed by the gender — where the two poles of the gender are thought to enable the user to form, when applicable, the plural. Consider the nouns in (2):

- (2) (a) kààfê [12/4] cf var kàfê
 (b) kaakù [1/2a] grootouder; voorouder; ~ mukàjì [ud] grootmoeder; ~ mulùme [ud] grootvader
 (c) kabèji [12/13; dim ◀ dibèji] papiertje; ~ kàà bwanga (dokters)voorschrift

Here the genders 12/4, 1/2a and 12/13 are thought to enable the target users to infer the plural forms.

When it comes to verbs, the treatment of a typical verb in LCN is illustrated in (3).

- (3) -dyà [tww; cf spw3, 5] eten; ~
kuukuta [ud] eten en verpaid z
 ▶ bidyà; cidìilu; cyàkudyà; -dìi-
 ka; -dìikiibwa; -dìila; -dìilsha;
 mudì; Mudiila-mpiku

As can be seen in (3), a verb is lemmatised in the imperative in LCN, thus without its class prefix *ku-* — a rather traditional approach, as bringing all the verbs together under *ku/kw* would only result in an artificial overcrowding of one particular lexicon-section.

From (2) and (3) one could get the impression that articles stand in relative isolation in LCN. This view is incorrect, as a comprehensive network of cross-references can link any reference position within an article with any internal or external reference address. We will briefly look at two novelties, "verb stem node" and "noun node". In LCN one encounters two kinds of large networks, each one with its respective "reference nodes". The first network (when present) is centred around the stem of verbs. Its purpose is to use this form as a node to link all lemma-signs that are connected. However, one is only referring to lemma-signs *within* the lexicon, and nothing is claimed about other possibilities in Cilubà. The "tail slot" was created for this purpose. It is to be found at the end of an article and starts with an arrow pointing to the right followed by one or more lemma-signs connected with this single "verb stem" (the head of the article). For instance, for the verb *kudyà* in (3), one finds nine lemma-signs in the tail. Under their proper alphabetical position, all of these nine evidently contain a cross-reference back to the node, mostly through the use of an arrow pointing to the left. Such cross-references can be seen in (4) where a sample of the lemma-signs from *kudyà*'s tail slot are shown as they can be found under their proper alphabetical position:

- (4) (a) cidìilu [7/8 ◀ app -dyà] 1 eetzaal;
 2 kribbe
 (b) cyàkudyà [cn sub 7/8 ◀ -dyà]
 voedsel; *mukàndà wà byà-*
kudyà menu(kaart)
 (c) -dìika [iww, sta -dyà] eetbaar z;
 gegeten w; afgeknaagd w; ver-
 teren
 (d) mudì [1/2 ◀ -dyà] eter; verslin-
 der

Although one can construct tail slots for several part-of-speech types, a second important network uses nouns as nodes. The method being analogous, it suffices to consider one example of a "noun node" together with the cross-references back to the node. One such small network is shown in (5):

- (5) (a) *diitu* [5/6] bos, woud
 ▶ -à *mwitu*; *kwitu*; *mwitu*
 (b) *kwitu* [loc 17/Ø ◀ *diitu*] cf var
 mwitu
 (c) *mwitu* [loc 18/Ø ◀ *diitu*; var
 kwitu; vgl *cisuku*] in 't bos/
 woud; a/h water
 □ -à ~ [cn adj] wild; iets dat
 groeit/verblijft in 't bos/woud;
 ngulube wa ~ wild varken/
 zwijn, everzwijn

Within the framework of simultaneous feedback, we maintained that we cannot agree with the statement that one is not able to experiment in lexicography. However, one cannot really experiment without making suggestions first. This is why the first parallel project, LCN, makes suggestions. And since LCN was actually published, suggestions had to be made for every possible aspect of the lexicon.

"There are various types of financiers that a lexicographer can approach," says Alberts (1999: 4). "Private funds may, as a first option, be raised by tapping one's own ... financial resources." This is exactly what was done for the first parallel project. The booklets themselves, which cost just 3US\$ each to produce, were given to the target users for free. From the moment the first parallel project was distributed, feedback was instantly received.

[3] Informal and formal feedback on the first parallel project

(a) Informal Files

On the day LCN was released, a series of Informal Files were opened to track every remark concerning the lexicon and its use. As far as dictionary users are concerned, we are convinced that these files represent a rare first-hand insight into target users' (initial) opinions regarding a dictionary's release and its subsequent use. It should be stressed that all the opinions were recorded in our capacity as natural participant observers. A different kind of feedback, namely academics' informal feedback, was received through discussions with scholars and following the reading of a paper at a Lexicography Seminar in Tervuren, Belgium (De Schryver 1998). The feedback in the Informal Files ranges from

plain suggestions to improve the layout, to unclear articles, codes and abbreviations, and valuable recommendations for different lemmatisations.

(b) The quest for formal feedback

Besides informal feedback, its formal counterpart was also collected. From the manifold standardised methods that are available to obtain formal feedback we opted for questionnaires through a mail survey. In order to view the different ways of receiving feedback from a somewhat broader perspective, one can make use of Galtung's methodological typology for differentiating between various forms of social research. According to him, there are nine "basic ways" of collecting data, as listed in (6).

(6) Basic ways of data-collection (Galtung 1967: 109-121)

	Non-verbal Acts	Verbal Acts	
		Oral	Written
Informal	1 participant observations	4 conversations / informants	7 letters / books / etc.
Formal & Unstructured	2 systematic observations	5 open interviews	8 open questionnaires
Formal & Structured	3 experimental observations	6 closed interviews	9 closed questionnaires

We see that the data collected in the Informal Files can be assigned to two types: the target users' remarks can be assigned to type 1, while all the scholars' remarks can be assigned to type 4. The type of data-collection to be expounded on now mainly belongs to type 9, and to a lesser extent also to type 8. Regarding the target users, it becomes apparent that we moved from one extreme (type 1) to the other (type 9). In this way we hope to have balanced the type of feedback collected.

Once one starts browsing the literature, one realises that the advantages of mailed questionnaires are legion. One huge drawback of mailed questionnaires, however, is the large percentage of nonresponses. Over the years, different procedures for reducing this large percentage of nonresponses have been suggested, the most tempting probably being Dillman's "Total Design Method" (1978), for short TDM. Dillman claims that his method boosts the response from 30 to 70%, a percentage comparable to face-to-face interviews. Basically, TDM is constructed around three keywords:

- (a) "personal character": one should try to give the mailed questionnaire a personal touch;
- (b) "trade tricks": one should use all the latest trade tricks, in that the questionnaire itself should be well-balanced and specifically designed for the specific target respondents; and

- (c) "follow-ups": one should incorporate three follow-up mailings.

TDM being a *total* design method, (a), (b) and (c) should *all* be implemented in order to attain the 70% response. With the variant we used, 69% of the questionnaires were eventually returned.²

(c) LCN Questionnaire

The Questionnaire itself, given in full as Appendix B, was four pages long and was packed with fifty questions covering all the various aspects of LCN, as well as some potential future endeavours. In spite of this large number of questions — with many of them even subdivided into subquestions — nearly all respondents tried to reply to every single question. It is a golden rule in an anonymous questionnaire not to place "personal questions" at the beginning, since the respondents might otherwise doubt whether the questionnaire is indeed intended to be anonymous. This is why the three questions with a personal touch were placed at the end (Q46-48). In addition, the questions were grouped together according to subject: Generalities (Q1-5), Layout (Q6-12), Abbreviations and Symbols (Q13-16), Grammar "within" the *Lexicon* (Q17-22), Nouns (Q23-29), Adjectives (Q30), Verbs (Q31-34), Cross-references (Q35-36), Outlook (Q37-40) and a Round-up (Q41-50). The way these different groups were laid out had to please the eye (cf. Dillman's keyword (b) above).

All questions were kept very short and with as simple a wording as possible. This is why, e.g., *inbrenge* (enter) was used instead of lemmatise, or *trefwoord* (headword) instead of lemma-sign. The use of easy phraseology and examples simplified communication between compilers and target users.

Closed questions ("yes/no" or "multiple-choice", where one simply had to tick off one or more alternatives) were the most common type. Other questions were more open-ended, in that different possibilities had to be rated using numerals. Lastly, only one tenth of the questions were true open questions (in which the respondents were asked to write their answers down).

(d) Analysing the formal target users' feedback on LCN

The analysis of the target users' formal feedback on LCN can be classified as follows:

- (i) the target users' needs and expectations;
- (ii) the target users' reference skills; and
- (iii) the target users' desires.

In this tripartite structure one instantly recognises the theoretical framework — [1] in (1), and the preliminary analysis. One or more representative questions

from the Questionnaire will now be used briefly to illustrate each of these classifications.

(i) With Q3 we went to the core of the target users' needs: We wanted to know for which type of data they consulted LCN. The results are shown in (7).³

(7) Why take to LCN?

LCN QUESTIONNAIRE		RESULTS (in %)		
Question	Answer ... suggested by us added by target user ...	Level		
		elem.	int.	adv.
3. For which purposes did you use the <i>Lexicon</i> most? (please place a '1' in front of the most important purpose, a '2' in front of the second most important, a '3' etc.)	... to verify the spelling of a word	16	22	21
	... to check the tones of a word	20	24	24
	... to get the meaning of a word	27	32	27
	... to have a translation equivalent at hand	8	9	11
	... to consult one of the tables	12	10	12
	... to study vocabulary	3	3	3
	... other:	—	—	—
	... to solve exercises ...	6	—	—
	... to find the stem of a verb ...	8	—	—
	... to look for words we have seen up to now ...	—	—	3

It is clear that every competence level (whether elementary, intermediate or advanced) used LCN primarily for semantic purposes (27%, 32% and 27% respectively). In addition, all levels rank the tonal dimension in the second place, orthography in the third, and tables in the fourth.

(8) Slot order

LCN QUESTIONNAIRE		RESULTS (in %)		
Question	Answer ... suggested by us added by target user ...	Level		
		elem.	int.	adv.
39. If you were to look up a word in a new edition of the <i>Lexicon</i> , which kind of information would you like to find first? (please place a '1' in front of the kind of information you would like to find first, a '2' in front of what you would like to find next, a '3' etc.)	... word class (if it is a verb, a noun, ...)	16	12	18
	... etymology (e.g.: a loan from language X)	7	7	2
	... from which word it is derived	11	12	10
	... a list of derivations	9	9	11
	... a list of synonyms	9	14	13
	... variant spelling(s)	12	13	13
	... a translation/description	19	19	20
	... an example to illustrate the headword	10	8	7
	... a petrified expression	7	4	6
	... other:	—	—	—

One cannot claim that this is also what the target users actually need. Since the target users were asked for which purposes they consulted LCN, the chosen data types were restricted to those available in LCN. To determine the target users' expectations in a lexicon of this kind, one can consider Q39. In Q39 we wanted to learn what particular kind of information the target users would like to find first in a future Cilubà-Dutch Dictionary. The results are shown in (8).

All three levels expect a translation/description first. The answers to Q39 thus confirm the conclusion drawn from Q3; that the target users' first and foremost need is semantic information.

(ii) As for the target users' *reference skills*, the Questionnaire dealt with both general reference skills and reference skills in a Lubà context. Regarding the general skills, the results confirm the preliminary analysis that these are rather well developed. As for the skills in a Lubà context, the answers to one particular question are very revealing. In Q15 the understanding of the "place-holder symbol", also called "repetition symbol" (cf. e.g. Hartmann and James 1998: 119), was tested in a Lubà context. This nontypographical structural marker, the tilde (~) in LCN, simply replaces the lemma-sign within a particular article. However, considering the variety of answers to Q15 shown in (9), one is surprised.

(9) Target users' reference skills in a Lubà context

LCN QUESTIONNAIRE			RESULTS (in %)		
Question	Answer	added by target user	Level		
			elem.	int.	adv.
15. In many dictionaries the headword is replaced by a tilde (~) within an article. This is also done in the <i>Lexicon</i> . Under the headword <i>bàna</i> one will find:	<i>bàna</i> [1/2] cf <i>ev mwàna</i> ; ~	<i>bàna</i>	78	73	64
	<i>bààbò</i> [1/2] cf <i>ev mwanaàbò</i> ;	<i>bàna</i> + remark	—	7	9
	...	<i>bàna</i>	6	20	9
	Which word should be read	?	6	—	—
	instead of the tilde (~)?	<i>baa</i>	—	—	9
 <i>bààbò</i>	(no answer)	11	—	9

In order to understand the confusion raised by Q15, one needs to take account of the following Lubà rule. A word-final high tone becomes low when the preceding syllable contains a falling tone or a low tone *and* when it is followed by a pronominal prefix carrying a low tone. With this rule, the above becomes *bàna* *bààbò*. Although this simple rule is mentioned within the first weeks of the Cilubà courses, the answers to Q15 show that the rule is *not* applied by the majority of target users. Even if one makes provision for this potential change in tone in the front matter of the dictionary by pointing out that one should take "floating tones" into account — as was done in LCN on p. xv, § 4. — this is not sufficient. The answers show that the place-holder symbol cannot be used whenever a word-final high tone changes into a low tone, as only one out of ten target users possesses this one "reference skill in a Lubà context".

(iii) The bulk of the Questionnaire dealt with the target users' *desires*. The Questionnaire was rather exhaustive in this respect, as it dealt with dictionary aspects ranging from grammar in the lexicon to nouns, adjectives, verbs, cross-references, layout, and a final round-up. Since it is impossible to review all this feedback here, we will restrict our discussion to two specific issues: the lem-

matisation of nouns on the macrostructural level and the "node networks" centred around verb stems.

In the Questionnaire, seven questions dealt specifically with nouns, of which four dealt with the lemmatisation of nouns on the macrostructural level. The latter four are shown in (10).

(10) Nouns: macrostructure

LCN QUESTIONNAIRE		RESULTS (in %)		
Question	Answer	Level		
		elem.	int.	adv.
23. In the <i>Lexicon</i> nouns are entered under their singular. Some plurals, however, had to be provided for (cf. p. xiv, § 3.). ⁴ How would you like to look up nouns?	<input type="checkbox"/> suggested by us			
	<input type="checkbox"/> solely as singular	11	7	9
	<input type="checkbox"/> as singular, with provision for some plurals	58	93	91
	<input type="checkbox"/> always both as singular and plural	25	—	—
	<input type="checkbox"/> other:	—	—	—
	(no answer) <input type="checkbox"/>	6	—	—
24. In your view, where should irregular plural forms be presented?	<input type="checkbox"/> within the article of their singular	28	40	27
	<input type="checkbox"/> under their plural (as a separate headword)	28	20	27
	<input type="checkbox"/> other:	—	—	—
	under both <input type="checkbox"/>	39	33	36
	(no answer) <input type="checkbox"/>	6	7	9
25. In the <i>Lexicon</i> the class numbers for singular and plural are given. Instead of 'numbers' one could use the 'nominal prefixes' themselves. According to you, are:	<input type="checkbox"/> class numbers handier than nominal prefixes	58	40	82
	<input type="checkbox"/> nominal prefixes handier than class numbers	31	60	9
	<input type="checkbox"/> class numbers & nominal prefixes redundant	—	—	—
	<input type="checkbox"/> other:	—	—	—
	use both, e.g. 1 «mu-» <input type="checkbox"/>	6	—	9
	(no answer) <input type="checkbox"/>	6	—	—
28. One could also enter nouns under their stem. In this way one would for instance find <i>muntu, bantu, buntu, cintu, bintu</i> and <i>kantu</i> under <i>-ntu</i> . This would be:	<input type="checkbox"/> extremely user-friendly	11	7	—
	<input type="checkbox"/> much too complicated	67	87	91
	<input type="checkbox"/> other:	—	—	—
	the <i>Lexicon</i> would become too large <input type="checkbox"/>	6	—	—
	better: both as stem and noun <input type="checkbox"/>	6	—	—
	interesting to add the stem under the noun <input type="checkbox"/>	6	—	—
	it reveals lexical relations <input type="checkbox"/>	—	7	9
	(no answer) <input type="checkbox"/>	6	—	—

In Q23, after reminding target users of the lemmatisation approach utilised in LCN, we sought to learn how they would like to look up nouns. For only 58% of the elementary learners, LCN's approach corresponded to the way they would like to look up nouns. For intermediate and advanced learners, LCN is on the right track, as 93% and 91% of them respectively mark LCN's approach as conforming to their desires. Since 25% of the elementary learners opted for both singulars and plurals, whilst *none* of the more advanced learners did, this suggests a certain unfamiliarity with the noun class system among beginners. In addition, one target user rightly remarked that entering both singulars and plurals would be handy but uneconomical because of the space it would take up in the dictionary.⁵

The target users' desire is thus: *For intermediate and advanced learners one can safely lemmatise nouns under their singulars as long as provision is made for*

irregular plurals, but for elementary learners additional guidance on the noun class system will have to be provided for if this is to be a successful approach. Alternatively, one might consider lemmatising both singulars and plurals for inexperienced users.

In Q24 we wanted to know where the target users would like to find irregular plural forms. When one considers the large number of respondents who added "under both", a certain logic appears. Moreover, one target user (rightly) observed that cross-references should link plurals and singulars.

Hence, the target users' desire is: *Irregular plural nouns must be mentioned both within the article of their singular forms and be lemmatised under their proper alphabetical position, whilst cross-references should link both forms.*

In Q25 we sought to learn whether target users need gender information, and if so, in which format. What is extremely interesting here is the option not one single target user chose, namely that class numbers as well as nominal prefixes are redundant. Though some respondents suggested including both class numbers and nominal prefixes, roughly half of the elementary and intermediate learners and nearly all the advanced learners opted for class numbers.

The target users' desire is thus: *Noun gender information is crucial and should, for advanced learners, preferably be coded using class numbers, whilst elementary and intermediate learners might desire to receive some additional guidance.*

In Q28 we wanted to know the target users' opinions about the "stem tradition". 67% of the elementary, 87% of the intermediate and no fewer than 91% of the advanced learners consider this tradition to be much too complicated. If one looks at the respondents' additions, they suggested that it might reveal lexical relations, that it might be useful to add the stem within each noun article, that it might be better to include both stems and nouns, and that the lexicon might become too large with the inclusion of stems. Also, one target user (rightly) suggested that this approach would simply require too much look-up time.

Hence, the target users' desire is: *Under no circumstances should nouns be lemmatised according to the "stem tradition".*

The true power of a questionnaire is well illustrated through these four questions. As a second topic to illustrate this power, one can consider the "node networks" centred around verb stems. Without going into the underlying theoretical details of the need for harmonising lumping and splitting, Q35 asked the target users whether they thought the node networks were useful. The results are shown in (11).

Even though the answers to Q35 are somewhat dispersed around different "useful aspects", especially for the intermediate level, the target users' judgement stands out clearly: the node networks *are* useful. The favoured aspect is that one is able to see "connections".

(11) Cross-references

LCN QUESTIONNAIRE		RESULTS (in %)			
Question	Answer		Level		
	<input type="checkbox"/> suggested by us	<input type="checkbox"/> added by target user	elem.	int.	adv.
35. Under the verb -dyà one encounters a 'tail' of headwords (bidyà; cidillu; cyákudyà; -dilka; -dilkiibwa; -dilla; -dilsha; mudli; Mudilla-mpiku) which are all connected to the verb stem -dyà. Each of those headwords from the tail has a cross-reference back to -dyà. All this is:	<input type="checkbox"/> not useful, because	<input type="checkbox"/> I don't know why one would need this (no explanation) <input type="checkbox"/>
	<input type="checkbox"/> useful, to see connections		72	52	86
	<input type="checkbox"/> useful, to remember words more easily		11	9	9
	<input type="checkbox"/> useful, to know what is in the <i>Lexicon</i>		6	6	5
	<input type="checkbox"/> other:
	<input type="checkbox"/> useful, with a short meaning <input type="checkbox"/>		—	7	—
	<input type="checkbox"/> useful, but it is impossible to include all references without giving a single meaning <input type="checkbox"/>		—	7	—
	<input type="checkbox"/> (no answer) <input type="checkbox"/>		11	7	—

It is hoped that these examples illustrate that a carefully designed questionnaire enables the lexicographer to track, evaluate and adjust all possible aspects of the dictionary process.

When simultaneous feedback was described at the start of this article, it was not only claimed that a dictionary project could be launched within this framework, but also that it should run concurrently with the development of electronic corpora. The best time to start compiling the main electronic corpus is between the release of the first parallel project and the stage when the feedback on that first parallel project is analysed. In doing so, the progress of the project is not delayed. In truth, simultaneous feedback from the parallel projects, combined with the wealth of information derived from the electronic corpus, constitute the basic building blocks for the main dictionary. The creation of an electronic Lubà corpus is precisely the subject of the next section.

[4] Main electronic corpus: Word-frequency studies, and The lemma-sign list

The *Collins COBUILD English Language Dictionary* (Sinclair 1987), known as COBUILD1, was derived from the earliest electronic megacorpora, the 7,3-million-word COBUILD Main Corpus (Renouf 1987). As COBUILD1 contained roughly 70 000 "references", this meant that approximately 100 running words in the corpus were required for each reference in the dictionary. Our objective was to apply this *same ratio* to the second parallel project. Increasing LCN's lemma-sign list by 20%, from 2 500 to 3 000 lemma-signs, seemed feasible. Applying the ratio derived from COBUILD1 meant that we had to create an electronic corpus of 300 000 words.

Furthermore, from the various corpora types discussed in the literature on corpus linguistics (cf. e.g. Kennedy 1998: 2, 20, 52, 62, Summers 1993: 186, 190)

we retained the notion of a "structured corpus", and decided to follow a combined genre/topic stratification.

The resulting corpus, which has been named "Recall's Cilubà Corpus", henceforth RCC, consists of seven subcorpora (i.e. magazines, traditional stories, informal literature, textbooks, scientific works, religious works and miscellaneous sources) of circa 40 000 words each and one subcorpus (i.e. poetry and proverbs) of circa 20 000 words.⁶ The outline of this "structured Lubà corpus of 300 000 words" is thus straightforward. Each genre/topic area is given equal weight, except for the block poetry and proverbs, which is only half the size of the others. Moreover, the different spoken sections in RCC amount to 17,4%.

A recent article by Kilgarriff (1997) is titled "Putting Frequencies in the Dictionary". As the title suggests, frequencies are *put in* the dictionary. This, of course, might be a nice addition to an existing dictionary. In our understanding of the utilisation of word-frequency studies, we would rather use frequency data to help in deciding *what* to include *how* in a dictionary. As such, a lemmatised frequency list is a crucial aspect of simultaneous feedback. If one wishes to compile, say, a parallel dictionary containing 2 000 lemma-signs, one can start by selecting the 2 000 most frequent lemmata in the corpus. If, however, the aim is to compile a parallel dictionary containing 5 000 lemma-signs, one can select the top 5 000 lemmata in the corpus, etc.

To enable the compilation of the lemma-sign list for the second parallel dictionary using data from word-frequency studies, we needed a "lemmatised frequency list" and information on the "distribution" of those lemma-signs across the eight subcorpora. Statistical software tools were thus required with which RCC could be analysed. We singled out WordSmith Tools, for short WST (cf. URL WordSmith), for its versatility in handling corpus queries.⁷

In order to generate a lemmatised frequency list for the most frequent types we simply went through the top 1 000 items of the corpus and lemmatised them "by hand". For nouns this meant that when we encountered a singular form, we added the frequency of the plural form (or vice versa), where applicable. For verbs this meant that we kept track of those verbs we had already encountered and added the frequency of every single "conjugated form" we encountered subsequently. In addition to this "true lemmatisation" we joined divergent orthographies. The latter was done for all possible parts of speech. In this way we obtained a new list in which well over the first 600 types had been lemmatised.

The inclusion or omission of particular lemmatised forms cannot be based solely upon overall frequencies. This is best illustrated by the counts in (12), which show the distribution of all the unlemmatised constituents of *one* particular lemma across the eight subcorpora.

Whether the singular noun **munsantu** or the plural form, or whether the word-final **-u** or the word-final **-o** variant, the distribution is in all instances extremely uneven, with a clear bias towards religious works. Therefore, lem-

(12) Distribution of one lemma across the eight subcorpora

rank	item	#	total	magaz.	tradi.	inform.	textb.	scient.	refig.	miscel.	poetry
434	munsantu	2	78	21	0	0	0	0	57	0	0
855	munsanto	3	42	3	0	0	2	0	37	0	0
1,536	bansantu	1	22	0	0	0	0	0	22	0	0
2,015	bansanto	1	16	0	0	0	0	0	16	0	0

mata of which all constituents show such an uneven distribution certainly do not qualify for inclusion among the "truly frequent" lemma-signs.

This procedure enabled us to determine the top 600 lemmata in the 300 000-word RCC. This data provides us with the means to mark the 600 most frequent lemma-signs in the second parallel dictionary project. While work on the electronic corpus was progressing, a last type of feedback on the first parallel dictionary project was received, namely formal academics' feedback.

[5] Reanalysis of the target users' desires, and Formal academics' feedback

Regarding this form of simultaneous feedback, the idea is simply to contrast the target users' desires with formal academics' feedback. Since a number of lemmatisation strategies proposed in LCN are a kind of package deal, for example the node networks discussed above, one cannot implement *all* the target users' desires *at once* without taking the formal academics' critical reviews into account. On the node networks suggested in LCN, Gouws and Prinsloo (1998: 17, 31-32) wrote:

The mediostructure, that is the system of cross-referencing, is a lexicographic device that can be used to establish relations among different components of a dictionary ... An excellent example in African language lexicography where mediostructure has been employed as a powerful access structure is the *Lexicon Cilubà-Nederlands* (LCN) compiled by De Schryver and Kabuta. This dictionary is highly successful in interconnecting the knowledge elements represented in different sectors of the dictionary on several levels of lexicographic description to form a network ... The compilers of LCN are aware of the benefits of "keeping together what semantically and grammatically belong together" but also of the need (a) to avoid extremely long entries and (b) to ensure proper treatment of each derivation in terms of grammatical, tonal and lexical information ... The compilers of LCN thus succeeded in harmonising lumping and splitting, capturing the advantages of both these approaches. It can, of course, be argued that the listing of the different derivations [in the tail] occupies precious space in the dictionary. However, by substantially reducing the font size, this redundancy is diminished.

Simultaneous informal and formal target users' feedback, interlinked with

simultaneous informal and formal academics' feedback, in combination with the active use of electronic corpora, should ultimately produce a "good dictionary answering all potential questions". The second parallel dictionary was such an attempt.

[6] Second parallel dictionary based on a second parallel test-corpus

The second parallel dictionary, the *Beknopt woordenboek Cilubà-Nederlands* (De Schryver and Kabuta 1998), henceforth BCN, is probably best introduced with some sample pages. Two pages can be found in Appendix C. Notwithstanding our conviction that the prime function of a lemmatised frequency list is to be the basis for the compilation of the lemma-sign list of a dictionary, we can obviously also use this data to *indicate* frequencies in the dictionary. In terms of Kilgarriff (1997) it means "putting frequencies in the dictionary". On BCN's page 23 for example, shown in Appendix C, one sees that *dîba* is preceded by a ①, which means that it belongs to the 200 most frequent lemmata; *dibòko* is preceded by a ②, which means that it belongs to the 400 most frequent ones; and *dib.* is preceded by a ③, which means that it belongs to the 600 most frequent ones. Together with all the lemmata that have not been given a symbol (80% of them), this illustrates the four-way categorisation utilised in BCN. It should be noted that this procedure is a variant of the approach found in the *Longman Dictionary of Contemporary English* (Summers 1995³), known as LDOCE3. As such, this is probably the first African language dictionary in which frequencies are indicated.

Of the other new features introduced in BCN, we will now focus on one which illustrates well how the target users' feedback can effectively be taken into account, whilst not jeopardising the academics' feedback. It will be recalled that we needed a way to provide additional guidance on the noun class system for elementary learners (and to a lesser extent also for intermediate learners). From the sample pages in Appendix C a page on the left-hand side of BCN, thus an even page, displays an inserted text which lists the most important affixes. In addition, a page on the right-hand side of BCN, thus an odd page, aims to display the most important abbreviations and symbols (also meeting a target users' need). Since these two "running footers" are repeated throughout the central lemma-list, we propose calling them the even-repetitive inserted text, and the odd-repetitive inserted text, respectively ERIT and ORIT.

The affixes tabulated in ERIT are straightforward. With the canonical form of the nominal prefixes for both participants and classes as a point of departure, ERIT then lists the subject concords, followed by the object concords and the pronominal prefixes, and ends with the morphemes used to form negative tenses. In other words, *the 100 items tabulated in ERIT synthesise the core of the Cilubà concordance system at a single glance.*

The abbreviations and symbols tabulated in ORIT were carefully chosen. Firstly, some guidance with respect to the frequency-annotation symbols ①, ②,

and ③ is given. Secondly, in order to decide which abbreviations to include in ORIT and which not, the frequencies of all the abbreviations in BCN were counted. Finally, the most important symbols used throughout BCN are listed in ORIT's last column.

With these two repetitive inserted texts, it is truly hoped that *all* the target users will more easily find their way, not only through the noun class system, but also through the microstructure of the articles and the complex network of cross-references. Just as was the case with the first parallel dictionary, the release of this second parallel dictionary immediately provided feedback.

[7] Informal and formal feedback on the second parallel dictionary

At present, feedback on the second parallel dictionary, BCN, is informal. A second series of Informal Files is being assembled in which this feedback is documented. The picture looks very positive. Nonetheless, it is still too early to analyse them as was done for LCN's Informal Files. As for academics' informal feedback, a number of very encouraging e-mails from various scholars were received in which BCN has been labelled a "super book" or even a "chef d'oeuvre". These scholars point out "many interesting features", of which the two repetitive inserted texts and the experimentation with explicit frequency information seem to be favoured. One scholar even claimed that Cilubà "is currently better served, corpus-wise, than most of the big US dictionary publishers (who remain doggedly in the precorpus era)".⁸

[8] Main electronic corpus: Concordance lines and Dictionary articles

Regarding *macrostructural* planning, the second parallel project, BCN, has reached a point where virtually all target users are satisfied with the general dictionary structure and layout, and especially with the means utilised to lemmatise the different parts of speech. As such, simultaneous feedback in combination with the counts derived from the electronic corpus proved to be of enormous value in solving some of the long-standing lemmatisation problems for African languages.

On the *microstructural* level the electronic corpus also proves to be invaluable. Indeed, concordance lines derived from the corpus constitute the main tool for enhancing the dictionary articles, in respect of examples of usage, typical collocations, ordering of senses, etc.

[9] Continuous monitoring of the target users' desires, and Recruiting of additional staff

Van der Merwe (1999: 4) claims: "It is not easy to determine the needs of dictionary users. In the literature a variety of methods used by researchers are

mentioned, for example questionnaires, interviews, tests and observation. Not one of these methods has been successful up till now in doing a complete and efficient needs analysis." She is entirely in the right when she maintains that "not one of these methods" can have the intended result, and this is precisely why, within the framework suggested here, various methods are used and contrasted with one another. Moreover, the continuous monitoring of the target users' desires is built into the framework of simultaneous feedback. This ensures that the target users' desires transcend the target users' needs, expectations and reference skills.

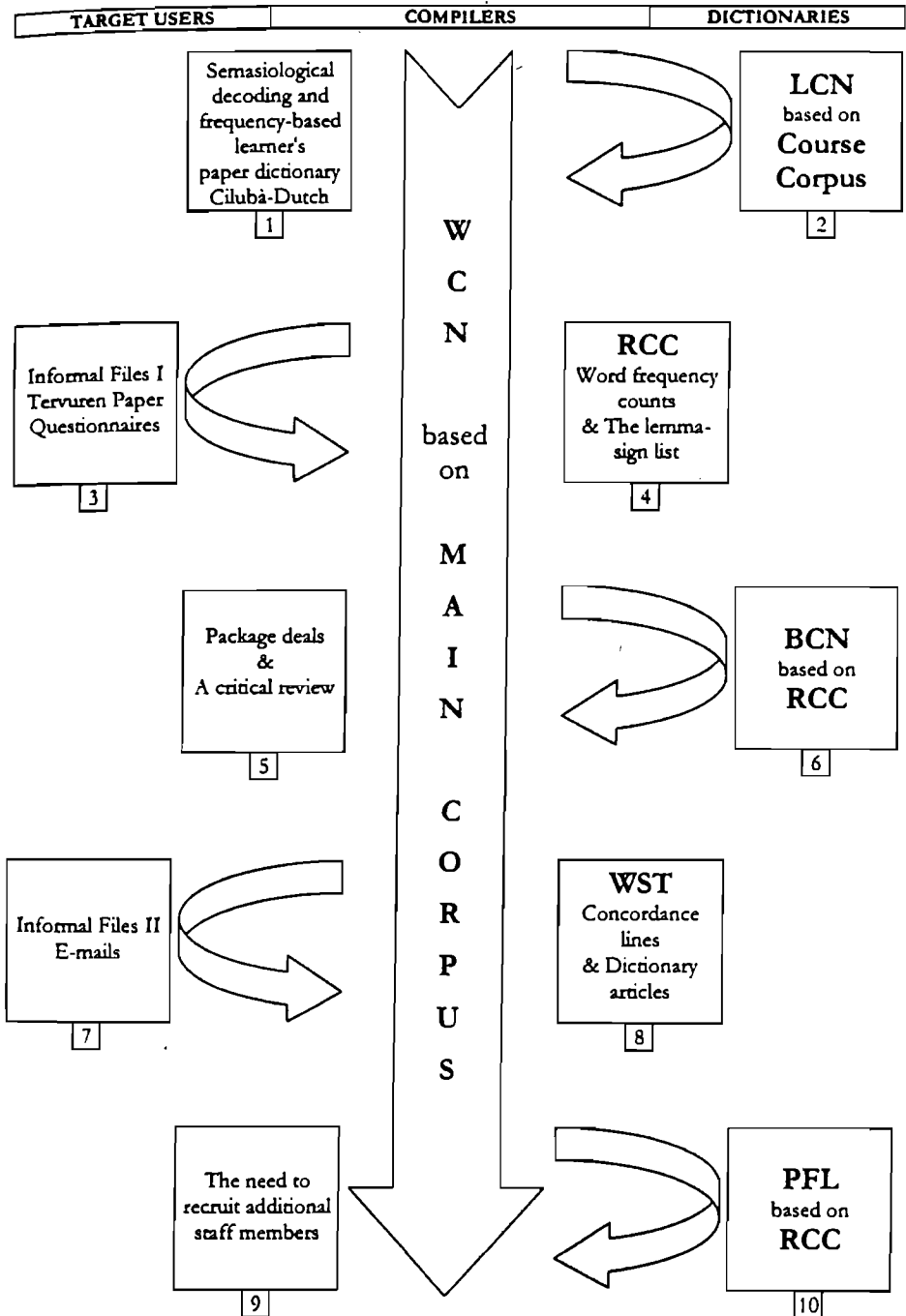
The features provided by WordSmith Tools (WST) with which an electronic corpus can be analysed, are impressive. Yet, at the same time, the corpus-query options of this software are so numerous that the amount of data generated becomes too much to cope with by any small team of lexicographers. Viewed differently, while momentum has increased after each and every feedback batch, the staff has remained the same size. If future parallel projects are to extract the full potential out of concordance information to enable the creation of corpus-based dictionary articles, it is high time to enlarge the team with *several* mother-tongue speakers of Cilubà. Any dictionary project which has gone through two parallel projects will feel the need for recruiting of additional staff while the project grows. Hence the question is how one should read the last component of the theoretical framework in this light — [10] *more* parallel dictionaries?

[10] More parallel dictionaries used as experimental tools in order to refine the main dictionary

Taking on additional staff members requires, in one way or another, a substantial amount of funding. As long as this is not available, the entire project seems bound to remain a small dictionary project. However, this does not mean that *more* parallel projects cannot be launched in order to plan aspects of the main dictionary project.

Potential additional parallel projects could revolve around word-frequency studies. This led us to introduce a new approach to the phonetic description of a language's lexicon, the findings of which were published as *Cilubà Phonetics: Proposals for a "Corpus-based Phonetics from Below"-approach* (De Schryver 1999). Of particular interest to the main dictionary project, is that this study resulted in two short wordlists, a *Phonetic Frequency-Lexicon Cilubà-English* consisting of 350 lemma-signs (De Schryver 1999: 55-68), and a converted *Phonetic Frequency-Lexicon English-Cilubà* (De Schryver 1999: 69-87), both for short PFL. Provided that the target users know the conventions of the International Phonetic Alphabet (IPA), these two wordlists enable them to "retrieve", "learn" and "pronounce" the 350 most frequent *words* from the Lubà language.

(13) The concept of simultaneous feedback as implemented in the current dictionary project



Round-up of [1] - [10]

The theoretical plan of the concept of simultaneous feedback as presented in (1) might have seemed a bold statement. The feasibility of this concept, however, has been illustrated from a very practical perspective through a small-scale Cilubà-Dutch dictionary project. The way the concept has been implemented thus far in this undertaking is summarised in (13).

In conclusion

In this article it was argued that the feedback from target users, indispensable in the compilation of any modern dictionary, should be obtained while the compilation of the dictionary is still in progress. This is achieved by launching several small-scale parallel dictionaries from which informal as well as formal feedback from users and academics is immediately channelled back to the compilation process of a main dictionary. Such simultaneous feedback is not only invaluable in the compilation of the macrostructure but also provides useful information for improvement on the microstructural level. Furthermore, the ever-increasing role of the corpus as a lexicographic aid has been emphasised and it was indicated how corpus creation should be planned to fit into the concept of simultaneous feedback by introducing subcorpora.

The parallel dictionaries fulfil the urgent need that the target users can quickly be provided with lexis while awaiting an unabridged electronic corpus-based main dictionary. Moreover, channelling feedback from parallel projects into the main project enables the compilers to integrate all criticism into the compilation methodology itself. We are consequently convinced that the concept of simultaneous feedback successfully provides present-day lexicographers with an inexpensive head start and a tool for producing dictionaries compiled according to the latest trends and most modern approaches in lexicography.

It was emphasised throughout this article that the concept of simultaneous feedback can in principle be applied to any language. It is expected, however, that it will make a significant impact on dictionary compilation for the *African languages*. We therefore conclude with the words of Gouws (1998):

One of the characteristic features of the Renaissance was the emergence of dictionaries and their spreading of the light of learning so desperately needed in the aftermath of the Dark Middle Ages. One of the characteristic features of the African Renaissance could be the emergence of dictionaries for all the African languages and productive and fruitful metalexigraphic co-operation between all the stakeholders.

It is hoped that the concept of simultaneous feedback be part of that African Renaissance.

Notes

1. Since this article is being submitted for publication in a South African journal, sensitivity with regard to the term *Bantu* languages is exercised in our choice to use the term *African* languages instead. However, keep in mind that the latter includes more than just the Bantu language family.
2. This response compares favourably with similar projects by other experts. For example, of the 1 600 questionnaires distributed in the EURALEX- and AILA-sponsored *Research Project into Dictionary Use* (Atkins and Varantola 1998: 24), 1 140 responses (71,25%) were eventually received.
3. In order to turn the respondents' rankings into percentages, a formula was used whereby different weights were given to each number, with number "1" the highest weight and "... " (the absence of a number) the lowest weight. This same formula was used to quantify all answers in which the respondents were asked to give a ranking.
4. This is a reference to LCN's front matter where four types of "singulars with a cross-reference to their plural form" are discussed.
5. Single, lengthy additional remarks by respondents, like this one, have not been tabulated throughout the analysis of the Questionnaire.
6. The publication of LCN in December 1997 marked the inauguration of the Research Centre of African Languages and Literatures (Recall), based at the Department of African Languages and Cultures of the University of Ghent, Belgium. *Recall's Cilubà Corpus* (RCC) was the first electronic corpus built at Recall.
7. This suite of lexical analysis tools was presented by M. Rundell (1998: 16/3.19) during the Afrilex-Salex'98 Tutorial.
8. Since this quote originates from informal communication, the author will not be named.

Acknowledgement

Financial assistance from the National Research Foundation (NRF) towards the presentation of "The Concept of 'Simultaneous Feedback'" as a paper at Euralex 2000 is hereby acknowledged.

Cited references

URL (Universal Resource Locator)

WordSmith <http://www.liv.ac.uk/~ms2928/wordsmith/screenshots>

References

- Alberts, Mariëtta. 1999. The Importance of a Business Plan when Planning a Lexicographic Project. Paper presented at the Fourth International Conference of the African Association for Lexicography, held at the University of Pretoria, Pretoria, 5-6 July 1999. *Lexikos* 9: 188-197.

- Atkins, B.T. Sue (Ed.). 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Max Niemeyer.
- Atkins, B.T. Sue and Krista Varantola. 1998. Language Learners Using Dictionaries: The Final Report on the EURALEX / AILA Research Project on Dictionary Use. Atkins, B.T. Sue (Ed.). 1998: 21-81.
- Barnhart, Clarence L. 1967² [1962]. Problems in Editing Commercial Monolingual Dictionaries. Householder, Fred W. and Sol Saporta (Eds.). 1967² [1962]. *Problems in Lexicography*: 161-81. Bloomington: Indiana University.
- De Clercq, Auguste and Emile Willems. 1960³. *Dictionnaire Tshilubà-Français*. Leopoldville: Imprimerie de la Société Missionnaire de St. Paul.
- De Schryver, Gilles-Maurice and Ngo S. Kabuta. 1997. *Lexicon Cilubà-Nederlands. Een circa 2500-lemma's-tellend strikt alfabetisch geordend vertalend aanleerderslexicon met decodeer-functie ten behoeve van studenten Afrikaanse Talen en Culturen aan de Universiteit Gent*. Ghent: Recall.
- De Schryver, Gilles-Maurice and Ngo S. Kabuta. 1998. *Beknopt woordenboek Cilubà-Nederlands & Kalombodi-mfundilu kàà Cilubà (Spellingsgids Cilubà). Een op gebruiksfrequentie gebaseerd vertalend aanleerderslexicon met decodeerfunctie bestaande uit circa 3.000 strikt alfabetisch geordende lemma's & Mfundilu wa myakù idi itàmba kumwènèka (De orthografie van de meest gangbare woorden)*. Ghent: Recall.
- De Schryver, Gilles-Maurice. 1998. *Cilubà Lexicography: A First Approach*. Unpublished paper presented at a Lexicography Seminar, Linguistic Department of the Musée royal de l'Afrique centrale, Tervuren, 20 February 1998.
- De Schryver, Gilles-Maurice. 1999. *Cilubà Phonetics: Proposals for a "Corpus-based Phonetics from Below"-approach*. Ghent: Recall.
- De Schryver, Gilles-Maurice. 1999a. *Bantu Lexicography and the Concept of Simultaneous Feedback. Some Preliminary Observations on the Introduction of a New Methodology for the Compilation of Dictionaries with Special Reference to a Bilingual Learner's Dictionary Cilubà-Dutch*. Unpublished M.A. dissertation. Ghent: University of Ghent.
- Dillman, Don A. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- Galtung, Johan. 1967. *Theory and Methods of Social Research*. Oslo: University of Oslo.
- Gouws, Rufus H. 1998. Chairperson's Annual Report: 1997-1998. Unpublished report read at the Third International Conference of the African Association for Lexicography, Potchefstroom University for Christian Higher Education, Potchefstroom, 29-30 June 1998.
- Gouws, Rufus H. and D.J. Prinsloo. 1998. Cross-referencing as a Lexicographic Device. *Lexikos* 8: 17-36.
- Hartmann, Reinhard R.K. and Gregory James. 1998. *Dictionary of Lexicography*. London: Routledge.
- Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London: Longman.
- Kilgarriff, Adam. 1997. Putting Frequencies in the Dictionary. *International Journal of Lexicography* 10(2): 135-155.
- Lombard, F.J. 1994. Lexicographer, Linguist and Dictionary User: An Uneasy Triangle? *Lexikos* 4: 204-214.
- Osselton, Noel E. 1983. On the History of Dictionaries. The History of English-language Dictionaries. Hartmann, R.R.K. (Ed.). 1983. *Lexicography: Principles and Practice*: 13-21. London: Academic Press.

- Prinsloo, D.J. and Gilles-Maurice de Schryver. 1999. The Lemmatization of Nouns in African Languages with Special Reference to Sep di and Cilubà. *South African Journal of African Languages* 19(4): 258-275.
- Renouf, A. 1987. Corpus Development. Sinclair, John M. (Ed.). 1987. *Looking Up: An Account of the Cobuild Project in Lexical Computing*: 1-40. London: Collins ELT.
- Rundell, Michael. 1998. Corpus Exploration: Extracting Lexicographic Data from a Corpus. Atkins, B.T. Sue, Michael Rundell and Rufus H. Gouws. 1998. *Afrilex-Salex '98. A Training Course in the Compilation of Bilingual Dictionaries*: 16/3.1-19. Unpublished course material of a tutorial held at the University of Pretoria, Pretoria, 7-18 September 1998.
- Sinclair, John M. (Ed.). 1987. *Collins COBUILD English Language Dictionary*. London: HarperCollins Publishers.
- Summers, Della. 1993. Longman/Lancaster English Language Corpus — Criteria and Design. *International Journal of Lexicography* 6(3): 181-208.
- Summers, Della (Director). 1995³. *Longman Dictionary of Contemporary English*. Harlow: Longman Dictionaries.
- Tono, Yukio. 1992. The Effect of Menus on EFL Learners' Look-up Processes. *Lexikos* 2: 230-253.
- Van der Merwe, Michele. 1999. The User Perspective in South African Lexicography. Unpublished paper presented at the Fourth International Conference of the African Association for Lexicography, held at the University of Pretoria, Pretoria, 5-6 July 1999.
- Wiegand, Herbert E. 1998. *Wörterbuchforschung: Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. 1. Teilband. Berlin: Walter de Gruyter.

Abbreviations used in reference to dictionaries and lexica

BCN	<i>Beknopt woordenboek Cilubà-Nederlands</i> . De Schryver and Kabuta. 1998.
COBUILD1	<i>Collins COBUILD English Language Dictionary</i> . Sinclair (Ed.). 1987.
LCN	<i>Lexicon Cilubà-Nederlands</i> . De Schryver and Kabuta. 1997.
LDOCE3	<i>Longman Dictionary of Contemporary English</i> . Summers (Director). 1995 ³ .
PFL	<i>Phonetic Frequency-Lexicon Cilubà-English-Cilubà</i> . De Schryver. 1999: 55-87.
WCN	<i>Woordenboek Cilubà-Nederlands</i> . In progress.

bakulú bábá	Lexicon Cilliers- Mouton 1986	bábá
bakulú bábá [1/2] cf ev. <i>mulwábá</i>		hoeveel; bantu -? hoeveel personen?
bálá [x/6 < Fr(Eng) "bar"] bar; discotheek		-banga [tw] 1 [vgl -twádija] beginnen; [asp hww] <initiatief> <i>bangá didyá</i> [IMP GER] begin te eten; 2 gereedmaken, voorbereiden; op voorhand beschikken
-bala' [tw] lezen; <i>mukándá wá kubala</i> (lees)boek		> <i>bangabanga</i> , -bangile; -bangulula; -bangile; -bangulu; -mbeangile
-bala' [tw] tellen; opsmen		<i>bangabanga</i> [bw < db -banga] aanvankelijk; vooraleer; vroeger; eerst
> <i>bédika</i> ; <i>balulula</i> ; <i>dibédika</i> ; <i>babalama-táka</i>		<i>banganyi</i> [ny12] wie; dewelke
-bála [iww] schijnen; blinken, glinsteren, schitteren		-bangila I [tw, app -banga] voor iemand voorbehouden of gereedmaken; vooraf bespreken; II [adj < I] verloofd
-bále [adj] schaars, zeldzaam, (zeer) weinig (talrijk); <i>mákúta ma- heel weinig geld</i>		-bangulula [tw, rep -banga] herbeginnen, opnieuw beginnen
-balulula [tw, rep -bala'] tellen		-banjija [tw, cau -banja] 1 de pasgehuude in haar huis houden installeren (door 't voorzien v huisraad); 2 aanzien geven
balúma [1/2] cf ev <i>mulúma</i>		-banza [tw, vgl <i>bubanja</i> , <i>mubanja</i>] huishouden beginnen (na de definitieve voltrekking v/h huwelijk)
□ -á - [cn adj; vgl -lúma] mannelijk		> -banjija; <i>tábánat</i>
-bamba [iww] zich verbergen, schuilhouden		<i>Basangana</i> [v/mpn < -sangana] De-gevondenen
> -bambila		<i>basédyanganyi</i> [1/2] 1 cf ev <i>musédyanganyi</i> ; 2 echtpaar; 3 schoonfamilie
-bamba [tw] stevig vasthouden; vastmaken		<i>Bátutsi</i> [eig] de Tutsi
-bambila [tw, app -bamba] zich verbergen/schuilhouden voor		<i>báya</i> - [1/2a] echtgenoot, man
bána [1/2] cf ev <i>mwána</i> ; - <i>bábá</i> [1/2] cf ev <i>mwánábá</i> ;		> <i>báyendá</i>
- <i>báányi</i> [1/2] cf ev <i>mwánáányi</i> ; - <i>béétú</i> [1/2] cf ev <i>mwanéétú</i> ; - <i>béndé</i> [1/2] cf ev <i>mwánéndé</i> ; - <i>mulongo</i> [ud] veel kinderen		<i>báyendá</i> [1/2a < báya-] haar man
-bánda [i/tww] klimmen, stijgen, naar omhoog g; opstappen		<i>bé</i> [bw] 1 veel, talrijk; 2 [op 't einde v/e zin(stuk) of na adj] zeer; <i>kuulu - zeer hoog</i>
> -bándila; <i>eibándilu</i>		-bedi [adj; syn -á <i>kumpála</i>] eerste; <i>musangu mu-</i> eerste keer
-bándila [tw] aanwijzen, tonen; beschouwen, bekijken		> <i>dyábedi</i> ; <i>mbedi</i>
-bándila [iww, app -bánda] zich verheffen tegen iemand		<i>béábé</i> [' pos2+2pe] jouw; <i>bána</i>
<i>bángá'</i> [bw] in tegenstelling tot wat gezegd of verwacht werd; echt, inderdaad; dus, zo		
<i>bángá'</i> 1 [ngálpm] hoeveel; 2 [ngá2pm] hoeveel; 3 [ngá2]		

-jiija	Lexicon Cilliers- Mouton 1986	-ka/á-
-jiija [tw, cau -já] doen dansen; dansfeest geven; [fig] bewegen		-jingila [tw, app -jingal] (op)rollen rond; omringen
-jiika [tw] begraven, i/d grond steken, bedelven		-jiwuka [iww] blootgesteld z aan gevaar (v zich te kwetsen); verongelukken
-jiila [tw, app -já] dansen ter ere v		-jiya [iww] betrouwen, diep bedroefd z
-jika [iww] op z; gedaan, voorbij z; tot 'n eind k		<i>jóógráfi</i> [1/4b < Fr "géographie"] aardrijkskunde
> -jikija		-juuka [iww] 1 rechtstaan; 2 zich voordoen
-jika [tw] afwenden v; tegenhouden, belemmeren; beschermen (tegen)		-jyama [tw] vastzitten
> -jikila		> -jyamisha
-jikija [tw, cau -jika] beëindigen, afmaken, gedaan maken met; [asp hww] <voltooiend> <i>pawáájikijá kwimba</i> [rCOT] wanneer je klaar zal z met zingen		-jyamisha [tw, cau -jyama] stevig vastzetten
-jikila [tw, app -jika; vgl -jaalama 3] verrassen; betrappen		
-jila [tw] 1 iemand 't woord niet richten; 2 vasten, zich onthouden v sommige spijzen		
-jimá [adj] heel, gans, volledig; <i>diba di-</i> [ud] <i>gans uur</i>		
-jima I [i/tww] (uit)doven; (uit)blussen; uitgaan; afzetten, uitzetten; II [adj < I] doven; uitblussen		
> -jimaja; -jimaina; <i>ujimá</i>		
-jimija [tw, cau -jima] verliezen; uitdoven; doen verdwijnen; gans uitwissen		
-jimina [iww, app -jima] verdwijnen; (uit zichzelf) (uit)doven; verloren g; [fig] sterven		
-jinga [tw; vgl -nanga I, -swa/á 1, 2; cf <i>spw13</i>] wensen; verlangen		
> <i>dijinga</i>		
-jinga [tw] oprollen; opwinden		
> -jingila		
		K
		<i>ka-</i> [NP12] <i>kakesé</i> [bw] 'n beetje; <i>kapyá</i> [12/13] vuur
		<i>ka-</i> [BP P] komen; gaan; <i>kafilá</i> [IMP] <i>ga brengen!</i> ;
		<i>kanúmoné</i> [POT] komen jullie eens zien
		<i>ka-</i> 1 [neg prt bij KL2 tot KL18 = <i>kaVP-</i>] niet; 2 [OP1 ENA] <i>katwána wij z niet</i>
		<i>ká-</i> I [NP12] <i>kábidi</i> [bw] opnieuw; <i>kálasá</i> [12/13] <i>klas</i> ; <i>kárcé</i> [12/4] <i>wijk</i> ; II [VP12] <i>kabéji kádi pa méésá</i> [PDI] 't papiertje ligt op tafel; III [' PPI2] <i>kabéji kádi pa méésá</i> [rPDI] 't papiertje dat op tafel ligt
		<i>ká-</i> [OP1] niet; <i>káákakula</i> [nAKA] z/hij sprak niet; <DOCH bij ENA is 't <i>ka-</i> >
		- <i>ka-</i> [BI] gaan (om te); <i>kukafila</i> [INF] <i>gaan geven</i> ; <i>ndi mukafúnda</i> [PDI PVD] <i>iemand is me gaan aanklagen</i>
		- <i>ka/á-</i> [PO (rENA)] niet; <i>bá-keená</i> [rENA] <i>zij die niet zijn</i> ; <i>ukáána</i> [rENA] <i>z/hij die niet is</i> ; <i>únkámányi</i> [nREL] <i>diegene die ik niet ken</i>

WOORDENBOEKJE LUBÀ QUESTIONNAIRE

Nat voor het kerstverlof ontving u een *Lexicon Cluhà-Nederlands*. Hopelijk heeft u het zo nu en dan kunnen gebruiken. Gezien dit woordenboekje kadert in mijn licentiaatsverhandeling had ik nu graag wat feedback van u-ontvangers. Het enige wat u daarvoor hoeft te doen is enkele minutjes vrijmaken om op onderstaande vragen te antwoorden. Hier wordt naar UW MENING gevraagd; er zijn géén juiste noch foute antwoorden! Mijn opinie doet hoegenaamd niets ter zake. Geef mij – uw medestudent – er dus maar gerust van langs.

Deze questionnaire heeft niets met het examen Lubà te maken: hij is NAAMLOOS en de gegevens zullen door mij enkel statistisch worden verwerkt. Zelfs indien dit woordenboekje juist reeds heeft wat van mijn tijd en energie in beslag nam, blijft het zonder uw hulp waarde-loos. Enkel uw mening kan mij helpen een goede theorie te ontwerpen voor mijn thesis. Sterker: enkel door met u met mening rekening te houden kan de volgende uitgave gebruiksvriendelijk(er) worden. Alvast bedankt voor uw medewerking, Gilles-Maurice.

ALGEMEENHEDEN

1. Maakte u gebruik van het <i>Lexicon</i> ?	<input type="checkbox"/> ja [spring naar vraag 3.]	<input type="checkbox"/> nee	
2. Waarom niet?	<input type="checkbox"/> u ontving géén <i>Lexicon</i> [spring naar vraag 44.]	<input type="checkbox"/> u lapte(e) Cluhà aan uw laars [spring naar vraag 46.]	
	<input type="checkbox"/> u vond het <i>Lexicon</i> overbodig [spring naar vraag 44.]	<input type="checkbox"/> andere [spring naar vraag 44.]	
3. Waarvoor gebruikte u het <i>Lexicon</i> het meest? (geefte een '1' te zetten voor de meest gebruikte reden, een '2' voor de tweede meest gebruikte, een '3' etc.)	... om de schrijfwijze van een woord na te gaan	... om de tonen van een woord te controleren	
	... om de betekenis van een woord te achterhalen	... om een vertaal-equivalent ter beschikking te hebben	
	... om één van de tabellen te raadplegen	... om woordenschat te blokkeren	
	andere		
4. Hoe vaak vond u de informatie naar dewelke u op zoek was?	<input type="checkbox"/> altijd	<input type="checkbox"/> zelden	
	<input type="checkbox"/> vaak	<input type="checkbox"/> nooit [spring naar vraag 6.]	
	<input type="checkbox"/> soms	<input type="checkbox"/> andere	
5. Duurde het lang vooraleer u vond wat u zocht?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
a. in het algemeen	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
b. wanneer u moest kiezen tussen gelijkgespelde trefwoorden	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
c. wanneer een artikel met Romeinse cijfers was ingedeeld	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
d. wanneer een artikel met Arabische cijfers was ingedeeld	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
e. wanneer een artikel veel (pun)komma's bevatte	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
f. andere	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere

LAY-OUT

6. Het <i>Lexicon</i> bestaat uit drie delen: voorbeschouwingen, de woordenlijst zelf, en nabeschouwingen. Was u zich daarvan bewust?	<input type="checkbox"/> ja, en ik wist in welk deel ik wat moest zoeken	<input type="checkbox"/> ja, maar ik wist niet goed hoe met de delen om te gaan	<input type="checkbox"/> nee	
7. Wat las u uit de voorbeschouwingen, en hoe?	<input type="checkbox"/> grondig	<input type="checkbox"/> vluchtig	<input type="checkbox"/> niet	<input type="checkbox"/> andere
a. Woord vooraf	<input type="checkbox"/> grondig	<input type="checkbox"/> vluchtig	<input type="checkbox"/> niet	<input type="checkbox"/> andere
b. Gebruiksaanwijzing	<input type="checkbox"/> grondig	<input type="checkbox"/> vluchtig	<input type="checkbox"/> niet	<input type="checkbox"/> andere
c. Mini-grammatica	<input type="checkbox"/> grondig	<input type="checkbox"/> vluchtig	<input type="checkbox"/> niet	<input type="checkbox"/> andere
d. Afkortingen en symbolen	<input type="checkbox"/> grondig	<input type="checkbox"/> vluchtig	<input type="checkbox"/> niet	<input type="checkbox"/> andere
8. Wat gebruikte u uit de nabeschouwingen, en hoe?	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
a. Spreekwoorden	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
b. Kensepreuken	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
c. Tabel affixen	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
d. Tabel genera	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
e. Tabel pronominaalvormen	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
f. Tabel possessieven	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
g. Bibliografie	<input type="checkbox"/> vaak	<input type="checkbox"/> soms	<input type="checkbox"/> nooit	<input type="checkbox"/> andere
9. Wat vindt u van de beschikking van het centrale deel (de woordenlijst zelf)?	<input type="checkbox"/> goed	<input type="checkbox"/> kan beter	<input type="checkbox"/> slecht	<input type="checkbox"/> andere
10. Is het onderscheid 'vet' voor 'Cluhà', en 'gewoon' voor 'Nederlands' duidelijk?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
11. Is het onderscheid 'carrière' voor 'voorbeelden', en 'recht' voor 'al het overige' duidelijk?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
12. Het lettertype van deze questionnaire (Times New Roman) is verschillend van dat in het <i>Lexicon</i> (Courier). Welk lettertype acht u geschikt voor een woordenboek?	<input type="checkbox"/> Times New Roman (zoals deze letters)	<input type="checkbox"/> Courier (zoals in het huidige <i>Lexicon</i>)	<input type="checkbox"/> andere	

AFKORTINGEN EN SYMBOLEN

13. Afkortingen:	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
a. zijn ze over het algemeen duidelijk?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
b. zijn er teveel?	<input type="checkbox"/> welke zijn volstrekt onbegrijpelijk en/of overbodig?		
14. Symbolen:	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
a. zijn ze over het algemeen duidelijk?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
b. zijn er teveel?	<input type="checkbox"/> welke zijn volstrekt onbegrijpelijk en/of overbodig?		
15. In het <i>Lexicon</i> worden woorden binnen een artikel het trefwoord vervangen door een tilde (~). Dit gebeurt ook in het <i>Lexicon</i> . Onder het trefwoord <i>bána</i> vindt men:	<i>bána</i> [1/2] cf ev <i>mawáa</i> ; - <i>AAABáá</i> [1/2] cf ev <i>mawankáá</i> ; ... Welk woord vervangt de tilde (~) hier? <i>AAABáá</i>		
16. In het <i>Lexicon</i> worden de symbolen \emptyset en \square gebruikt.	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere
a. enig idee waar ze eigenlijk voor staan?	<input type="checkbox"/> ja		
b. zijn deze symbolen (\emptyset en \square) storend?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere

GRAMMATICA 'IN' HET LEXICON

17. Affixen (prefixen, suffixen, e.a.) staan zowel getabelariseerd, als op hun alfabetisch plaats in de woordenlijst:	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
a. vindt u dit soort informatie nuttig?	<input type="checkbox"/> tabel	<input type="checkbox"/> woordenlijst	<input type="checkbox"/> beiden	<input type="checkbox"/> geen van beiden
b. waar raadpleegde u dit soort informatie?	<input type="checkbox"/> ja			
18. Sommige combinaties van affixen hebben een vaste betekenis. Zo bijvoorbeeld: ci...-idi/lu = plaats waar men ... Gebruikte u dit soort trefwoorden?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
19. Voorbeelden worden vaak vergezeld van TRINSE informatie. Zo wordt het OI (=OC) -mij- gestuurd m.b.v. een resultaatief. Vindt u deze informatie (hier [RES]) nuttig?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
20. Waar zou u affixen, pronominaalvormen, postposities, e.d. het liefst aantreffen?	<input type="checkbox"/> allen enkel op hun alfabetisch plaats in de woordenlijst			
	<input type="checkbox"/> allen enkel in een tabel in een appendix			
	<input type="checkbox"/> allen zowel in de woordenlijst als in een appendix			
	<input type="checkbox"/> de belangrijkste in de woordenlijst, allen in een appendix			
	<input type="checkbox"/> andere			
21. Waar hoort grammaticale informatie in het algemeen (affixen, tensevormen, deverbaalief-patroon, tense-formules, etc.) volgens u thuis?	<input type="checkbox"/> zoveel mogelijk in de centrale woordenlijst			
	<input type="checkbox"/> zoveel mogelijk sameengebracht in een appendix			
	<input type="checkbox"/> andere			
22. Moet dit <i>Lexicon</i> een beknopte grammatica bevatten?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	

SUBSTANTIEVEN

23. In het <i>Lexicon</i> worden substantieven onder hun enkelvoud ingebracht. Sommige meervouden moesten echter hier en daar voorzien worden (cf. p. xiv, punt 3.). Hoe zou u substantieven het liefst willen opzoeken?	<input type="checkbox"/> uitsluitend als enkelvoud	<input type="checkbox"/> als enkelvoud, met voorziening voor enkele meervouden	<input type="checkbox"/> steeds zowel als enkelvoud als meervoud	<input type="checkbox"/> andere
24. Waar moet men onregelmatige meervouden volgens u kunnen aantreffen?	<input type="checkbox"/> binnen het artikel van hun enkelvoud			
	<input type="checkbox"/> onder hun meervoud (en dus als apart trefwoord)			
	<input type="checkbox"/> andere			
25. In het <i>Lexicon</i> worden de klassenummers van enkelvoud en meervoud gegeven. I.p.v. 'nummers' zouden de 'nominale prefixen' zelf kunnen gebruikt worden. Zijn volgens u:	<input type="checkbox"/> klassenummers handiger dan nominale prefixen			
	<input type="checkbox"/> nominale prefixen handiger dan klassenummers			
	<input type="checkbox"/> zowel klassenummers als nominale prefixen overbodig			
	<input type="checkbox"/> andere			
26. In het <i>Lexicon</i> wordt verondersteld dat uit de klassenummers de morfemen voor concordantie kunnen afgeleid worden. Men zou bij elk substantief het VP (= SC) en OI (= OC) steeds opnieuw kunnen herhalen. Dit zou:	<input type="checkbox"/> nuttig zijn			
	<input type="checkbox"/> overmatig zijn			
	<input type="checkbox"/> andere			
27. Moet een voorbeeldin elk substantief vergezeld zijn?	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> andere	
28. Men zou substantieven ook kunnen inhengen onder hun stem. Zo zou men <i>muntu</i> , <i>buntu</i> , <i>cintu</i> , <i>biutu</i> en <i>kantu</i> dan allen onder <i>-ntu</i> aantreffen. Dit zou:	<input type="checkbox"/> heel gebruiksvriendelijk zijn			
	<input type="checkbox"/> veel te ingewikkeld zijn			
	<input type="checkbox"/> andere			
29. Wat is het meervoud van volgende substantieven: <i>tábiti</i> [1/2a], <i>nzóólo</i> [1/4a], <i>ditámá</i> [3/6], en <i>lumwénu</i> [1/4/4]?	meervouden =			

ADJECTIEVEN

30. Adjectieven gebruiken voor hun concordantie steeds een NP. In het <i>Lexicon</i> werden adjectieven zonder prefix ingebracht. Men had echter ook frequent voorkomende adjectieven met hun prefix kunnen inhengen. Dit zou:	<input type="checkbox"/> nuttig zijn (men zoekt het adjectief gewoon met prefix op)	<input type="checkbox"/> redundant zijn (adjectieven staan meermalen vermeld)	<input type="checkbox"/> andere
--	---	---	---------------------------------------

WERKWOORDEN	
31. In het <i>Lexicon</i> werden werkwoorden zonder een klasseffix ku- ingebracht. In deze werkwijze:	<input type="checkbox"/> nuttig <input type="checkbox"/> verwarrend <input type="checkbox"/> andere
32. Werkwoorden die met een vocaal beginnen (als gevolg van het weglaten van ku-) zien deze vocaal verdubbeld, tenzij men met een verzwaarde positie te maken heeft. Zo vindt men spreken onder -aakula (verdubbeling V), maar zeggen onder -amba (geen verdubbeling V). Is dit:	<input type="checkbox"/> onduidelijk <input type="checkbox"/> duidelijk <input type="checkbox"/> duidelijk doch moeilijk te hanteren <input type="checkbox"/> duidelijk en goed te hanteren <input type="checkbox"/> andere
33. Door het toevoegen van affixingsuffixen aan de kern van een werkwoord maakt men snel een hele resem nieuwe werkwoorden. Waar treft u deze derivaties het liefst aan?	<input type="checkbox"/> in het artikel van de kern zelf van het werkwoord <input type="checkbox"/> alfabetisch geschikt tussen de andere trefwoorden <input type="checkbox"/> andere
34. Na toevoeging van sommige affixingsuffixen (zoals een applicatief -il-, of causatief -ish-) of een reflexiefnifix (-RC) -di- is de betekenis vaak voorspelbaar. Wanneer de betekenis voorspelbaar is moet men volgens u:	<input type="checkbox"/> het werkwoord niet opnemen in het woordenboek <input type="checkbox"/> enkel het bestaan vermelden in het artikel van de kern <input type="checkbox"/> het grondig behandelen binnen het artikel van de kern <input type="checkbox"/> het silees vermelden als trefwoord in de woordenlijst <input type="checkbox"/> het als volwaardig trefwoord opnemen in de woordenlijst <input type="checkbox"/> andere
VERWIJZINGEN	
35. Onder het werkwoord -dyà treft men een 'staart' van trefwoorden aan (bidya; cidili; cyakudya; -dikki; bidwa; -dilla; -disha; mudi; Mudila-mpiku) die in verband staan met de kern -dyà. Elk van die trefwoorden uit de staart bevat een verwijzing terug naar -dyà. In dit alles:	<input type="checkbox"/> onnuttig, want <input type="checkbox"/> nuttig om verbanden te zien <input type="checkbox"/> nuttig om woorden makkelijker te onthouden <input type="checkbox"/> nuttig om te weten wat er wel en niet in het <i>Lexicon</i> staat <input type="checkbox"/> andere
36. Heeft u enig idee waarom van het enkelvoud dila naar het meervoud mala wordt verwezen?	<input type="checkbox"/> ja <input type="checkbox"/> nee
HET VERVOLG	
37. Het <i>Lexicon</i> krijgt zeker een staartje. Aan wat moet volgens u nu eerst aandacht geschonken worden? (gelieve een '1' te zetten voor wat volgens u het belangrijkste is, een '2' voor wat daarna zou moeten behandeld worden, een '3' etc.)	... aan de correctie van fouten in het <i>Lexicon</i> ... aan de toevoeging van extra woorden aan het <i>Lexicon</i> ... aan een lexicon Nederlands-Cilubà ... aan een lexicon Cilubà-Engels ... aan een lexicon Engels-Cilubà ... andere
38. Wat moet er volgens u nog allemaal toegevoegd worden aan het huidige <i>Lexicon</i> .	<input type="checkbox"/> uitspraak van het trefwoord (fonetisch of andere)) <input type="checkbox"/> aanduiding van de gebruiksfrequentie van het trefwoord <input type="checkbox"/> gereconstrueerde proto-Bantoe vorm van het trefwoord <input type="checkbox"/> meer tabellen (welke?) <input type="checkbox"/> beknopte grammatica (hoevel bladzijden?) <input type="checkbox"/> andere
39. Wanneer u in een nieuwe uitgave van het <i>Lexicon</i> een woord zou opzoeken, over welk aspect zou u het snelst informatie wensen aan te treffen?	... woordsoort (of het een werkwoord, substantief, ... is) ... etymologie (uit welke taal het werd geleend) ... van welk woord het trefwoord werd afgeleid ... welke woorden uit het trefwoord worden afgeleid ... welke woorden synoniem zijn met het trefwoord ... variante spelling(en) van het trefwoord ... een vertaling/omschrijving van het trefwoord ... een voorbeeld om het trefwoord te illustreren ... een verstaende uitdrukking met het trefwoord ... andere
40. Het spreekt voor zich dat het <i>Lexicon</i> verre van foutloos is. Kunt u mij op enkele storende foutjes wijzen? (Heel wat Nederlandse woorden werden bijvoorbeeld verkeerd gepluist (zoals werk-en, donk-er, trouw-en, etc.). Dit soort onvergeetlijke fouten werd in het nieuwe manuscript reeds verbeterd. Maar u kent zeker andere ...)
ROUND-UP	
41. Al bij al:	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
a. vond u het <i>Lexicon</i> nuttig?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
b. vond u het <i>Lexicon</i> ingewikkeld in gebruik?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
c. vond u de vertaling goed in het <i>Lexicon</i> ?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
d. vond u de voorbeelden adequaat in het <i>Lexicon</i> ?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
e. vond u genoeg gegevens in het <i>Lexicon</i> ?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
42. Gebruikte u naast het <i>Lexicon Cilubà-Nederlands</i> , andere Cilubà lexica of woordenboeken?	<input type="checkbox"/> ja <input type="checkbox"/> nee [spring naar vraag 46.]

43. Zijn er bepaalde zaken waarvoor u enkel van het <i>Lexicon</i> gebruik maakte, en niet van een andere bron?	<input type="checkbox"/> ja, welke?
44. Welke bron(ne)n gebruikte u het vaakst? (gelieve een '1' te zetten voor de meest gebruikte bron, een '2' voor de tweede meest gebruikte, een '3' etc.)	<input type="checkbox"/> nee ... Dictionnaire Tshi -Fra. (De Clercq & Willems 1960) ... Dictionnaire Fra.-Tshi. (Willems 1986) ... Vocabulaire Tshi.-Fra. et Fra.-Tshi. (Willems 1993) ... <i>Lexicon Cilubà-Nederlands</i> (de Schryver & Kabuta 1997) ... andere
45. Welke bron(ne)n gebruikte u het liefst? (gelieve een '1' te zetten voor de eerste voorkeur, een '2' voor de tweede voorkeur, een '3' etc.)	... Dictionnaire Tshi.-Fra. (De Clercq & Willems 1960) ... Dictionnaire Fra.-Tshi. (Willems 1986) ... Vocabulaire Tshi.-Fra. et Fra.-Tshi. (Willems 1993) ... <i>Lexicon Cilubà-Nederlands</i> (de Schryver & Kabuta 1997) ... andere
46. Maakt u vaak gebruik van woordenboeken?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
a. voor taalvakken aan de RUG (Lubà/Swahili/Arabisch/...)	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
b. tijdens het lezen van syllabi/papers in het Engels/Frans/...	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
c. andere	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
47. Vindt u van uzelf dat u goed overweg kunt met woordenboeken?	<input type="checkbox"/> ja <input type="checkbox"/> nee <input type="checkbox"/> andere
48. Met welk niveau zou u uw huidige kennis van het Cilubà omschrijven?	<input type="checkbox"/> elementair niveau <input type="checkbox"/> gevorderd niveau <input type="checkbox"/> tussenliggend niveau <input type="checkbox"/> andere
49. Een lexicon Cilubà-Engels/ Engels-Cilubà i.p.v. een lexicon Cilubà-Nederlands / Nederlands-Cilubà zou:	<input type="checkbox"/> even goed zijn <input type="checkbox"/> minder bruikbaar zijn <input type="checkbox"/> andere
50. Is er nog iets dat u kwijt wilt? Laat u gerust gaan

Hartelijk dank voor uw medewerking en veel succes met uw examens!

P.S. Laat uw antwoorden niet slingeren: maak gebruik van de voorgedrukte envelop om ze mij zo snel mogelijk te laten geworden.

cyondò	Behept woordenboek Ghùbà-Nederlands	22	Ketebesohi-rufùfùti kàl Ghùbà	-di
--------	--	----	----------------------------------	-----

cyondò [7/8] trom <om berichten door te sturen; vervaardigd uit 'n uitgeholde boomstam>; **kwela** ~ [ud] *op de boom spelen*
cyongo (**wa minanga**) [7/8] 1 januari; 2 maanmaand volgend op de **ciswà**-(**munène**) <korte droge periode rond december-januari>
cyonò [7/8] gesnurk; **kwela byonò** [ud] *snurken*
cyónso [al7] elk
cyowedi [7/8 < app -oowa'] bad-plaats; badkamer
cyulà [7/8] pad <dier>
cyùma [7/8] rijkdom; waardevol voorwerp

▷ byùma

Cyunza [mpn] Cyunza

D

dèsèr [1/4b < Fr "dessert"] nage-recht
di- [NP5] **dici** [5/6] *oor*; **diskè** [5/4] *plaat/platen*; **dyákabi** [5/Ø] *mislukking*
□ **di**...-a/a [NP5 en FN a/a ⇒ GER] het; **didya** [5/Ø] *het eten*; **difla** [5/Ø] *het geven*; **dijà** [5/Ø] *het dansen*
di/i- [RP5] <tjd>; **diwàávwà** [rAT] *de dag waarop je zult le*
di- I [NP5] **dikalù** [5/6] *fiets*
II [SC5] **dingeyà didi pa mèèsà** [PDI] *de mango ligt op tafel*

III [PP5] **dibà dyàsùmbuyè** [rRES] *'t uurwerk dat z/hij kocht*
-di-¹ [OC5] **bwejà dikalù! nkàadi mudibwèje** [IMP; PDI AND] *breng de fiets binnen! ik heb die reeds binnen-gebracht*
-di-² [reflexiefconcordant] zich(zelf); **kudifila** [INF] *zich geven/wijden*
◊-di [dww PDI; cf spw5, 7, 12, 15] zijn; **bàdi zè zjin**; **udi je bent**;

- ZIJN **nè** [ud] hebben; **ùdi nè mudimu z/hij heeft werk**
□ ZIJN **nè bwà/cyà** [ud] moeten; **bàdi nè bwà kumanya zè moeten weten**; **ndi nè cyà kuya ik moet vertrekken**
□ ZIJN **mwà** [ud] kunnen <potentieel>; mogen <toelattend>; **kéna mwà kupátuka z/hij kan/mag niet uitgaan**
□ ZIJN **munkaci mwà** [ud] aan 't <progressief>
□ ZIJN -**kaa/kàà-pa** [ud] weldra <imminent>; **wámánya, mulongeshi úkaadi pa kulwa** [AT, PDI INF] *opgelet, de lergever staat op 't punt te le*
□ **neg(ZIJN)+ND(-anji)+INF** [bw] nog niet; **kávwa mwanji kuledí-bwa z/hij was nog niet geboren**
□ ZIJN-**ku/ù** [SS17, PFa] <existentiele betekenis>; **nsátánà ùdikù anyi? bestaat de duivel?**

-di/i	Behept woordenboek Ghùbà-Nederlands	23	Ketebesohi-rufùfùti kàl Ghùbà	diibidi
-------	--	----	----------------------------------	---------

◊-di/i [dww rPDI; cf spw1, 2, 7, 12, 17, ksp3] die is/zijn; **bàdi diegenen die zjin**; **udi diegene die is**
◊-di voor constructies met ZIJN
◊**dib**. [afk v **dibèji**] bladzijde
◊**diba** [5/6] 1 zon; 2 tjd; uur; horloge, uurwerk; **dibà dyà ku cimanu wandeloke**; **dibà dyà ku dibòko polshorloge**; **dibà dyà ngòlù gouden uurwerk**; **dibà dyà nzembu elektrisch uurwerk**; ~ **dijimà** [ud] 'n gans uur; **màkàzè àà mèba uurwerken-winkel**
◊ **dibà dyà ...** [ud] gedurende ...
◊ **~ kaayi** [vud] wanneer; hoe laat; **tudi ~ kaayi?** [vud] *hoe laat is 't?*
dibadika [5/Ø; GER < imp -bala²] (volks)telling
◊dibàkà [5/6 < -bàka] huwelijk; **fetè wa ~ huwelijksfeest**; **kwenza ~** [ud] *trouwen*
dibànzà [5/6] schuld; **kufuta mabànzà** [ud] *schulden betalen*
dibata' [5/6; syn mpaatu] eend
dibata' [5/6] vuist
dibaya [5/6; vgl tàblò] (school)bord; plank; hout
dibàyà [5/6] vrouwenbloes
dibèèle [5/6; vgl cyàdi] vrouwenborst; **kwamwa ~** [ud] *de borst nemen*; **mutù wà ~ tepel**
▷ **mabèèle**
◊dibèji [5/6] blad <v boom of boek>; (geld)briefje; biljet; blad-zijde [afk **dib**.]
▷ **cibèjèji**; **kabèji**

-**dibèya** [rww < -bèya] zich scheren
◊dibòko [5/6] arm; **dibà dyà ku ~ polshorloge**
dibùki [5/6] pak(je)
dibungi [5/6] mist
dibùulà [5/6] bloembed, smal tuin-bed; boogje (zand op 'n akker)
dibwe [5/6] 1 steen: rot; 2 parcl; 3 batterij; ~ **dyà nzembu batterij**
▷ **kabwe**(**bwe**(**bwe**))
dibwè [5/6; cf spw24] palmboom; **maalà àà ~ palmwijn**
dici [5/6] oor
dididiishi [5/6 < Ks "dirisha"] venster; vensterluik
didimà [5/6 < Prt "limão"; vgl **cimuma cyà mpùsù, dilalaà**] citrusvrucht <citroen; sinaasappel>
difi [5/6] leugen; **mwena mafi** [ud] *leugenaar*
difikiltè [1/4b en x/6 < Fr "difficulté"] moeilijkheid
◊difù [5/6] buik; **kusaama ~** [ud] *buikpijn h*; **kwimita (~)** [ud] *zwanger w*
difùku [5/6] cf var **ditùku**
-**difùndisha** [rww < cau -fùnda] inschrijven
difutu [5/6 < -futa] loon, salaris; **kuvudijila ~** [ud] *opslag geven*
dii- [NP5] **diikùmi** [tw 5/6] tien; **diiskè** [5/4] *plaat/platen*
diibaka [5/Ø; GER < -iibaka] op-richting
diibidi [5/6 < -ibidi] dinsdag

Affixen	1pe	1pm	2pe	2pm	1	2	3	4	5	6	7	8	11	12	13	14	15	16	17	18	
NF	tsu	ba	tsu	ba	mu	ba	mu	tsu	di	ma	ci	bi	lu	ka	ru	bu	ku	pa	ku	mu	
SC	tsu	ba	tsu	ba	mu	ba	mu	tsu	di	ma	ci	bi	lu	ka	ru	bu	ku	pa	ku	mu	
OC	tsu	ba	tsu	ba	mu	ba	mu	tsu	di	ma	ci	bi	lu	ka	ru	bu	ku	pa	ku	mu	
PP	u	ba	u	ba	u	ba	u	i	di	i	ci	bi	lu	ka	ru	bu	ku	pa	ku	mu	
OM	ci	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku	ku

Afmerkingen & Symbolen	FN final	AKA ver-voerden met -ika	INF infinitief	app applicatief / Zamenlatief
① 200 frequentie lenuta	NP nominalprefix	AT à tense	ND nominal deelwoord	asa associatief
② 400 frequentie lenuta	SC subject-concordant	ENA negatief voor -na	PDI present met -di	◊ vrentening
③ 600 frequentie lenuta	OC object-concordant	FUT futurum	RES resultatief	□ zamenlatief
g vorm v/h ww "gaan"	PP pronominaalprefix	GER gerundum	VD verbaal deelwoord	◊ afgeleid uit
h vorm v/h ww "komen"	OM ontkenningsoorform	IMP imperatief	VVA praeterium met -vwa	ud uitdrukking
				✱ sandachi

Adressierung in der ein- und zweisprachigen Lexikographie. Eine einführende Übersicht über die Forschungs- und Problemlage

Herbert Ernst Wiegand, *Germanistisches Seminar, Universität Heidelberg,
Heidelberg, Bundesrepublik Deutschland*

Zusammenfassung: Obwohl die Adressierungsbeziehung eine wichtige textuelle Beziehung insbesondere in Wörterbuchartikeln ist, ohne deren theoretisches Verständnis die Gewinnung von lexikographischen Informationen anhand von hochverdichteten und stark standardisierten Wörterbuchartikeln nicht erklärt werden kann, gibt es bisher nur wenige Beiträge, in denen die Adressierung theoriebezogen behandelt wurde.

Zunächst werden einige theoretische Voraussetzungen aus einer Theorie der Wörterbuchform eingeführt. Erläutert wird *erstens* die Standardisierung an Beispielen zur lexikographischen Beschreibungssprache, anhand der inneren Textverdichtung, an Beispielen zum Wörterbuchgegenstand und bezüglich der textuellen Strukturen von Wörterbuchartikeln. Daraufhin wird *zweitens* gezeigt, wie man sich den Prozeß der Textverdichtung als Textverarbeitungsprozeß denken kann, der von einem Volltext zu einem verdichteten Artikeltext führt; es wird deutlich, daß hierbei die syntaktische Struktur des Volltextes weitgehend getilgt wird, so daß im Artikeltext keine Kohäsionsanweisungen und syntaktischen Funktionen erkennbar sind. Damit die syntaktischen Konstituenten (die Angaben) aufeinander bezogen werden können, müssen die Angaben an andere Angaben adressiert sein. Adressierungsbeziehungen sind daher der Ersatz für die syntaktischen Funktionen und kohäsiven Bezüge. Schließlich werden *drittens* Textkonstituentenstrukturen eingeführt, und zwar wird insbesondere dargelegt, wie Mikrostrukturen zu analysieren sind.

Auf der Basis der eingeführten Voraussetzungen werden dann zuerst einsprachige Wörterbuchartikel hinsichtlich ihrer Adressierungsstrukturen exemplarisch untersucht. Dabei werden verschiedene Arten der artikelinternen Adressierung unterschieden, z.B. lemmatische, vollständig lemmatische, sublemmatische, nichtlemmatische, standardisierte, nichtstandardisierte Adressierung sowie Links- und Rechtsadressierung und einige weitere Arten. Weiterhin wird dargelegt, daß die Adressierungsrelation in Wörterbuchartikeln mit vollständiger lemmatischer Adressierung eine Funktion ist und wie die Angabestruktur eines Wörterbuchartikels als Zusammenfügung von Mikro- und Adressierungsstruktur erklärt und dargestellt werden kann.

Schließlich wird im letzten Abschnitt auf Aspekte der beiden wichtigsten Adressierungsprobleme in zweisprachigen Wörterbuchartikeln eingegangen. Diese entstehen (i) dadurch, daß die Äquivalentbeziehung meistens eine bedingte Beziehung ist, die nur gilt, wenn andere Sachverhalte gelten, und (ii) weiterhin dadurch, daß es zu einem ausgangssprachlichen Lemmazeichen mehrere zielsprachliche Äquivalente geben kann. Dadurch werden Äquivalentunterscheidungsangaben

nötig, mit denen komplexe Adressierungen notwendig werden. Es wird vorgeschlagen, diese als lexikographisch vertextete *wenn-dann*-Beziehungen zu verstehen und entsprechend in den Benutzungshinweisen zu erklären.

Stichwörter: ADRESSIERUNG, ADRESSE, ANGABE, WÖRTERBUCHARTIKEL, STANDARDISIERUNG, TEXTVERDICHTUNG, MIKROSTRUKTUR, ADRESSIERUNGSSTRUKTUR, ANGABESTRUKTUR

Abstract: Addressing in Mono- and Bilingual Lexicography. An Introductory Survey of the State of Research and Problematics. The addressing relationship is an important textual relationship, especially in dictionary articles. Without the theoretical understanding thereof, lexicographical information obtained from highly condensed and greatly standardised dictionary articles cannot be interpreted. Up to now, however, there have only been a few contributions in which addressing has been theoretically treated.

To begin with, some theoretical prerequisites from a theory of dictionary form is introduced. Firstly, standardisation is explained using examples from lexicographic descriptive language, referring to internal textual condensation, using examples from the dictionary object and referring to the textual structures of dictionary articles. Secondly, it is shown how the process of textual condensation as text processing procedure can be conceived of as leading from the full text to the condensed text of the article. It becomes clear that in this process the syntactic structure of the full text is considerably reduced, so that no indicators of cohesion and syntactic functions are recognisable in the text of the article. Items should be addressed to other items so that the syntactic constituents (the items) can be referred to one another. Addressing relationships are therefore substitutes for syntactic functions and cohesive relations. Finally, the text constituting structures are introduced, and it is explained how microstructures can be analysed.

On the basis of the introduced prerequisites, the addressing structures of monolingual dictionary articles are then firstly examined by way of example. In this process different types of article internal addressing are distinguished, e.g. lemmatic, full lemmatic, sublemmatic, non-lemmatic, standardised, nonstandardised addressing, as well as left and right addressing and several other types. It is further explained that the addressing relationship in dictionary articles with full lemmatic addressing is a function and how the entry structure of a dictionary article can be explained and presented as a junction of the micro- and addressing structure.

In the last section aspects of the two most important addressing problems in bilingual dictionaries are discussed. These originate (i) from the fact that the equivalence relationship is usually a qualified relationship which is only valid when other conditions apply, and (ii) also from the fact that one source language lemma sign may have more than one target language equivalent. This makes equivalent discrimination entries necessary, which necessitates complex addressing. It is suggested that these should be understood as lexicographically textualised if-then relationships and be explained as such in the usage notes.

Keywords: ADDRESSING, ADDRESS, ITEM, DICTIONARY ARTICLE, STANDARDISATION, TEXT CONDENSATION, MICROSTRUCTURE, ADDRESSING STRUCTURE, ITEM STRUCTURE

1. Zum Forschungsstand und zum theoretischen Rahmen

Die Adressierungsbeziehung ist eine der wichtigsten textuellen Beziehungen in lexicographischen Texten mit einem Leitelementträger, insbesondere in hochverdichteten und stark standardisierten Wörterbuchartikeln in gedruckten ein- und zweisprachigen Sprachwörterbüchern, und zwar sowohl in tabellarischen Wörterbuchartikeln (i.S.v. Wiegand 2000d) als auch in nichttabellarischen. Grob gesprochen, wird durch Adressierungsbeziehungen der Bezug von lexicographischen Daten zu anderen lexicographischen Daten und damit zum Wörterbuchgegenstand (i.S.v. Wiegand 1998: 302, D 3-5) geregelt. Nur ein Benutzer, der in der Lage ist, z.B. eine Wörterbuchangabe (kurz: Angabe) korrekt auf eine andere Angabe zu beziehen, kann anhand der Angaben eines Wörterbuchartikels lexicographische Informationen (als kognitive Gegebenheiten) erschließen.

Trotz der zentralen Bedeutung der Adressierungsbeziehungen für das Verständnis von Teilen der Wörterbuchform und der Informationsgewinnung bei der Wörterbuchbenutzung gibt es bisher nur relativ wenige Arbeiten zur Adressierung. Die Feststellung von Hausmann und Wiegand (1989: 353): „Systematic studies of the addressing practices of dictionaries all over the world are a desideratum of dictionary research for several reasons“ gilt auch heute noch. Vor dem Erscheinen des dreibändigen Handbuches zur Lexikographie (HSK 5.1-5.3) in den Jahren 1989-1991 wurde die Adressierung m.W. nicht zusammenhängend in einer eigenen Publikation thematisiert. Auch in den wichtigsten Monographien, die vor HSK 5.1-5.3 erschienen sind, wie z.B. Rey-Debove (1971), Zgusta (1971), Landau (1984), Wolski (1986) u.a., bzw. in solchen, die vorher abgeschlossen waren, wie z.B. Svensén (1993), spielt die Adressierung keine Rolle. Selbst in Übersichtswerken nach 1991, wie z.B. in Bergenholtz und Tarp (1995), wird die Adressierung nicht behandelt.

Für die Berücksichtigung der Terminologie im Bereich der Adressierung in den Wörterbüchern zur Lexikographie gilt folgendes: In Martínez de Sousa (1995) findet sich kein Lemma, das als spanisches Äquivalent zu *Adresse* oder *Adressierung* gelten kann. Auch in Burkhanov (1998) ist die Adressierung nicht berücksichtigt. In Bergenholtz et al. (1997) findet man folgende Lemmata: **adresse** (mit der Verweisangabe „→ adressering, angivelse, bearbejdingsenhet“), **adressering** (mit der Verweisangabe „→ adresse“) und **adresseringsstruktur** (mit der Verweisangabe „→ bearbejdingsenhet“). Die in Bergenholtz et al. (1997) begonnene lexicographische Aufarbeitung der neueren metalexikographischen Terminologie im Bereich der Adressierungsphänomene wird in Hartmann und James (1998) nicht angemessen fortgesetzt; hier wurde nur das Lemma **address** angesetzt.¹ In den neueren linguistischen Wörterbüchern ist die metalexikographische Terminologie notorisch unterrepräsentiert; in Glück (2000) als dem deutschen sprachwissenschaftlichen Wörterbuch, das die meisten metalexikographischen Termini primär gebucht hat, findet man zwar einen Verweisartikel „Adresse ↗ Lemma“, unter dem Lemma **Lemma** steht aber nichts zu Adressen oder zur Adressierung.

Seit der expliziten Behandlung der Adressierung in Hausmann und Wiegand (1989) für die einsprachige, in Hausmann/Werner (1991) für die zweisprachige Lexikographie sowie in Wiegand (1989a) im systematischen Zusammenhang mit anderen textuellen Strukturen von Wörterbuchartikeln wurden Probleme der Adressierung nur relativ selten angesprochen. Einige Ausführungen finden sich in Wiegand (1990: 97-107) und neuerdings in Meyer und Wiegand (2000) sowie in Wiegand (2000e). Der Grund für die Zurückhaltung der Forschung bei der Bearbeitung von Fragen der Adressierung ist möglicherweise mit folgendem Zitat angesprochen: „A theoretical well-founded treatment of the addressing in the general monolingual dictionary belongs to the most difficult fields of an explicit metalexikographic description“ (Wiegand 1990: 97). Für die zweisprachige Lexikographie ist eine explizite Darstellung der Adressierungsbeziehungen noch erheblich komplizierter (vgl. Wiegand 2000e); es muß m.E. derzeit sogar als offen gelten, ob die Adressierungsbeziehungen in zweisprachigen Wörterbuchartikeln in der bisherigen Forschung überhaupt hinreichend genau verstanden sind.

Eine separate ad-hoc-Behandlung der Adressierungsproblematik ohne Anbindung an eine Wörterbuchtheorie ist wenig nützlich. Im folgenden wird daher an eine weitgehend ausgearbeitete Theorie lexikographischer Textträger angeschlossen.² Diese Theorie besteht aus zwei Teiltheorien: einer Theorie des Wörterbuchgegenstandes und einer Theorie der Wörterbuchform (vgl. Wiegand 1998: 7ff.). Im Rahmen der letztgenannten Theorie wird auch die Adressierung behandelt. Die Theorie der Wörterbuchform bildet auch für diesen Beitrag den engeren Argumentationsrahmen. Da ich eine detaillierte Kenntnis dieser Teiltheorie einer allgemeinen Theorie der Lexikographie nicht voraussetzen möchte, werden im nächsten Abschnitt zuerst einige Begrifflichkeiten informell eingeführt, die hier benötigt werden, um eine zwar einführende, aber dennoch theorieorientierte Darstellung der Adressierung geben zu können. Es sei ausdrücklich darauf hingewiesen, daß die einführenden Ausführungen im nächsten Abschnitt relativ grobmaschig sind, so daß für ein vertieftes Verständnis die jeweils genannte Literatur zu berücksichtigen ist.

2. Einige theoretische Voraussetzungen

In diesem Abschnitt werden folgende Theoriebereiche berücksichtigt, allerdings nur insoweit, als sie die Wörterbuchartikel in gedruckten Sprachwörterbüchern betreffen:

- die Standardisierung,
- die Textverdichtung, und
- die Textkonstituentenstrukturen von Wörterbuchartikeln mit Schwerpunkt auf den Mikrostrukturen.

2.1 Zur Standardisierung

Eine Übersicht über die Möglichkeiten und Probleme der lexikographischen Standardisierung ist Wiegand (1997). Die Standardisierung in der Lexikographie erfolgt aufgrund von Standardisierungsinstruktionen. Diese beziehen sich auf die lexikographischen Handlungen im lexikographischen Prozeß. Sie legen fest, daß die Handlungen in bestimmter Weise anhand von Mustern, vorgegebenen Schemata, nach bestimmten Standards etc. ausgeführt werden. So werden beispielsweise Wörterbuchartikel nach vorgegebenen Musterartikeln verfaßt. Bei Wörterbuchartikeln bezieht sich die Standardisierung vor allem auf folgende artikelinterne Phänomene:

- (a) die lexikographische Beschreibungssprache,
- (b) die lexikographische Textverdichtung,
- (c) den Artikelgegenstand als Teil des Wörterbuchgegenstandes, und
- (d) die textuellen Strukturen und Architekturen des Artikels.

Zu (a): Die Standardisierung der Beschreibungssprache betrifft u.a. die Ergebnisse der Textverdichtung, z.B. einheitliche Abkürzungen linguistischer Termini oder einheitliche Ersetzung der Lemmazeichen in Beispielangaben (nicht einmal durch eine Tilde „~“, ein anderes mal durch einen Divis „~“) u.v.a.m. Häufig wird in die Beschreibungssprache an kulturspezifische Standardisierungstraditionen in der Sprachwissenschaft angeschlossen. Beispielsweise steht in etymologischen Wörterbüchern der Asterisk mit der Bedeutung von *erschlossene Form* und das Zeichen „<“ in der Bedeutung von *seit* (z.B. Band n. (< 8. Jh.). Mhd. *bant*, ahd. *bant*, as. *band* aus g. **banda-* n. [...]; Kluge 1995: 77).

Auch zahlreiche wörterbuchspezifische Besonderheiten der lexikographischen Beschreibungssprache können standardisiert sein. Beispiele dafür finden sich in dem BW-Artikel in Abb. 1.

Abb. 1: wa₁ aus BW.

'an|bin|den (V. 111) 1(500) j m d n. od. etwas ~ mit *Faden, Schnur usw. befestigen*: binden Sie erst den Hund an!; Rosen am Spalier ~; ein Boot an einen Pfahl ~; ein Segel ~ 1.1 sich nicht ~ lassen (fig.) *seine Freiheit bewahren* 2(800) mit j m d m. ~ (umg.) *Streit beginnen. Händel suchen* 3 → a. *angebunden*

Die dreistelligen Zahlen „111“, „500“ und „800“ in wa₁ sind BW-spezifisch standardisierte Elemente der lexikographischen Beschreibungssprache dieses allgemeinen einsprachigen Wörterbuchs. Nur wer ein kundiger Benutzer des BW (i.S.v. Wiegand 1998: 506f.) ist, so daß er die Metatexte des BW gut kennt, weiß etwas mit diesen Zahlen anzufangen. Was er wissen muß, wird im folgenden theoriesprachlich ausgedrückt (also nicht so wie ein Benutzer über das Wissen verfügt bzw. es exteriorisieren kann).

Ein Benutzer muß wissen, daß z.B. „500“ eine verdichtete Verweisangabe (v.VerwA) ist, die nur aus der elementaren Verweisadressenangabe (Verw-AdA) besteht, weil die Verweisbeziehungsangabe (VerwBA), realisiert z.B. als „↑“ oder „s.“ (= siehe), im Zuge der Textverdichtung weggelassen wurde. Der Benutzer muß weiterhin den Adressenfundort kennen, d.h.: er muß wissen, wo er die mit der Angabe „500“ erwähnte Verweisadresse innerhalb des Wörterbuches suchen muß, nämlich im Vorspann des BW, und zwar als Element einer numerischen äußeren Zugriffsstruktur zu einer Tabelle, in der die Satzmuster für Verben aufgelistet sind. Nach Ausführung einer Verweispflichthandlung (i.S.v. Wiegand 1998: 413) kann er die Verweisadresse „500“ (auf S. 26) erreichen und unter dieser Umtextadresse das Verweisziel, nämlich folgende Satzmusterangabe finden: „S + Vb + Akko“ (= Subjekt + finites Verb + Akkusativobjekt). Weiterhin muß der Benutzer aber auch wissen, daß er die Satzmusterangabe auf das Lemmazeichen von w_{a_1} , also auf *anbinden* beziehen muß; denn es handelt sich ja um eine Angabe, mit der auch ein Satzmuster für *anbinden* gegeben wird, und gerade dieser Sachverhalt war ja der Zweck der Verweisung. Woher aber kann der Benutzer wissen, daß er den genannten Bezug herstellen muß? Er kann dies nur dann wissen, wenn er weiß, daß die verdichtete Verweisangabe „500“ in w_{a_1} auch an die Lemmzeichengestaltung (LZGA) adressiert ist. Damit haben wir — sozusagen en passant — ein erstes Ergebnis zur Adressierung von Verweisangaben; es gilt nämlich: Jede artikelinterne Verweisangabe ist *erstens* explizit an mindestens eine Verweisadresse adressiert, die in der Verweisadressenangabe (als einer Teilangabe der Verweisangabe) erwähnt ist, und *zweitens* an eine andere artikelinterne Angabe, anhand derer (oder: mit deren Hilfe) die lexikographische Information, die anhand des Verweisziels ermittelt werden kann, vom Benutzer auf den Adressaten zu beziehen ist. Verweisangaben sind also stets mindestens zweifach adressiert. Nur dadurch können sie ihre Funktion erfüllen, dem Benutzer als Brücke zwischen zwei Textsegmenten zu dienen, die aufeinander zu beziehen sind, so daß eine lexikographische Information entstehen kann.

Da auf die lexikographische Textverdichtung unter 2.2 einzugehen ist, kann nachfolgend gleich auf einige Aspekte der Standardisierung des Gegenstands von Wörterbuchartikeln eingegangen werden.

Zu (c): Nach Wiegand (1998: 302, D 3-5) ist der Wörterbuchgegenstand eines bestimmten Wörterbuchs die Menge der in diesem Wörterbuch lexikographisch bearbeiteten Eigenschaftsausprägungen von einer oder mehreren sprachlichen Eigenschaften bei einer bestimmten Menge von im Wörterbuch erwähnten sprachlichen Ausdrücken, die zu einem bestimmten Wörterbuchgegenstandsbereich gehören. Bei einem Aussprachewörterbuch der neuhochdeutschen Standardsprache ist die sprachliche Eigenschaft die Aussprache und der Wörterbuchgegenstandsbereich die neuhochdeutsche Standardsprache. Der Wörterbuchgegenstand ist dann die Ausprägung dieser Eigenschaft ‚Aussprache‘ bei jedem einzelnen neuhochdeutschen Lemmzeichen, also die jewei-

lige Lemmazeichenausprache. Wird der Wörterbuchgegenstand eines Wörterbuchs (wie z.B. bei Aussprache- und Rechtschreibwörterbüchern) gerade durch *eine* Eigenschaft festgelegt, liegt ein *monoinformatives* Wörterbuch vor. Geschieht die Festlegung des Wörterbuchgegenstands durch *mehrere* sprachliche Eigenschaften (wie z.B. bei allgemeinen einsprachigen Wörterbüchern) liegt ein *polyinformatives* Wörterbuch vor. Mit der Festlegung der Eigenschaften, die zu bearbeiten sind, ist noch keine Standardisierung gegeben. Vielmehr ist diese erst dann etabliert, wenn relativ zu Lemmazeichentypen genau festgelegt ist, welche der Angabeklassen stets zu berücksichtigen sind und welche nicht berücksichtigt werden dürfen. Es ist klar, daß eine strikte Standardisierung bezüglich des Wörterbuchgegenstandes nicht bei allen Wörterbuchtypen erwünscht ist.

Zu (d): Die Standardisierung der textuellen Strukturen und Architekturen eines Wörterbuches ist durch die Standardisierung des Wörterbuchgegenstandes in bestimmter Weise präformiert. Denn durch die Festlegung der Angabeklassen pro Lemmazeichentyp ist zugleich festgelegt, welche Angabeklassen und wieviele zur Trägermenge einer abstrakten hierarchischen Mikrostruktur eines Artikeltyps gehören. Definiert man auf solchen Trägermengen die erforderlichen strukturprägenden Relationen (vgl. 2.3), dann legt man damit die artikelinterne Datendistribution fest, so daß eine textstrukturelle Standardisierung gegeben ist, die sich auf die Mikrostruktur als Teilstruktur der Artikelkonstituentenstruktur bezieht. Auch letztere Struktur kann standardisiert werden; entsprechende Standardisierungsinstruktionen legen dann zusätzlich fest, welche nichttypographischen Mikrostrukturanzeiger (z.B. Kommata, senkrechte Striche; runde/eckige/spitze Klammern u.a.) wie zu setzen sind.

Die Standardisierung der Mikrostruktur determiniert die Standardisierung der Adressierungsstruktur (wie wir noch genauer sehen werden) und damit auch die der Angabestruktur weitgehend.

Die artikelspezifische Standardisierung bezieht sich meistens auch auf die Präsentation der lexikographischen Daten als zweidimensionale Textgestalten und Text-Teilgestalten im zweidimensionalen Druckraum, d.h. standardisiert sind z.B. die artikelinterne Absatzbildung, die Textblockbildung, die Auslagerung bestimmter Angaben auf den Außensteg der Wörterbuchseite u.v.a.m. Anders ausgedrückt heißt das, daß die Mikroarchitekturen (vgl. zu diesen z.B. Wiegand 2000: 259ff.) in ihren verschiedenen Ausprägungen standardisiert sind.

Der Grad der Standardisierung in Wörterbuchartikeln kann unterschiedlich sein. Auch kann in einem Wörterbuch vorgesehen sein, daß bestimmte Bereiche stark, andere weniger stark standardisiert sind. Beispielsweise sind die Formkommentare (fk) im BW stark standardisiert, wie die folgenden vier Beispiele zeigen:

- fk₁: 'Bei-stand <m.; -(e)s; ♂ e>
 fk₂: 'Wein-brand <m.; -(e)s; ♂ e>
 fk₃: 'Spiel-ball <m.; -(e)s; ♂ e>
 fk₄: 'Stief-sohn <m.; -(e)s; ♂ e>

Die Befolgung der Standardisierungsinstruktionen hat bei den Formkommentaren fk₁ bis fk₄ dazu geführt, daß alle linguistischen Aussagen zu den Lemmazeichen, die zur gleichen Aussageklasse gehören (außer den Aussagen zur Rechtschreibung), so lexikographisch vertextet wurden, daß die vier Formkommentare stark standardisiert sind und damit insgesamt untereinander weitestgehend ähnlich, weil sie in folgenden Hinsichten gleich sind:

- Die Anzahl der Angaben (aus derselben Klasse von Angaben mit gleichem allgemeinem genuinem Zweck) ist gleich.
- Die Reihenfolge dieser Angaben ist gleich.
- Die Zuordnung der typographischen Mikrostrukturanzeiger (halbfett und normal) zur Angabeform der elementaren Angaben ist gleich.
- Die nichttypographischen Mikrostrukturanzeiger (spitze Klammer, Semikolon und Komma) sind gleichwertig verwendet und damit nach Position und Funktion gleich.
- Die positionsgebundenen Angabesymbole (freistehender Akzent, Mittelpunkt, Bindestrich und hochgestellter Zweierpunkt) sind gleichartig verwendet und damit nach Position und Funktion gleich.
- Die Anwendung der Kondensierungsmethoden (i.S.v. Wiegand 1996c) ‚Einschachteln‘ (vgl. „-(e)s“), ‚Abkürzen‘ (vgl. „m.“) und ‚Ersetzen‘ (vgl. „-(e)s“ und „-e“) hat zu gleichen Textsegmenten geführt.
- Entsprechendes kann für die Auslassungsmethode relativ zu Volltexten gezeigt werden (aus der linguistischen Aussage „Beistand ist maskulin“ wird zunächst „Beistand m.“; vgl. 2.2).
- In fk₁ bis fk₄ sind die Morphologieangaben bei Substantiven (MorA.S) — also die nichtelementare Angabe in „< >“ — sogar hinsichtlich ihrer Angabeform und ihrer Angabefunktion (und damit als ganze Angabe) vollständig gleich.

2.2 Zur Textverdichtung

Im folgenden geht es nur um die *innere* Textverdichtung (i.S.v. Wiegand 1998a), also um die, welche sich auf die lexikographischen Teiltexthe mit einem Leitelementträger und damit auf Wörterbuchartikel bezieht; weiterhin wird nur die *Textkondensierung*, nicht die *Textkomprimierung* betrachtet. Denn nur die erste Art der Textverdichtung, in der das Verhältnis von Schriftzeichen zu propositionalen Gehalten so modifiziert wird, daß die propositionale Dichte größer wird, ist für das Verständnis der Adressierung von Interesse. Jeder stark stan-

dardisierte Wörterbuchartikel, der auch hinsichtlich der Textkondensation standardisiert ist, so daß ein in einem bestimmten Grad verdichteter (z.B. hochverdichteter) und stark standardisierter Wörterbuchartikel gegeben ist, kann als ein zweiteiliges textuelles Kondensat (bestehend aus einem vorderen und hinteren Teilkondensat; vgl. Abb. 2) aufgefaßt werden, welches durch die Anwendung von Textverdichtungsmethoden auf einen Volltext, wie *Kürzen, Abkürzen, Auslagern, Ersetzen, Zusammenrücken, Zusammenfassen* und *Ineinerschachteln*, und unter Berücksichtigung des Codes für die typographischen und nichttypographischen Mikrostrukturanzeiger entstanden ist. Der Volltext zu einem hochverdichteten und stark standardisierten Wörterbuchartikel (wie z.B. wa_1 bis wa_5) besteht aus einem Titel und dem zugehörigen Kotext.

Bei der Anwendung von Textverdichtungsmethoden auf den Volltext von wa_2 wird der Titel des Volltextes zum vorderen Teilkondensat und der Kotext zum hinteren Teilkondensat. Durch Anwendung des Codes für die typographischen und nichttypographischen Mikrostrukturanzeiger wird sodann das vordere Teilkondensat zum Lemma, und das hintere Teilkondensat wird zu einer geordneten Menge von funktionalen Textsegmenten, nämlich Angaben und nichttypographischen Mikrostrukturanzeigern.

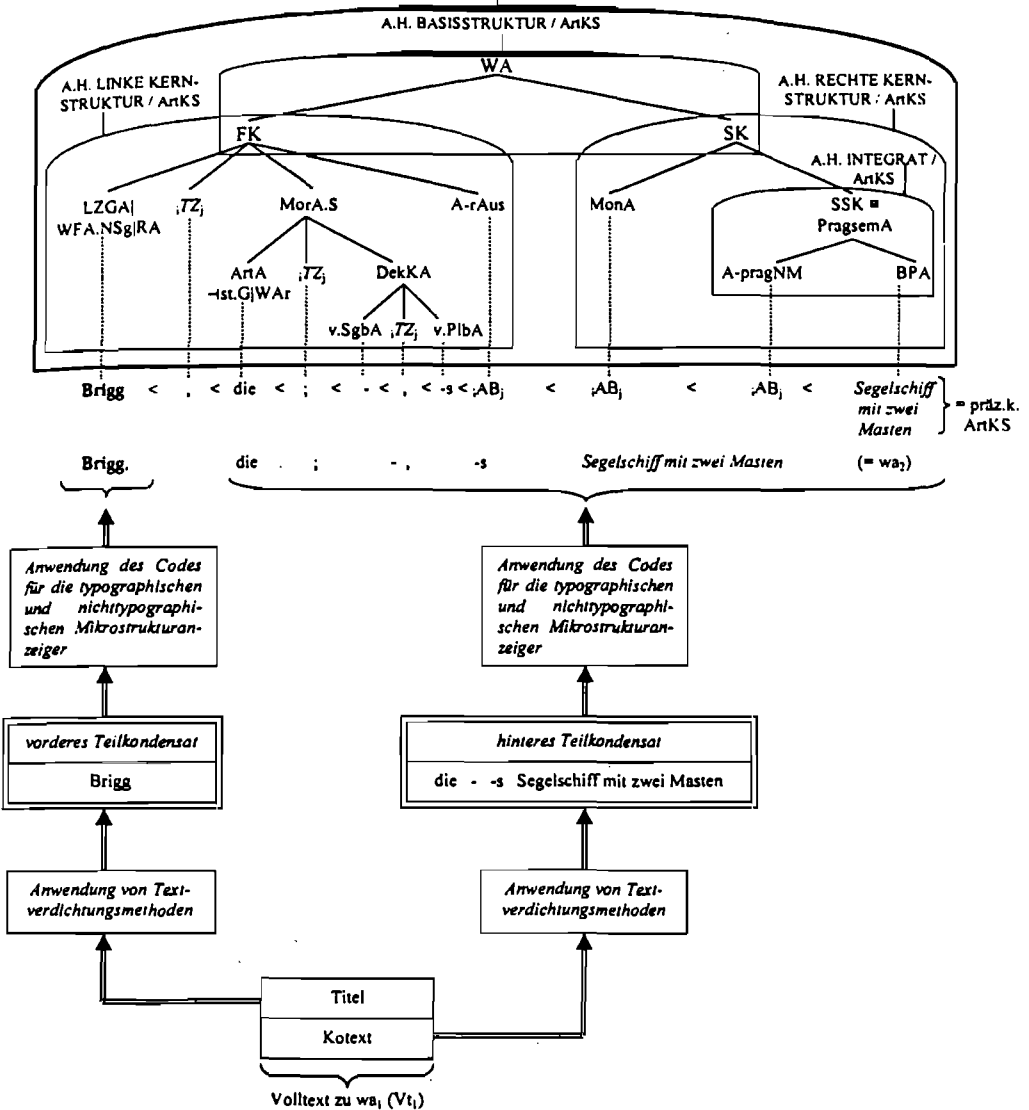
Dies ist in Abb. 2 anhand des Einzeilenartikels wa_2 aus dem HWDG veranschaulicht; zusätzlich findet sich in Abb. 2 ein erweitert kommentierter Strukturgraph zu den beiden Artikelkonstituentenstrukturen (der abstrakten und der isomorphen konkreten) von wa_2 , auf den unter 2.3 kurz zurückzukommen ist.

Das Konzept der Volltexte ist ein notwendiges Konstrukt der Theorie der Wörterbuchform, das eine angemessene Erklärung der lexikographischen Textkondensierung ermöglicht. Wenn davon gesprochen wird, Wörterbuchartikel seien verdichtet (oder: kondensiert), dann ist eine geeignete Interpretation dieser Redeweise, daß ein Textverarbeitungsprozeß vorausgesetzt wird, in dem ein Text_i zu einem Text T_j (unter Anwendung von Textverdichtungsmethoden) verdichtet wurde; T_j ist dann einer von mehreren möglichen Volltexten, die ausgehend von einem Wörterbuchartikel konstruiert werden können. Wird die Volltextbildung standardisiert, so daß nur ein Volltexttyp gebildet werden darf (was hier nicht vorgeführt werden kann; vgl. dazu Wiegand 1998a: 15ff.), werden die Wörterbuchartikel hinsichtlich der Textverdichtung vergleichbar, und der Textverdichtungsgrad wird berechenbar. Ein möglicher Volltext zu wa_2 ist Vt_1 .

- Vt_1 : [1] Wörterbuchartikel zu *Brigg*.
 [2] Die Rechtschreibung von *Brigg* ist | *Brigg* |.
 [3] Der Artikel zu *Brigg* ist *die*.
 [4] Die Form des Ng. Sg. ist *Brigg*.
 [5] Die Form des G. Sg. ist *Brigg*.
 [6] Die Form des Nom. Pl. ist *Briggs*.
 [7] *Brigg* bedeutet soviel wie *Segelschiff mit zwei Masten*.

Abb. 2: Veranschaulichung zur Textverdichtung als Übergang von einem Volltext zu einem hochverdichteten Wörterbuchartikel als zweiteiligem Kondensat

A.H. EINFACHE INTEGRIERTE ARTIKELKONSTITUENTENSTRUKTUR



Erläuterungen: „x—Y“ bedeutet (von unten nach oben gelesen) soviel wie x ist ein Teil von Y; „x⋯Y“ bedeutet (von unten nach oben gelesen) soviel wie x ist ein Element von Y (= x ∈ Y); „x ≡ y“ bedeutet soviel wie x ist (hier) identisch mit y; „x < y“ bedeutet soviel wie x geht y voraus; „x \xrightarrow{u} y“ bedeutet soviel wie x wird durch u zu y.

Abkürzungen: ArtKS = Artikelkonstituentenstruktur; präz. = präzedenitiv; k. = konkret; WA = Wörterbuchartikel; FK = Formkommentar; SK = semantischer Kommentar; SSK = semanti-

scher Subkommentar; LZGA = Lemmazeichengestaltangabe; „|“ bedeutet soviel wie *zugleich*; WFA.NSg = Wortformenangabe für den Nom. Sg.; RA = Rechtschreibangabe; ;TZ_j = Trennzeichen mit den Vorgänger- und Nachfolgervariablen „i“ und „j“; MorA.S = Morphologieangabe bei Substantiven; ArtA = Artikelangabe; „-“ bedeutet soviel wie *anhand derer erschließbar ist*; st.G = stabiles Genus; WAr = Wortart; DekKA = Deklinationsklassenangabe; v.SgbA = verdichtete Singularbildungsangabe; v.PlbA = verdichtete Pluralbildungsangabe; ArAus = Angabe zur regelmäßigen Aussprache; MonA = Monosemieangabe; PragsemA = pragmatisch-semantische Angabe; A-pragNM = Angabe zur pragmatischen Nullmarkierung; BPA = Bedeutungsparaphrasenangabe; ;AB_j = Angabeblank mit den Vorgänger- und Nachfolgervariablen „i“ und „j“; A. = ABSTRAKT; H. = HIERARCHISCH.

In Vt₁ ist [1] der Titel, und [2] bis [7] bilden den Kotext des Volltextes zu wa₁. Werden auf Vt₁ die geeigneten Textverdichtungsmethoden Satz für Satz angewendet, entsteht wa₂. Bei dem Textverdichtungsprozeß geht die syntaktische Struktur von Vt₁ weitgehend verloren und damit auch die syntaktischen Funktionen und die kohäsiven Mittel. Damit gehen auch die Kohäsionsanweisungen verloren, die es dem Benutzer ermöglichen, die mikrostrukturellen Konstituenten des Wörterbuchartikels aufeinander zu beziehen. Für den Verlust der textuellen Voraussetzungen, um die artikelinternen Angaben aufeinander beziehen zu können, sind die artikelinternen Adressierungsbeziehungen — wie wir noch genauer sehen werden — der für das Verständnis eines Wörterbuchartikels notwendige Ersatz.

2.3 Zu den Mikrostrukturen

Hochverdichtete und stark standardisierte Wörterbuchartikel weisen zahlreiche textuelle Strukturen auf.³ Im folgenden kann nur soweit auf die Mikrostrukturen eingegangen werden, wie es für das Verständnis der artikelinternen Adressierung erforderlich ist.

Alle konkreten Textkonstituentenstrukturen von Wörterbuchartikeln (Artikelkonstituentenstrukturen, Mikrostrukturen, Suchbereichsstrukturen) sind hierarchische Strukturen und dadurch erhältlich, daß Wörterbuchartikel durch die Anwendung einer der Varianten der funktional-positionalen Segmentationsmethode (i.S.v. Wiegand 1990: 20ff.) in funktionale Textsegmente zerlegt werden, so daß auf bestimmten genau festgelegten Mengen solcher Textsegmente strukturprägende Relationen definiert werden können. Die wichtigste Voraussetzung dafür besteht darin, daß man über ein geeignetes Inventar von Klassen von funktionalen Textsegmenten mit gleichem allgemeinen genuinen Zweck verfügt.

Mikrostrukturen sind Ordnungsstrukturen, so daß bestimmte Strukturausschnitte für den kundigen Benutzer als *innere Zugriffs-* oder *Schnellzugriffsstrukturen* fungieren können. Mikrostrukturen sind weiterhin Teilstrukturen von Artikelkonstituentenstrukturen. In Abb. 2 sind Strukturen des letzteren Typs, die zu wa₂ gehören, unter Zuhilfenahme formaler Darstellungsmittel

abgebildet; es handelt sich um die abstrakte hierarchische einfache integrierte Artikelkonstituentenstruktur und um die isomorphe konkrete Struktur des gleichen Typs, weiterhin um einige hierarchische Teilstrukturen dieser beiden Strukturen sowie schließlich um die konkrete präzedentive Artikelkonstituentenstruktur als Teilstruktur der konkreten hierarchischen Artikelkonstituentenstruktur. Die konkrete hierarchische Artikelkonstituentenstruktur eines bestimmten Wörterbuchartikels ist diejenige Textkonstituentenstruktur, die *erstens* die artikelinterne Distribution der Angaben festlegt und *zweitens* diese Distribution mittels der nichttypographischen Mikrostrukturanzeiger explizit kenntlich macht. Letzteres ist deswegen der Fall, weil zur Trägermenge einer konkreten hierarchischen Artikelkonstituentenstruktur auch die nichttypographischen Mikrostrukturanzeiger gehören. Im Falle von wa_2 gehören also das fette Komma nach dem Lemma, das Semikolon nach der Artikelangabe und das Komma als Teil der Deklinationsklassenangabe zur Trägermenge (vgl. Abb. 2). Im Unterschied zur Artikelkonstituentenstruktur legt die konkrete hierarchische Mikrostruktur nur die artikelinterne Distribution der Angaben fest.

Im folgenden betrachten wir Mikrostrukturen anhand von wa_3 (vgl. Abb. 3) etwas genauer:

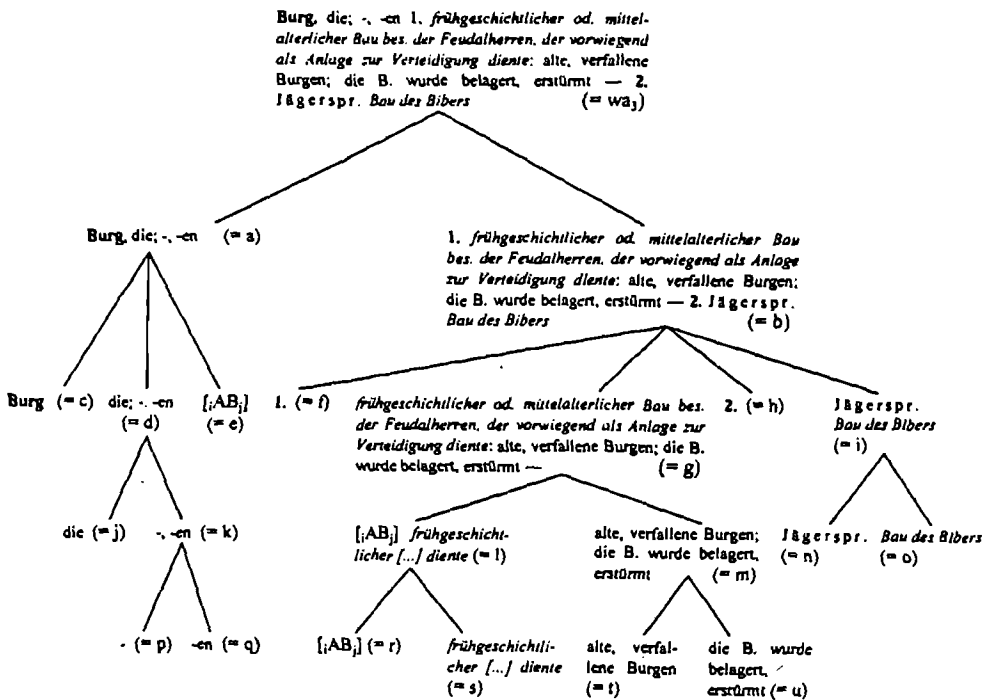
Abb. 3: wa_3 aus HWDG

Burg, die; -, -en 1. *frühgeschichtlicher od. mittelalterlicher Bau bes. der Feudalherren, der vorwiegend als Anlage zur Verteidigung diente: alte, verfallene Burgen; die B. wurde belagert, erstürmt* — 2. *Jägerspr. Bau des Bibers*

Ausgehend von einem Inventar von Klassen von Angaben mit gleichem allgemeinen genuinen Zweck segmentieren wir wa_3 unter Anwendung der Methode der nichtexhaustiven funktional-positionalen Segmentation (i.S.v. Wiegand 1990: 22). Das Segmentationsergebnis findet sich in dem kommentierten Baumgraphen in Abb. 4; die Kommentierung besteht darin, daß jeder durch die Segmentation entstehenden elementaren und nichtelementaren Angabe in runden Klammern ein Kleinbuchstabe als Individuenname zugeordnet wird. Das verwendete Inventar von Angabeklassen und damit auch die Segmentation ist sprachtheoretisch motiviert (vgl. Wiegand 2000: 236f.). Sprachliche Zeichen — wie im vorliegenden Fall *Burg* — haben eine Form und eine Bedeutung. In wa_3 werden mit der Angabe a zunächst explizite lexikographische Aussagen zur Form von *Burg* gemacht; daran schließen sich Aussagen an, die sich vor allem auf die Bedeutung von *Burg* beziehen. Daher wird die erste Segmentationsfuge zwischen a und b gelegt; a und b sind dann unmittelbare Textkonstituenten von wa_3 . Alle unmittelbaren Textkonstituenten eines ganzen Wörterbuchartikeltextes heißen (aus Gründen, die hier nicht relevant sind) *Kommentar*, wenn sie nicht nur aus einem nichttypographischen Struk-

turanzeiger bestehen; a ist der *Formkommentar* (FK) von wa₃, b der *semantische Kommentar* (SK). Weist ein Wörterbuchartikel nur einen Form- und einen semantischen Kommentar auf (wie z.B. wa₃), heißen die zugehörigen Mikrostrukturen *einfach*. Den einfachen stehen die *erweiterten* Mikrostrukturen gegenüber. Beispielsweise hat wa₅ (vgl. Abb. 13) einen Postkommentar und weist damit eine *rechtserweiterte* Mikrostruktur auf (vgl. Wiegand 1989b). Die Abkürzungen FK und SK gelten als Klassensymbole, so daß gilt: a ∈ FK und b ∈ SK. Beide Kommentare haben den Status von nichtelementaren Angaben, d.h.: sie sind weiter funktional-positional segmentierbar, ohne daß nichtfunktionale Textreste entstehen. In Abb. 4 ist das Ergebnis der funktional-positionalen Segmentation von wa₃ insgesamt dargestellt. Lediglich der nächste Segmentationschritt beim Formkommentar wird hier noch genauer erläutert.⁴ Im nächsten Segmentationschritt wird a ∈ FK in drei Angaben segmentiert, so daß im Strukturgraphen in Abb. 4 drei Kanten vom a-Knoten wegführen, und

Abb. 4: *Kommentierter Strukturgraph* zur konkreten hierarchischen einfachen integrierten Mikrostruktur von wa₃



Erläuterungen: Die Kleinbuchstaben „a“ bis „u“ in den runden Klammern sind Individuenamen für (nichtelementare und elementare) Angaben aus wa₃, und „wa₃“ ist der Individuenname für den HWDG-Artikel zu *Burg*, so daß „= x“ hier zu lesen ist wie *hat als Namen x*.

zwar zu c, d und e; c vereinigt alle lexikographischen Aussagen zum Formativ des Lemmzeichens *Burg*. Anhand der Lemmzeichengestaltangabe (LZGA) **Burg** (die im vorliegenden Fall die gleiche Form wie das Lemma aufweist) erfährt der Benutzer die Rechtschreibung und — gemäß eingespielter Gepflogenheiten in Printwörterbüchern des Neuhochdeutschen für die Präsentation von substantivischen Lemmzeichen — die Wortform im Nominativ Singular. Bei d handelt es sich um die Morphologieangabe bei Substantiven (*MorA.S*), so daß gilt $d \in \text{MorA.S}$. Auf d folgt e, und zwar als Angabeblank (AB) indiziert mit einer „Vorgänger-Variablen „i“ und einer Nachfolger-Variablen „j“ ($\{i, AB_j\}$). Ein Angabeblank wird bei der Darstellung einer konkreten Mikrostruktur stets dann in einer bestimmten textuellen Artikelposition angesetzt, wenn aus dem Fehlen einer Angabe einer bestimmten Angabeklasse in dieser Position für den kundigen Benutzer ein Schluß auf eine spezifische Eigenschaft des Lemmzeichens möglich ist. Bei HWDG-Artikeln ist dies deswegen möglich, weil im Metatext (S. XXV) steht: „Ausspracheangaben werden für alle Wörter verzeichnet, die von Ausspracheregeln der deutschen Sprache abweichen.“ Gemäß dieser Feststellung lautet der Formkommentar des HWDG-Artikels zu *Cape* wie folgt: fk_5 : *Cape*, das; -s, -s [ke:p].

In der Position der Ausspracheangabe ($ke:p \in \text{AusA}$) wird in der Strukturdarstellung zu wa_3 ein Angabeblank angesetzt, der zu folgenden Angabeklassen gehört: Klasse der Angaben zur regelmäßigen Aussprache (*A-rAus*). Es gilt daher: $v_{PlbA} AB_{PA} \in \text{A-rAus}$.

In einer Theorie der Wörterbuchform kann es nicht darum gehen, daß nur die konkreten Strukturen je spezifischer Wörterbuchartikel betrachtet werden. Um zu theoretisch fundierten Generalisierungen übergehen zu können, benötigt man abstrakte Kategorien. Im Bereich der Mikrostrukturen sind diese dadurch erhältlich, daß man mit Angabeklassen arbeitet. Dies bedeutet, daß man Trägermengen für abstrakte Mikrostrukturen bildet, deren Elemente n Angabeklassen (AK) sind (mit $n \geq 1$) und m Artikelklassen (mit $m = 1$).

Dadurch, daß wir die in Abb. 4 eingeführten Individuennamen („a“ bis „u“ und „ wa_3 “) verwenden, können wir nachfolgend wa_3 einer Artikelklasse und seine Angaben den entsprechenden Klassen von Angaben mit gleichem allgemeinen genuinen Zweck zuordnen, so daß folgende 22 Element-Klassen-Beziehungen gelten:

$wa_3 \in \text{WA}$
 $a \in \text{FK}$
 $b \in \text{SK}$
 $c \in \text{LZGA} \mid \text{WFA.NSg} \mid \text{RA}$
 $d \in \text{MorA.S}$
 $e \in \text{A-rAus}$

$f \in \text{PA}$
 $g \in \text{SSK}$
 $h \in \text{PA}$
 $i \in \text{SSK}$
 $j \in \text{ArtA} \mid \text{st.G} \mid \text{WA}_r$
 $k \in \text{DekKA}$

$l \in \text{PragsemA}$	$q \in v.\text{PlbA}$
$m \in \text{KBeiA}^2$	$r \in \text{A-pragNM}$
$n \in v.\text{FGA}$	$s \in v.\text{BPA}$
$o \in \text{BPA}$	$t \in \text{KBeiA}$
$p \in v.\text{SgbA}$	$u \in v.\text{KBei}^2\text{A}$

Damit verfügen wir über die Trägermenge ($M_{AK}(wa_3)$) der abstrakten (a) hierarchischen (h) (einfachen integrierten) Mikrostruktur (MiS), die wa_3 aufweist ($MiS_h^a(wa_3)$). Diese Trägermenge kann demnach wie folgt (in aufzählender Schreibweise) angegeben werden, wobei jede Angabeklasse nur einmal erscheint:

$M_{AK}(wa_3)$ für $MiS_h^a(wa_3) = \{\text{WA, FK, SK, LZGA} \mid \text{WFA.NSg} \mid \text{RA, MorA.S, A-rAus, PA, SSK, ArtA} \mid \text{st.G} \mid \text{WAr, DekKA, PragsemA, KBeiA}^2, v.\text{FGA, v.BPA, v.SgbA, v.PlbA, A-pragNM, BPA, KBeiA, v.KBei}^2\text{A}\}$.

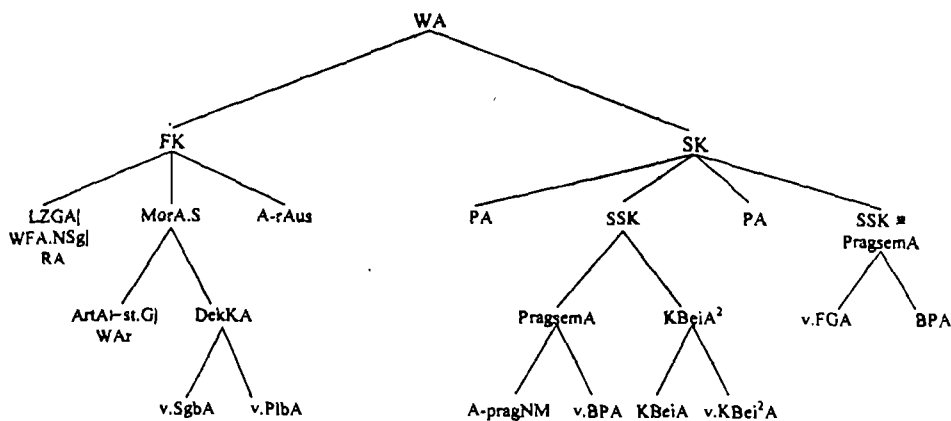
Auf der Menge $M_{AK}(wa_3)$ definieren wir nun zwei disjunkte strukturprägende Relationen, und zwar erstens eine zweistellige partitive Relation (Teil-Ganzes-Relation) und zweitens eine zweistellige Präzedenzrelation (Vorgänger-Nachfolger-Relation). Die erstgenannte Relation ist reflexiv, antisymmetrisch und transitiv; ihre Elemente sind 2-Tupel von Angabeklassen, z.B. $\langle \text{FK, WA} \rangle$, $\langle \text{MorA.S, FK} \rangle$, $\langle \text{PragsemA, SSK} \rangle$. Auf ein Paar von je einem Element aus den beiden verschiedenen Klassen eines 2-Tupels trifft der Relationsterm „x ist eine Teilangabe von y“ zu; beispielsweise ist: „ $d \in \text{MorA.S}$ ist eine Teilangabe zu $a \in \text{FK}$ “ eine wahre Aussage. Die zweitgenannte Relation, die Präzedenzrelation, ist irreflexiv, asymmetrisch und transitiv; ihre Elemente sind 2-Tupel von Angabeklassen, z.B.: $\langle \text{MorA.S, A-rAus} \rangle$, $\langle v.\text{FGA, BPA} \rangle$, $\langle v.\text{SgbA, v.PlbA} \rangle$. Auf ein Paar von einem Element aus den beiden verschiedenen Klassen eines 2-Tupels trifft der Relationsterm „x geht voraus y“ zu (nachfolgend symbolisiert als „ $x < y$ “); beispielsweise gilt: $(p \in v.\text{SgbA}) < (q \in v.\text{PlbA})$.

Die abstrakte hierarchische Mikrostruktur, die zur konkreten hierarchischen Mikrostruktur von wa_3 isomorph ist, läßt sich dann — wie in Abb. 5 — mittels eines Baumgraphen darstellen, für den die üblichen Darstellungskonventionen gelten, so daß (um der Übersichtlichkeit willen) bei der Darstellung der partitiven Relation der transitive Mantel nicht explizit mittels Kanten repräsentiert wird (z.B. fehlen die erschließbaren Kanten: „ $v.\text{PlbA} — \text{FK}$ “ oder $v.\text{FGA} — \text{WA}$), und eine eigene graphische Repräsentation der Präzedenzrelation unterbleibt (vgl. auch Wiegand 1989a). Vielmehr gelten die Konventionen des Schriftsystems sinngemäß, so daß das, was durch die Etikettierung eines am weitesten links liegenden Tochterknotens bezeichnet wird, dem vorausgeht, was durch die weiteren weiter rechts liegenden Tochterknoten des gleichen Vaterknotens bezeichnet wird.

Strukturgraphen für abstrakte textuelle Strukturen von Wörterbuchartikeln (wie z.B. den in Abb. 5) können auf zwei Arten kommentiert werden: Sie können mit einfachen oder erweiterten Kommentierungen versehen werden.

Einfache Kommentierungen bestehen aus Umrandungszeichen, die etikettiert sind; umrandet werden die jeweils dargestellte textuelle Struktur und einige oder alle ihre Teilstrukturen; die Etikettierungen bestehen aus den Strukturnamen und/oder Abkürzungen für diese. Erweiterte Kommentierungen sind erläuternde Zusätze zu einfach kommentierten Strukturgraphen. Weder die einfache noch die erweiterte Kommentierung tangieren die formalen Eigenschaften der Strukturgraphen.

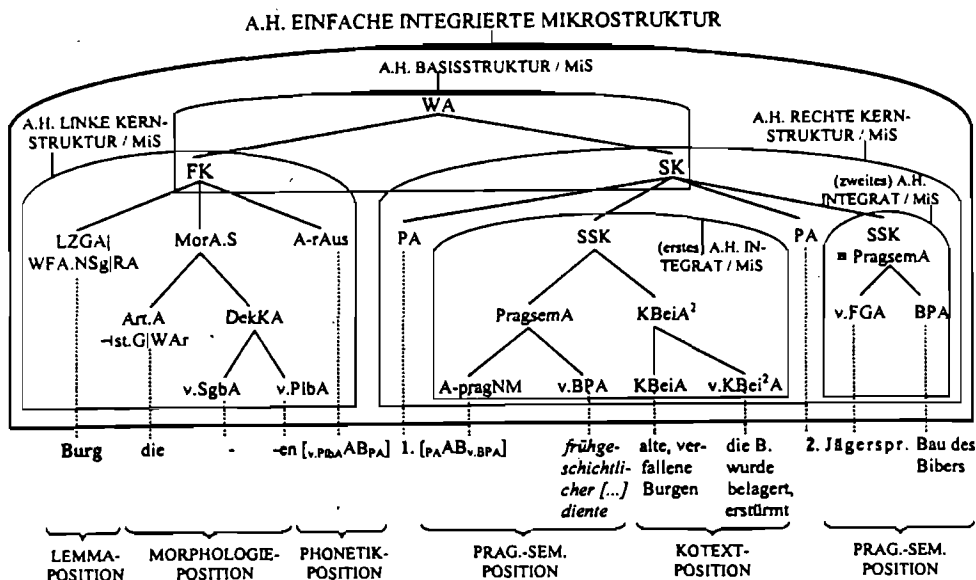
Abb. 5: Nichtkommentierter Strukturgraph zur abstrakten hierarchischen einfachen integrierten Mikrostruktur, die wa_3 aufweist



Abkürzungen: v.FGA = verdichtete Fachgebietsangabe; $KBeiA^2$ = Kompetenzbeispielangabe, die aus zwei Kompetenzbeispielangaben besteht; $v.KBei^2A$ = verdichtete Kompetenzbeispielangabe, anhand derer zwei Kompetenzbeispiele erschlossen werden können (oder: mit zwei Kompetenzbeispielen).

Zu einer abstrakten hierarchischen Textkonstituentenstruktur können stets mehrere konkrete hierarchische Textkonstituentenstrukturen verschiedener Wörterbuchartikel isomorph sein (vgl. z.B. Wiegand 1991: 459); dies bedeutet u.a.: Aussagen über Struktureigenschaften der abstrakten Struktur gelten auch für die isomorphen konkreten Strukturen. Der Isomorphismus und auf der Darstellungsebene die formalen Eigenschaften von Baumgraphen bilden die Voraussetzung dafür, daß eine abstrakte hierarchische zusammen mit der isomorphen konkreten hierarchischen Textkonstituentenstruktur eines Artikels in einem Strukturgraphen, der zwei Baumgraphen kombiniert, dargestellt werden können. Bei einer solchen Strukturdarstellung, die in der Theorie der Wörterbuchform am häufigsten verwendet wird, werden von der konkreten Textkonstituentenstruktur nur die terminalen Textkonstituenten explizit berücksichtigt; die nichtterminalen können aufgrund des Isomorphismus erschlossen werden. In Abb. 6 sind die beiden Strukturdarstellungen aus Abb. 4 und Abb. 5 zu einer Strukturdarstellung vereinigt und mit einer erweiterten Kommentierung versehen.

Abb. 6: Erweitert kommentierter Strukturgraph zur (abstrakten und zur isomorphen konkreten) hierarchischen einfachen integrierten Mikrostruktur, die wa₂ aufweist

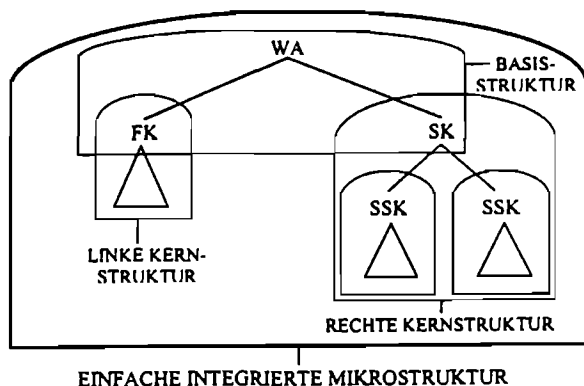


Erläuterungen: „x—y“ bedeutet (von unten nach oben gelesen) soviel wie x ist ein Teil von y; „x···y“ bedeutet (von unten nach oben gelesen) soviel wie x ist ein Element von (= x ∈ y); „x ≡ y“ bedeutet soviel wie x ist (hier) identisch mit y; „X / Y“ ist zu lesen wie X als Teilstruktur von Y.

Abkürzungen: A. = ABSTRAKT; H. = HIERARCHISCH; WA = Wörterbuchartikel; FK = Formkommentar; SK = semantischer Kommentar; MiS = Mikrostruktur; LZGA = Lemmazcichengestaltangabe; WFA.NSg = Wortformenangabe: Nominativ Singular; RA = Rechtschreibangabe; MorA.S = Morphologieangabe bei Substantiven; A-rAus = Angabe zur regelmäßigen Aussprache; ArtA -st.G|WAr = Artikelangabe, anhand derer das stabile (st.) Genus (G) sowie die Wortart (WAr) erschließbar sind; DekKA = Deklinationsklassenangabe; v.SgbA = verdichtete Singularbildungsangabe; v.PlbA = verdichtete Pluralbildungsangabe; SSK = semantischer Subkommentar; PragsemA = pragmatisch-semantische Angabe; A-pragNM = Angabe zur pragmatischen Nullmarkierung; v.BPA = verdichtete Bedeutungsparaphrasenangabe; KBeiA² = (nichtelementare) Kompetenzbeispielangabe, die aus zwei Kompetenzbeispielangaben besteht; KBeiA = Kompetenzbeispielangabe; v.KBei²A = verdichtete Kompetenzbeispielangabe, anhand derer zwei Kompetenzbeispiele erschließbar sind (oder: mit zwei Kompetenzbeispielen); v.FGA = verdichtete Fachgebietsangabe; AB = Angabeblank; v.PlbA ABPA = Angabeblank, dessen Vorgänger eine v.PlbA und dessen Nachfolger eine PA ist; PA AB v.BPA = Angabeblank, dessen Vorgänger eine PA und dessen Nachfolger eine v.BPA ist.

Zu jeder Strukturdarstellung von abstrakten Textkonstituentenstrukturen kann ein *allgemeines Strukturbild* und ein *erweitertes allgemeines Strukturbild* angegeben werden, die nach festgelegten Darstellungskonventionen erstellt werden (vgl. z.B. Wiegand 2000). Das allgemeine Mikrostrukturbild für wa_3 findet sich in Abb. 7. Die Strukturbilder dienen vor allem zur schnellen Orientierung über die Mikrostrukturart.

Abb. 7: Allgemeines Mikrostrukturbild für einfach integrierte Mikrostrukturen von Wörterbuchartikeln zu zweifach polysemen Lemmazeichen wie z.B. wa_3



Die bisher über Standardisierung, Textverdichtung und Mikrostrukturen von Wörterbuchartikeln eingeführten Kenntnisse sind zwar noch relativ eingeschränkt, dürften aber ausreichen, um die sich anschließende Darstellung der artikelinternen Adressierung verstehen zu können.

2. Adressierung in der einsprachigen Lexikographie

Wir betrachten nachfolgend vor allem die artikelinterne Adressierung, und zwar zunächst anhand von wa_3 (vgl. Abb. 3). Als erstes muß daran erinnert werden, daß der Ausdruck *Adresse* nicht nur — wie auch *adressieren* — in der Alltagssprache unterschiedlich verwendet wird (vgl. Wiegand 1989a: 445), sondern auch in verschiedenen theoretischen Zusammenhängen unterschiedlich terminologisiert ist. In der Dokumentationswissenschaft und in der Datenbanktheorie wird unter einer Adresse von (oder: für) Daten, die sich auf einem Speichermedium befinden, die Angabe des Speicherplatzes verstanden, die es ermöglicht, gezielt auf die Daten zuzugreifen. Zwar kann man auch in lexikographischen Zusammenhängen davon sprechen, daß ein Lemma, wie z.B. *Burg* in wa_3 , eine Adresse darstellt, unter der ein Benutzer-in-actu etwas sucht und gegebenenfalls etwas findet; diese eher werkstattsprachliche (durchaus gut verständliche) Redeweise ist aber für eine präzise lexikographietheoretische Sprechweise weniger gut geeignet. Denn ein Lemma in einem Sprachwörter-

buch kann nicht nur als eine Adresse im gerade genannten Sinn gelten, weil ein Benutzer anhand des Lemmas selbst auch Informationen über die Lemmazeichengestalt, z.B. über eine bestimmte Wortform oder über die Form eines Affixes sowie über die Rechtschreibung erschließen kann. Während also anhand eines Lemmas Informationen über den Wörterbuchgegenstand erschließbar sind, gilt das z.B. für eine Adresse einer Datenbank nicht notwendigerweise.

Wenn man von Adressen im Zusammenhang mit der artikelinternen Adressierung spricht, sind Adressen immer Angaben. In wa_3 sind alle Angaben (außer dem Formkommentar) an die Lemmazeichengestaltangabe adressiert (vgl. Abb. 15). Die zweistellige Beziehung „ x ist adressiert an y “ (mit „ x “ und „ y “ als Variablen für Angaben) heißt *Adressierungsbeziehung*. Beispielsweise gelten für wa_3 folgende Aussagen:

- MorA.S ist adressiert an die LZGA.
- DekKA ist adressiert an die LZGA.
- PragsemA ist adressiert an die LZGA.

Der Sachverhalt der Adressierung kann auch konvers mit „ y ist die Adresse von x “ wiedergegeben werden (mit „ y “ als Variable für Adressen und „ x “ als Variable für Angaben). Entsprechend gilt dann z.B.: Die LZGA ist die Adresse der MorA.S.

In der lexikographischen Werkstattsprache nennt man eine Adresse zusammen mit einer an sie adressierte Angabe auch eine *lexikographische Bearbeitungseinheit*. In wa_3 ist z.B. <Burg, Bau des Bibers> eine lexikographische Bearbeitungseinheit (vgl. auch 3). Es ist hierbei allerdings unbedingt zu beachten, daß die Konverse der Adressierungsrelation nicht mit der Bearbeitungsrelation verwechselt wird. Denn bearbeitet werden Phänomene des Wörterbuchgegenstandes und nicht Angaben, während Angaben an Angaben adressiert sind und nicht an Phänomene des Wörterbuchgegenstandes; daher muß auch die Angabebeziehung „ x ist Angabe zu y “ (auf die hier aus Platzgründen nicht näher eingegangen werden kann) strikt von der Adressierungsbeziehung unterschieden werden (vgl. Wiegand 2000e).

Die beiden Elemente einer lexikographischen Bearbeitungseinheit sind als verschiedene Angaben lexikographische Daten und keine lexikographischen Informationen. Wie bereits erwähnt, ist eine lexikographische Information eine kognitive Entität, die während der Ausführung von Benutzungshandlungen beim Benutzer-in-actu entsteht. Sieht man von Informationen über geschriebene Wortformen ab, die sich aus der jeweiligen Angabeform als solcher beim Lesen unmittelbar ergeben, dann sind lexikographische Informationen nur dadurch erhältlich, daß ein Benutzer-in-actu eine Angabe auf dasjenige bezieht, wozu sie eine Angabe ist (also die Angabebeziehung rekonstruiert). Die Herstellung dieses Bezugs wird durch die Kenntnis der Adressierungsbeziehung

möglich und kann nur so erfolgen, daß eine Angabe a_1 auf eine andere Angabe a_2 bezogen wird, weil a_2 die Adresse von a_1 ist. Erst danach kann der zweite Schritt der Informationsgewinnung erfolgen; dieser besteht in der kognitiven Konstruktion des Adressaten als dem wichtigsten Teil der Antwort auf eine Suchfrage.

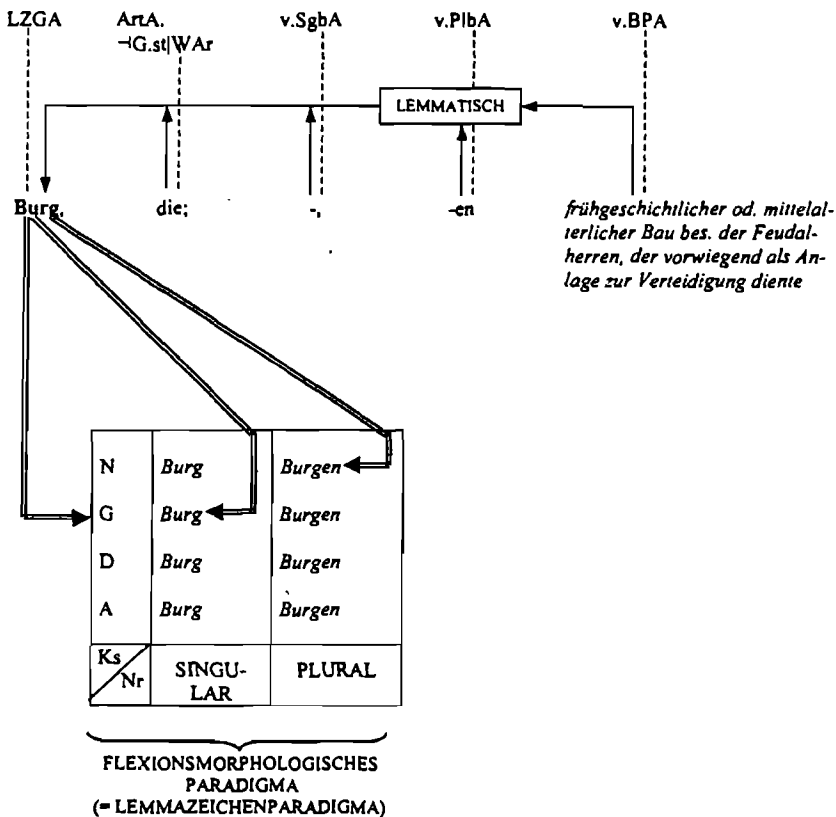
Will beispielsweise ein Benutzer-in-actu die verdichtete Singularbildungsangabe „-“ in wa_3 (also: - \in v.SgbA) für eine Antwort auf seine Suchfrage „Wie lauten die Singularformen von *Burg*?“ nutzen, dann muß er die verdichtete Singularbildungsangabe auf die Lemmazeichengestaltangabe, also auf *Burg* beziehen. Unter der Voraussetzung, daß der Benutzer-in-actu weiß, daß „-“ die Singularbildungsangabe ist, ergibt sich erst nach Ausführung der genannten kognitiven Operation eine lexikographische Information. Diese besteht in dem punktuellen Wissen, daß die vier Singularformen von *Burg*, die der Benutzer-in-actu bilden muß, *Burg* lauten (vgl. Wiegand 2000e).

In Wörterbuchartikeln einsprachiger Wörterbücher (wie z.B. wa_1 bis wa_5) ist die zentrale Adresse meistens die Lemmazeichengestaltangabe (LZGA). Diese weist in mehreren Hinsichten einen Sonderstatus auf. Sie ist *erstens* Träger des Leitelementes, also der Buchstabenfolge, nach der alphabetisiert wird (vgl. Wiegand 1989). In wa_3 beispielsweise ist die LZGA *Burg* hinsichtlich der Eigenschaften, die bei der Alphabetisierung eine Rolle spielen können, nämlich hinsichtlich der Anzahl und hinsichtlich der Reihenfolge der Buchstaben, mit dem Leitelement |*Burg*| gleich. Dies ist natürlich nicht notwendigerweise so, sondern nur dann, wenn die Methode der exhaustiven mechanischen Alphabetisierung (wie in den meisten modernen Wörterbüchern) angewandt wurde (vgl. dazu Wiegand 1989). Das Leitelement ist stets ein Element der Trägermenge der Makrostruktur und damit auch stets ein Element der (oder: einer der) äußeren Zugriffsstruktur(en). Die Einheit, die das Leitelement trägt und damit ein Zugriffselement ist, und anhand derer man zugleich etwas über die schriftkonstituierte Wortform des Lemmazeichens erfährt, heißt *Lemma*. Spricht man von der Lemmazeichengestaltangabe, dann wird von dem Zugriffsaspekt abstrahiert, d.h.: die Eigenschaft, ein Leitelementträger zu sein, spielt dann keine Rolle, sondern nur die Eigenschaft, daß die Gestalt des Lemmazeichens präsentiert wird. Zum Sonderstatus der Lemmazeichengestaltangabe gehört *zweitens*, daß sie neben dem Formkommentar, dessen Teilangabe sie ist, die einzige Angabe unter allen Wörterbuchangaben ist, die nicht eine andere Angabe als Adresse hat. Die Lemmazeichengestaltangabe ist adressenlos, ist aber dennoch eine Angabe.⁵

Angaben, welche die gleiche artikelinterne Adresse haben, können verschiedene Adressaten haben. So gilt beispielsweise für die in Abb. 8 berücksichtigten Angaben aus wa_3 folgendes: Die Artikelangabe, die verdichtete Singularbildungsangabe, die verdichtete Pluralbildungsangabe und die verdichtete Bedeutungsparaphrasenangabe sind an die Lemmazeichengestaltangabe adressiert; ihre Adressaten sind jedoch z.T. verschieden. Denn es gilt:

- Der Adressat der Artikelangabe ist das Lemmazeichenparadigma (alle acht Wortformen von *Burg* sind femininum).
- Der Adressat der verdichteten Singularbildungsangabe ist die Form des Genitiv Singulars.
- Der Adressat der verdichteten Pluralbildungsangabe ist die Form des Nominativ Plurals.
- Der Adressat der verdichteten Bedeutungsparaphrasenangabe ist das Lemmazeichenparadigma (allen acht Wortformen von *Burg* kommt die Bedeutung zu, die anhand der v.BPA erschließbar ist).

Abb. 8: Veranschaulichung zum Zusammenhang von Angaben, lemmatischer Adresse und Adressaten anhand von wa_3



Erläuterungen: „ $x \rightarrow y$ “ bedeutet soviel wie x ist artikelintern adressiert an y ; „ $u \Rightarrow v$ “ bedeutet soviel wie u ist die Adresse zum Adressaten v ; „ $z \text{-----} Z$ “ bedeutet soviel wie $z \in Z$ (z ist ein Element von Z); Abkürzungen: N = Nominativ; G = Genitiv; D = Dativ; A = Akkusativ; Ks = Kasus; Nr = Numerus.

Es lassen sich verschiedene Arten der artikelinternen Adressierung unterscheiden. Auf einige Arten wird nachfolgend kurz eingegangen. Die Unterscheidun-

gen gelten nur für Wörterbücher, deren Wörterbuchgegenstandsbereich Sprachen oder Varietäten bilden, deren Schriftsystem eine zeilengebundene Verschriftung von links nach rechts vorsehen. Eine Angabe ist artikelintern *linksadressiert*, wenn ihre Adresse gegen die Schriftrichtung (oder: Schreibrichtung) und damit auf den Zeilen nach links orientiert gesucht werden muß. Entsprechend ist eine Angabe artikelintern *rechtsadressiert*, wenn ihre Adresse mit der Schriftrichtung und damit auf den Zeilen nach rechts orientiert gesucht werden muß. Die Rechtsadressierung ist relativ selten. Sie findet sich beispielsweise im Campe-WdS. Hier haben zahlreiche Artikel einen Präkommentar, der aus Angabesymbolen besteht, die nach rechts an die Lemmazeichengestaltangabe adressiert sind. Beispiele sind die folgenden Einträge (e):

- e₁: † Beschlagen; v. intrs. unregelm [...]
 e₂: O Die Beisitzerstelle, Mz die -n [...]
 e₃: Δ Durchschläfen, v. ntr. [...]

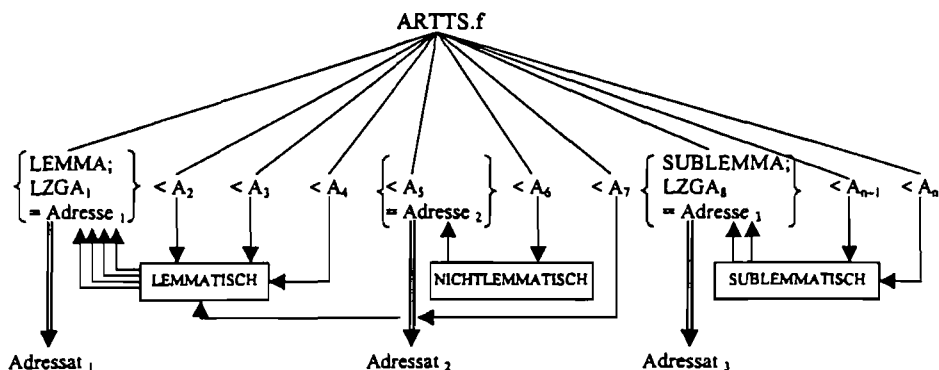
Auch die Artikelangabe „Die“ in e₂ ist nach rechts an die Lemmazeichengestaltangabe „Beisitzerstelle“ adressiert.

Die Unterscheidung von Links- und Rechtsadressierung gilt vor allem für die artikelinterne Adressierung. Artikelinterne Verweisangaben sind — wie wir gesehen haben — stets zweifach adressiert. Sie sind daher — wie alle anderen Angaben auch — zunächst einmal entweder artikelintern rechts- oder linksadressiert. Bezüglich ihrer Verweisadressen, die außerhalb des Artikels liegen, zu dem sie gehören, spricht man dagegen nicht von Links- oder Rechtsadressierung. Vielmehr wird in dem Fall, in welchem der Adressenfundort innerhalb der Lemmaliste des gleichen Wörterverzeichnis liegt, zu dem der Artikel mit der fraglichen Verweisangabe gehört, entweder von *Hinauf-* oder *Hinabadressierung* gesprochen. Besonders in älteren Wörterbüchern benutzte man entsprechend als Verweisbeziehungsangabe Pfeile mit ikonischen Eigenschaften („↑“, „↓“).

Bei der artikelinternen Adressierung kann weiterhin zwischen lemmatischer, sublemmatischer und nichtlemmatischer Adressierung unterschieden werden. Alle drei Arten können als Links- oder Rechtsadressierung ausgeprägt sein. Abb. 9 veranschaulicht die drei Möglichkeiten anhand der wesentlich öfter gegebenen Linksadressierung.

Lemmatische Adressierung liegt vor, wenn die artikelinterne Adresse die Lemmazeichengestaltangabe ist. *Sublemmatische Adressierung* ist gegeben, wenn die subartikelinterne Adresse die Zeichengestaltangabe zum Sublemmazeichen ist; und schließlich handelt es sich um *nichtlemmatische Adressierung*, wenn die Adresse eine artikelinterne Angabe ist, die nicht als Leitelementträger fungiert. Sind alle Angaben eines Wörterbuchartikels (außer der Lemmazeichengestaltangabe und dem Formkommentar, zu dem die LZGA gehört) an die Lemmazeichengestaltangabe adressiert, liegt *vollständige lemmatische Adressierung* vor.

Abb. 9: Veranschaulichung zu den Arten der artikelinternen Linksadressierung bei hochverdichteten und stark standardisierten Wörterbuchartikeln: Lemmatische, nichtlemmatische und sublemmatische Linksadressierung



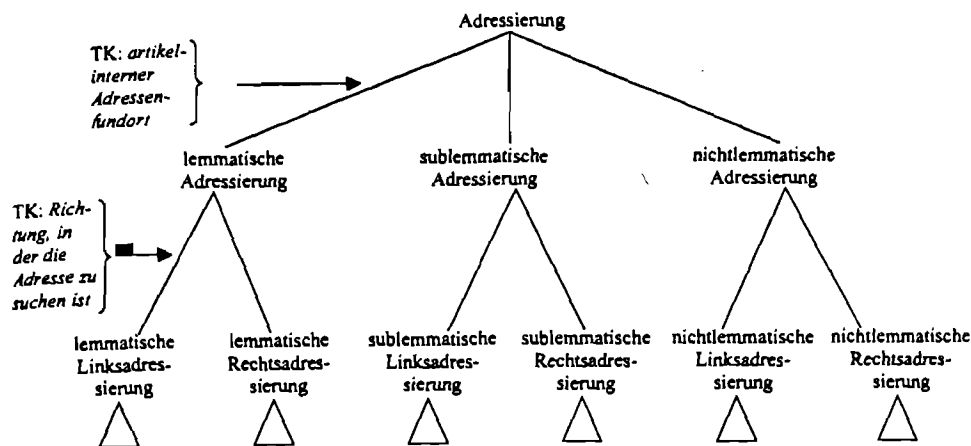
Erläuterungen: „ $x \rightarrow y$ “ bedeutet soviel wie x ist artikelintern adressiert an y ; „ $u \rightarrow v$ “ bedeutet soviel wie u ist die Adresse zum Adressaten v ; „ \leftarrow “ bedeutet soviel wie geht voraus; Abkürzungen: A = Angabe; ARTTS.f = funktionale Artikelteilstrecke; LZGA = Lemmzeichengestaltangabe.

Dies heißt zugleich, daß es keinen artikelinternen Themenwechsel gibt: Alle lexikographischen Aussagen, die in der verdichteten Form von lexikographischen Bearbeitungseinheiten gegeben sind, sind daher vom Benutzer-in-actu auf das Lemmzeichen und das Lemmzeichenparadigma zu beziehen. Für Subartikel gilt dies sinngemäß. Hier spricht man entsprechend von *vollständiger sublemmatischer Adressierung*.

Mit den bisherigen Unterscheidungen verfügen wir über erste Ansätze zu einer Typologie der artikelinternen Adressierungsbeziehungen, die in den Abb. 10 und 11 — zusammen mit einer weiteren noch zu erläuternden Differenzierung — in eine Übersicht gebracht sind.

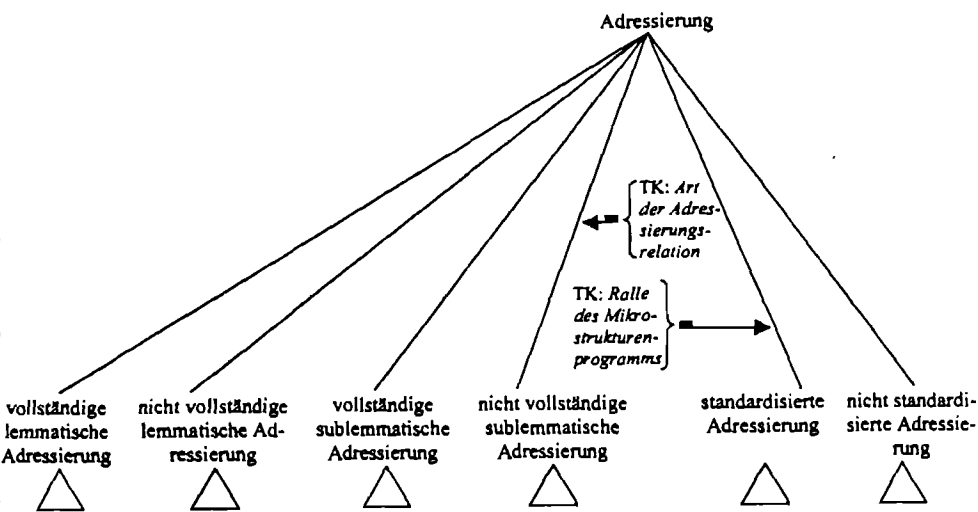
Die Standardisierung, wie sie unter 2.1 erläutert wurde, betrifft auch die Adressierungsbeziehungen (vgl. Abb. 11). Daher kann zwischen *standardisierter* und *nichtstandardisierter Adressierung* unterschieden werden. Eine Adressierungsbeziehung gilt als *standardisiert*, wenn nach dem Mikrostrukturenprogramm des Wörterbuchs die Klasse der Adressen für eine bestimmte Angebeklasse vorgeschrieben ist; dies ist z.B. dann der Fall, wenn es eine Standardisierungsinstruktion gibt, die besagt, daß alle Bedeutungsparaphrasenangaben im semantischen Kommentar an die Lemmzeichengestaltangabe zu adressieren sind. Adressen, die im Mikrostrukturenprogramm als Adressen für Angaben aus bestimmten Klassen festgelegt sind, sind *standardisierte Adressen*. *Nichtstandardisierte Adressen* sind dagegen solche, die vom bearbeitenden Lexikographen frei gewählt werden können (wobei es hier unterschiedliche Freiheitsgrade geben kann); für die zugehörigen Angaben, die an die nichtstandardisierten Adressen adressiert sind, gilt dies entsprechend. Die frei wählbaren

Abb. 10: Erster Ausschnitt aus einer Typologie der artikelinternen Adressierungsbeziehungen



Erläuterungen: „ $\blacksquare \rightarrow$ “ bedeutet soviel wie *die Anwendung des TK führt zu der Unterteilung*, Abkürzungen: TK = Typologiekriterium.

Abb. 11: Zweiter Ausschnitt aus einer Typologie der artikelinternen Adressierungsbeziehungen



Erläuterungen: „ $\blacksquare \rightarrow$ “ bedeutet das gleiche wie „ $\leftarrow \blacksquare$ “ (vgl. Abb. 10).

Angaben heißen *Glossate* (vgl. Wiegand 1989a: 447); die nichtstandardisierten Adressen heißen entsprechend *Glossat*-adressen. Je nach ihrer Position zu der Angabe, die die Glossatadresse darstellt oder von der ein Teil eine Glossat-

adresse ist, können *Prä-*, *Post-* und *Binnenglossate* unterschieden werden; *Präglossate* sind rechts-, *Binnen-* und *Postglossate* dagegen linksadressiert. *Glossate* kommen in vielen Wörterbüchern häufig vor. Zahlreiche *Glossate* finden sich z.B. in wa_4 in Abb. 12.

Abb. 12: wa_4 aus dem HWDG

sauber /Adj./ 1. *frei von Schmutz, Verunreinigungen*: saubere Finger(nägel), Kleider, Schuhe, Straßen; sauberes Wasser; ein sauberes Besteck; Hemd; etw. s. bürsten, putzen, waschen; sich s. halten; das Kind ist schon s. (*braucht keine Windeln mehr zu tragen*) – 2.1. *sorgfältig und ordentlich (ausgeführt)*: eine saubere Arbeit, Naht, Schrift; stw. s. aussägen, stopfen 2.2. *im Klang rein* (1.2): dieser Ton war nicht ganz s. 2.3. eine saubere (*angemessene und überzeugende, widerspruchsfreie*) Lösung des Problems 2.4. /nur adv./ umg. das hast du s. (*sehr gut*) hingekriegt! – 3.1. *moralisch einwandfrei*: eine saubere Gesinnung, Haltung; ein sauberer Charakter; iron. du hast ja saubere (*anfechtbare*) Ansichten!; dein sauberer Freund (*dein Freund, dieser gemeine Kerl*) hat uns das eingebrockt! 3.2. /nicht adv./ ein sauberes (*von Schuld unbelastetes*) Gewissen – 4. umg. süddt. österr. schweiz. *hübsch* (1.1), *schmuck*: ein sauberes Mädchen, ein sauberer Bursche; s. aussehen + (k)eine saubere ↗ Weste haben

Für wa_4 gelten folgende Aussagen:

- „*braucht keine Windeln mehr zu tragen*“ ist ein (linksadressiertes) Postglossat zur verdichteten Kompetenzbeispielangabe „das Kind ist schon s.“ als Glossatadresse.
- „*angemessene und überzeugende, widerspruchsfreie*“ ist ein nichtelementares (linksadressiertes) Binnenglossat zur Glossatadresse „saubere“, das aus zwei elementaren (= nicht weiter segmentierbaren) Binnenglossaten besteht.
- „*sehr gut*“ ist ein (linksadressiertes) Binnenglossat zur Glossatadresse „s.“.
- „*anfechtbare*“ ist ein (linksadressiertes) Binnenglossat mit der Glossatadresse „saubere“.
- „*dein Freund, dieser gemeine Kerl*“ ist ein (linksadressiertes) Binnenglossat mit der Glossatadresse „sauberer Freund“.
- „*von Schuld unbelastetes*“ ist ein (linksadressiertes) Binnenglossat mit der Glossatadresse „sauberes“.

Mit den sechs Glossaten in wa_4 sind sechs *nichtstandardisierte Themenwechsel* verbunden. Diese müssen von den standardisierten Themenwechseln unterschieden werden wie sie z.B. in wa_5 vorliegen.⁶

Abb. 13: wa_5 aus dem HWDG

Streich, der; -(e)s, -e aus Übermut, Mutwillen begangene Handlung, durch die jmd. geneckt, heringelegt, geärgert werden soll: ein lustiger, übermütiger, böser S.; freche, tolle Streiche machen, vollführen, verüben; er ist stets zu Streichen aufgelegt
 + jmd. spielt jmdm. einen S. (*legt jmdn. herein*);
 etw. spielt jmdm. einen S. *etw. läßt jmdn. im Stich*:
 sein Gedächtnis hat ihm einen S. gespielt

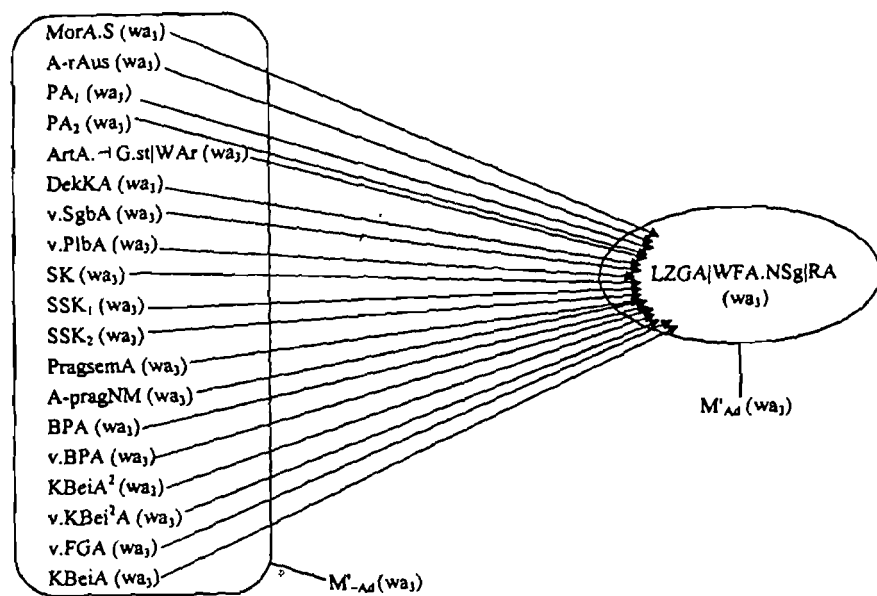
Im Postkommentar von wa_5 , der mit der Phrasemidentifizierungsangabe „+“ beginnt, sind „*legt jmdn. herein*“ und „*etw. läßt jmdn. im Stich*“ nicht etwa verdichtete Glossate, sondern nach dem Mikrostrukturenprogramm des HWDG verdichtete Bedeutungsparaphrasenangaben für Phraseme, so daß in wa_5 die verdichteten Phrasemangaben „jmd. spielt jmdm. einen S.“ und „etw. spielt jmdm. einen S.“ ihre standardisierten Adressen sind und somit in wa_5 ein standardisierter Themenwechsel vorliegt.

Bei Wörterbuchartikeln mit vollständiger lemmatischer Adressierung und bei solchen mit vollständiger sublemmatischer Adressierung ist die zweistellige Adressierungsrelation naheindeutig, so daß sie eine Funktion ist. Anhand von wa_3 (vgl. Abb. 3) sei dies etwas eingehender erläutert. Die Adressierungsfunktion kann z.B. mittels eines Pfeildiagramms dargestellt werden (vgl. Abb. 14).

Nach der Analyse der Mikrostrukturen von wa_3 in 2.3. verfügen wir sowohl über die Trägermenge der abstrakten hierarchischen als auch über die der konkreten (k) hierarchischen (h) Mikrostruktur (MiS_h^k) von wa_3 . Letztere Menge heiße $M_A(wa_3)$. Die 22 Elemente dieser Menge sind in 2.3. aufgezählt, und zwar unter Verwendung der in Abb. 4 eingeführten Individuennamen „a“ bis „u“ und „ wa_3 “. Wir vermindern nun $M_A(wa_3)$ um die Elemente $wa_3 \in WA$ und $a \in FK$ und erhalten $M'_A(wa_3)$. Auf dieser Menge $M'_A(wa_3)$ definieren wir nun die zweistellige irreflexive und asymmetrische Adressierungsrelation — sie heiße R_a ($R_a \subseteq M'_A(wa_3) \times M'_A(wa_3)$), auf deren 2-Tupel der Relationsterm $T_a = x$ ist *artikelintern standardisiert adressiert an y* zutrifft, wobei „x“ eine Variable für Angaben (die keine Adressen sind) ist und „y“ eine Variable für Adressen. Da wa_3 ein Artikel mit vollständiger lemmatischer Adressierung ist, kann für „y“ nur „LZGA | WFA.NSg | RA“ eingesetzt werden; d.h. in jedem der 2-Tupel der Adressierungsrelation R_a erscheint als zweite Koordinate LZGA | WFA.NSg | RA, z.B. in dem 2-Tupel $\langle \text{MorA.S, LZGA | WFA.NSg | RA} \rangle$, auf das

T_a zutrifft. In der Abb. 14 entspricht diesem 2-Tupel die Darstellungskomponente „MorA.S \rightarrow LZGA|WFA.NSg|RA“. Die Definition von R_a auf der Menge $M'_A(wa_3)$ führt zu deren Zerlegung in zwei disjunkte Teilmengen. Dies sind die beiden folgenden Mengen: $M'_{Ad}(wa_3)$, die Menge der Angaben von wa_3 , die Adressen sind (eine Einermenge), und $M'_{-Ad}(wa_3)$, die Menge der Angaben, die keine Adressen sind. Das Pfeildiagramm in Abb. 14 ist mithin eine formale Strukturdarstellung für die konkrete Adressierungsstruktur, die wa_3 aufweist und damit ein Beispiel für die Darstellung der konkreten Adressierungsstruktur eines Wörterbuchartikels mit vollständiger lemmatischer Adressierung.

Abb. 14: Pfeildiagramm zur vollständigen lemmatischen Adressierung in wa_3



Erläuterungen: MorA.S (wa_3) ist zu lesen wie die (konkrete) Morphologieangabe bei Substantiven in wa_3 ; „ $x \rightarrow y$ “ bedeutet soviel wie x ist artikelintern standardisiert adressiert an y .

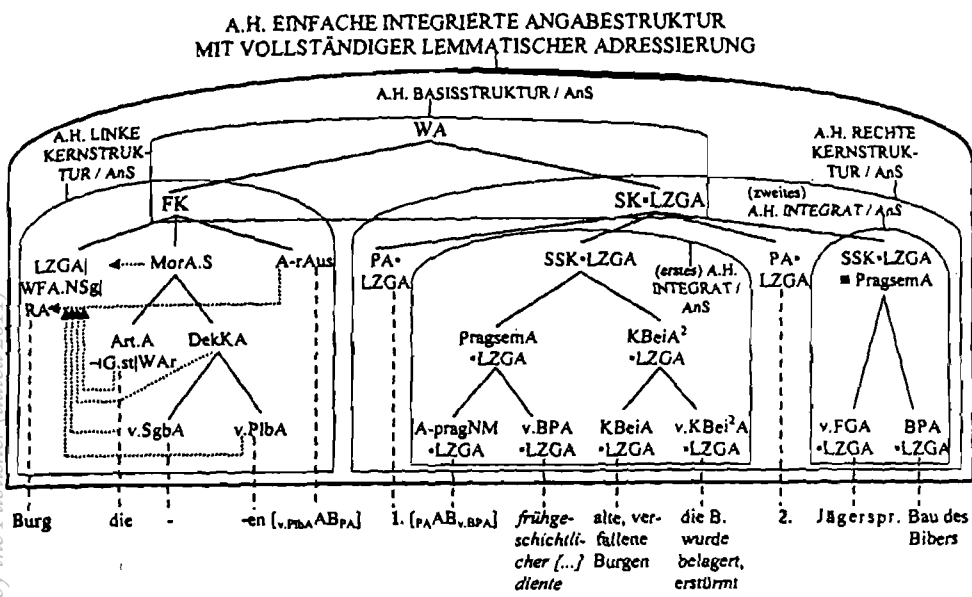
Abkürzungen: $M'_{Ad}(wa_3)$ = Menge der Angaben von wa_3 , die Adressen sind; $M'_{-Ad}(wa_3)$ = Menge der Angaben von wa_3 , die keine Adressen sind.

Es ist zu beachten, daß in Abb. 14, um die konkreten Angaben besser identifizieren zu können, andere Individuennamen als in Abb. 4 für die konkreten Angaben in wa_3 verwendet sind, nämlich z.B. „DekKA(wa_3)“; dieser Name ist zu lesen wie Deklinationsklassenangabe in wa_3 (nämlich $d \in$ DekKA). Bei einer Darstellung der abstrakten Adressierungsstruktur sind die Elemente der Träger-

menge Angabeklassen, so daß die Mächtigkeit der Trägermenge geringer ist als die Mächtigkeit von $M'_A(wa_3)$. Denn in wa_3 gibt es Angaben, die zu gleichen Angabeklassen gehören, z.B. gilt: $1. \in PA$ und $2. \in PA$.

Führt man die Mikrostrukturen von wa_3 , wie sie in Abb. 6 dargestellt sind, mit den Adressierungsstrukturen zusammen, dann erhält man die konkrete hierarchische und abstrakte hierarchische Angabestruktur von wa_3 , wie sie in Abb. 15 dargestellt sind.

Abb. 15: Erweitert kommentierter Strukturgraph zur (abstrakten und zur isomorphen konkreten) hierarchischen einfachen integrierten Angabestruktur mit vollständiger lemmatischer Adressierung, die wa_3 aufweist



Erläuterungen: „x.....►y“ im FK bedeutet soviel wie x ist artikelintern adressiert an y (mit „x“ und „y“ als Variablen für Angabeklassensymbole); „x.y“ im SK ist synonym mit „x.....►y“; Abkürzungen: AnS = Angabestruktur.

Im Unterschied zu einem Strukturgraphen, mit dem die Mikrostrukturen eines Artikels dargestellt sind, kann man anhand eines Strukturgraphen, mit welchem die Angabestrukturen dargestellt sind, erschließen, welche Angaben an welche Adressen adressiert sind, so daß damit auch die jeweilige Adressierungsart erkennbar ist. In Abb. 15 ist die Kennzeichnung der Adressierungsbeziehungen (aus Gründen der Veranschaulichung) mittels zweier unterschiedlicher Verfahren vorgenommen. Bei der Darstellung der Adressierungsbeziehungen im Formkommentar sind Pfeile mit gepunktetem Schaft verwendet. Bei größeren Artikeln führt die Anwendung dieser Methode zu unübersichtlichen Strukturdarstellungen. Daher ist das Verfahren vorzuziehen, das

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2014)

bei der Darstellung des semantischen Kommentars und der zugehörigen abstrakten hierarchischen rechten Kernstruktur als Teilstruktur der Angabestruktur angewandt wurde. Das Klassensymbol für die Angabeklasse, zu der die Adresse gehört, wird (getrennt durch einen Mittelpunkt) unmittelbar hinter dem Klassensymbol für die adressierte Angabe aufgeführt; es wird also z.B. das erweiterte Knotenetikett „PragsemA•LZGA“ geschrieben und (vereinfacht) gelesen: „Pragmatisch-semantische Angabe, adressiert an die Lemmazeichengestaltangabe“. Die Abkürzung für den Klassennamen einer Angabeklasse, die hinter dem Mittelpunkt steht, heißt *Adressensymbol*; im vorliegenden Fall ist das Adressensymbol „LZGA“. Anhand der Adressensymbole in einer Strukturdarstellung kann man unmittelbar ablesen, welche Adressierungsart in einem Wörterbuchartikel gegeben ist.

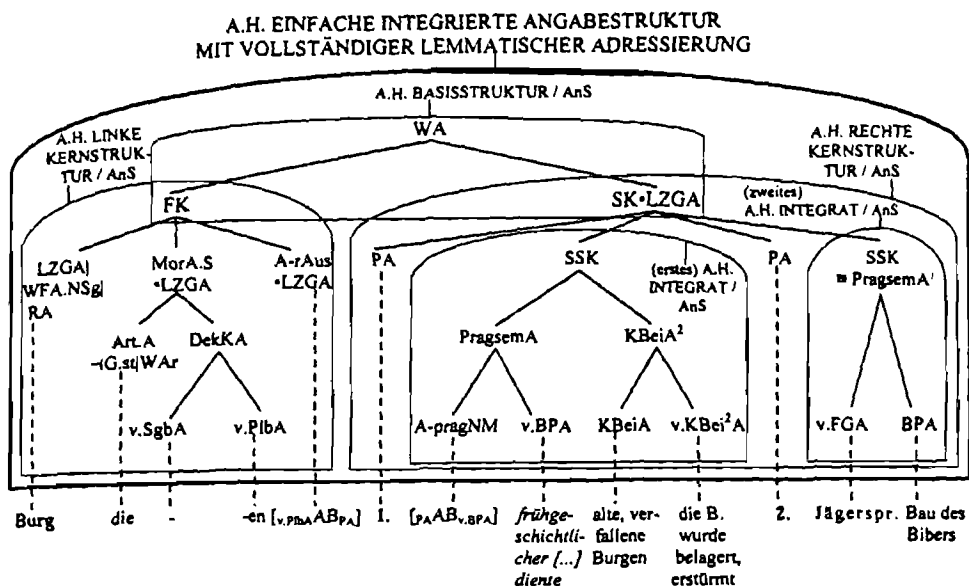
Für nichtelementare mikrostrukturelle Textkonstituenten, wie z.B. Formkommentare, semantische Kommentare, Morphologieangaben bei Substantiven usw. kann gelten, daß sie entweder *adressenhomogen* oder *adressenheterogen* sind. Eine nichtelementare mikrostrukturelle Textkonstituente ist adressenhomogen genau dann, wenn sie selbst und alle ihre Teilangaben an die gleiche Adresse adressiert sind; ist das nicht der Fall und kommen n Adressen (mit $n \geq 2$) vor, ist die mikrostrukturelle Textkonstituente adressenheterogen. In wa_3 ist z.B. die Morphologieangabe bei Substantiven (MorA.S) und der semantische Kommentar (SK) adressenhomogen. Der Formkommentar (FK) in wa_3 ist dagegen adressenheterogen, da er selbst nicht, wie alle seine Teilangaben, an die Lemmazeichengestaltangabe adressiert ist. Um die Darstellung der hierarchischen Angabestrukturen möglichst übersichtlich gestalten zu können, kann daher von folgender Notationskonvention Gebrauch gemacht werden: Ist eine nichtelementare mikrostrukturelle Textkonstituente T_i adressenhomogen, wird bei der Darstellung der abstrakten hierarchischen Mikrostruktur nur das Angabeklassensymbol für T_i um das Adressensymbol erweitert. Dies besagt, daß alle Teilangaben von T_i die gleiche Adresse haben wie T_i . In der Abb. 16 ist von dieser Notationskonvention Gebrauch gemacht.

Die einführende Übersicht zur Adressierung in der einsprachigen Lexikographie sei hier nicht ohne den Hinweis abgebrochen, daß aus Platzgründen u.a. (i) zahlreiche Detailprobleme nicht berücksichtigt wurden, daß (ii) eine Analyse eines Artikels mit standardisierter und nichtstandardisierter nichtlemmatischer Adressierung nicht vorgenommen wurde, daß (iii) weiterhin die Adressierung in tabellarischen Wörterbuchartikeln (i.S.v. Wiegand 2000d) ausgeblendet wurde und daß schließlich (iv) zur Adressierung bei Verweisangaben nur einige Hinweise gegeben wurden.

3. Adressierung in der zweisprachigen Lexikographie

In einem zweisprachigen Wörterbuchartikel kommt es wesentlich darauf an, daß wenigstens eine Äquivalentbeziehung vom Lemmazeichen der lexikographischen Ausgangssprache (A) zu einer (möglichst lexikalisierten) Einheit der

Abb. 16: Strukturgraph mit vereinfachter Etikettierung zur (abstrakten und zur isomorphen konkreten) hierarchischen einfachen integrierten Angabestruktur mit vollständiger lemmatischer Adressierung, die wa₃ aufweist



lexikographischen Zielsprache (Z) etabliert wird. Die Äquivalentbeziehung kann mit dem zweistelligen Relationsterm $R_a = x \text{ ist äquivalent mit } y$ benannt werden, wobei „x“ eine Variable für ein ausgangssprachliches Lemmazeichen und „y“ eine Variable für ein zielsprachliches Äquivalent ist. In vielen einfachen kleineren zweisprachigen Wörterbüchern (wie z.B. Küen 1993; Prinsloo und Sathekge 1996) gibt es Hunderte von Wörterbuchartikeln, zu welchen der Volltext (neben dem Volltexttitel) aus gerade einem Satz besteht, der durch die Belegung der Variablen in R_a gebildet werden kann. Solche Wörterbuchartikel sind z.B. wa₆ bis wa₁₅ aus Prinsloo und Sathekge (1996) in Abb. 17.

Abb. 17: wa₆ bis wa₁₅ aus Prinsloo und Sathekge (1996)

wa ₆ :	aim	lebanya	wa ₁₁ :	boy	mošimane
wa ₇ :	air	môya	wa ₁₂ :	brandy	poranti
wa ₈ :	altar	altare	wa ₁₃ :	card	karata
wa ₉ :	arrest	swara	wa ₁₄ :	professor	moprôfesa
wa ₁₀ :	bathroom	bohlapêlô	wa ₁₅ :	pronoun	lešala

Der genannte Volltextsatz zu wa₆ lautet:

[6] aim ist äquivalent mit lebanya.

Durch Textverdichtung wird nach dem in 2.2 erläuterten Verfahren aus [6] der zweisprachige Artikel wa_6 für die anderen Beispiele gilt dies entsprechend. Zweisprachige Bearbeitungseinheiten wie beispielsweise $\langle aim, lebanya \rangle$, $\langle card, karata \rangle$ sind in zweisprachigen Wörterbuchartikeln unverzichtbar. Auf ein 2-Tupel wie $\langle aim, lebanya \rangle$ als Element der Bearbeitungsrelation trifft der Relationsterm $R_b = x$ ist bearbeitet durch y zu, so daß gilt:

(6') *aim* ist bearbeitet durch *lebanya*.

Mit *aim* in (6') ist nicht die Lemmazeichengestaltangabe *aim* gemeint, denn es ist nicht die LZGA, die in wa_6 mit „lebanya“ bearbeitet ist, sondern das Lemmazeichen *aim* und mit ihm das gesamte Lemmazeichenparadigma (also der Adressat). Auf $\langle aim, lebanya \rangle$ trifft weiterhin der Relationsterm $R_{kAD} = y$ ist die Adresse zu x zu, so daß gilt:

(6'') *aim* ist die Adresse zu *lebanya*.

Im Unterschied zu (6') ist mit *aim* in (6'') die Lemmazeichengestaltangabe gemeint. Der Sachverhalt, daß das 2-Tupel $\langle aim, lebanya \rangle$ (und entsprechend alle anderen 2-Tupel) sowohl ein Element der Bearbeitungsrelation als auch ein Element der konversen Relation zur Adressierungsrelation sein kann (bzw. sein können), hat zu erheblichen Konfusionen geführt (wie man z.B. bei Hausmann und Werner 1991 sehen kann).

Klar ist dagegen, daß in wa_6 die Wortäquivalentangabe „lebanya“ an die Lemmazeichengestaltangabe *aim* adressiert ist, so daß mit wa_6 gerade das ausgesagt wird, was mit [6] auch ausgesagt wird, nämlich der propositionale Gehalt $p =$ „daß *aim* und *lebanya* äquivalent sind“.

Wörterbuchartikel wie wa_6 bis wa_{15} wollen nur eine erste Orientierung geben. Sie erwecken allerdings den falschen Eindruck, als gelte die behauptete Äquivalentbeziehung für alle Fälle des Gebrauchs von *aim*, *air* usw. Erfahrungsgemäß ist dies aber nicht der Fall. Vielmehr wirken sich ko- und kontextuelle Bedingungen auf die Äquivalentbeziehung aus. Die meisten Äquivalentbeziehungen gelten daher bedingt (= eingeschränkt durch Bedingungen) und nicht unbedingt (= uneingeschränkt), wie man nach wa_6 bis wa_{15} glauben könnte.

Damit ist der erste wichtige Sachverhalt angesprochen, der dafür verantwortlich ist, daß die Adressierungsbeziehungen selbst in kurzen zweisprachigen Wörterbuchartikeln relativ komplex und z.T. schwer durchschaubar sein können. Denn um eine Äquivalentbeziehungen eindeutig identifizieren zu können, müssen sowohl beim Ausgangssprachlichen Lemmazeichen als auch beim Zielsprachlichen Äquivalent bestimmte Eigenschaftsausprägungen berücksichtigt werden; dies bedeutet, daß gegebenenfalls A-intern lemmatisch adressierte und Z-intern nichtlemmatisch adressierte Angaben zu diesen Eigenschaftsausprägungen erforderlich werden. Mit solchen Angaben werden

z.T. auch Bedingungen gemacht, unter denen die Äquivalentbeziehung gilt. Daraus ergibt sich: Im Wörterbuchartikel müssen *wenn-dann*-Beziehungen, die im Wörterbuchgegenstandsbereich gegeben sind, lexikographisch vertextet sein, so daß bestimmte Angaben nur korrekt sind, wenn andere Angaben korrekt sind. Dieser Zusammenhang wirkt sich auch auf die Adressierung und ihr theoretisches Verständnis aus, was nun am Beispiel von wa_{16} in Abb. 18 erläutert werden soll.

Abb. 18: wa_{16} aus Duden-Oxford (1990)

bestatten [bɛˈStatɪ] *tr. V. (geh.) inter (formal)*; **bury**: **bestattet werden** *be laid to rest*

Im Anschluß an Hausmann und Werner (1991: 2732) weist wa_{16} folgende Bearbeitungseinheiten auf:

- (1) Zweisprachige Bearbeitungseinheiten
 - <bestatten, inter>
 - <bestatten, bury>
 - <bestattet werden, be laid to rest>
- (2) Einsprachige Bearbeitungseinheiten
 - (2.1) Ausgangssprachliche Bearbeitungseinheiten
 - <bestatten, bɛˈStatɪ>
 - <bestatten, tr. V.>
 - <bestatten, geh.>
 - (2.2) Zielsprachliche Bearbeitungseinheit
 - <inter, formal>

In wa_{16} finden sich mithin sieben Bearbeitungseinheiten, so daß die auf der Menge aller Angaben von wa_{16} definierte Bearbeitungsrelation sieben 2-Tupel als Elemente aufweist. Entsprechend weist nun die Adressierungsrelation für wa_{16} — sie heiße R_{Ad} — sieben 2-Tupel auf und kann wie folgt extensional bestimmt werden:

$R_{Ad} = \{ \langle \text{inter, bestatten} \rangle, \langle \text{bury, bestatten} \rangle, \langle \text{be laid to rest, bestattet werden} \rangle, \langle \text{bɛˈStatɪ, bestatten} \rangle, \langle \text{tr. V., bestatten} \rangle, \langle \text{geh., bestatten} \rangle, \langle \text{formal, inter} \rangle \}$. Auf alle 2-Tupel trifft der Relationsterm x ist adressiert an y zu.

Um die Darstellung nicht zu überlasten, berücksichtigen wir im folgenden nur folgenden Ausschnitt von wa_{16} :

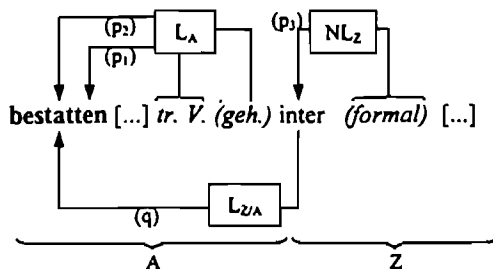
e_4 : **bestatten** [...] *tr. V. (geh.) inter (formal)* [...]

Zu diesem Ausschnitt bilden wir nun folgenden partiellen Volltext (pVt₂):

pVt₂: [p₁] *bestatten* ist ein transitives Verb.
 [p₂] *bestatten* wird gehoben verwendet.
 [p₃] *inter* wird offiziell (\approx *formal*) verwendet.
 [q] *inter* ist Wortäquivalent zu *bestatten*.

pVt₂ besteht aus vier Aussagen, wobei q die Aussage über die Äquivalentbeziehung ist; q gilt aber nur dann, wenn p₁, p₂ und p₃ gelten. Wir können mithin feststellen: Wenn *bestatten* ein transitives Verb ist, und wenn es gehoben verwendet wird, und wenn *inter* offiziell verwendet wird, dann sind *bestatten* und *inter* Wortäquivalente. In Abb. 19 werden nun die Adressierungsbeziehungen in e₄ veranschaulicht.

Abb. 19: Veranschaulichung zu einigen Adressierungsbeziehungen in e₄



Abkürzungen: A = lexikographische Ausgangssprache (hier: Dt.); Z = lexikographische Zielsprache (hier: Engl.); L_A = A-intern lemmatisch adressiert; NL_Z = Z-intern nichtlemmatisch adressiert; L_{Z/A} = von Z nach A lematische adressiert; **Erläuterungen:** „x \xrightarrow{u} y“ bedeutet soviel wie *x ist artikelintern u-adressiert an y mit „u“ als Variable für L_A, NL_Z und L_{Z/A}.*

Wir können nun feststellen: Daß die Adressierung von „inter“ an „bestatten“ (also die lematische Adressierung von Z nach A) korrekt ist, hat zur Bedingung, daß (i) die Adressierung von „tr. V.“ und von „geh.“ an *bestatten* (also die A-interne lematische Adressierung) korrekt ist und weiterhin, daß (ii) die Adressierung von „formal“ an „inter“ (also die Z-interne nichtlemmatische Adressierung) korrekt ist.

Damit können wir folgende beiden Arten von Adressierungsbeziehungen unterscheiden:

- bedingende Adressierungsbeziehungen, und
- bedingte Adressierungsbeziehungen.

Die Adressierungsbeziehung „inter ist von Z nach A lemmatisch an bestatten adressiert“ ist eine *bedingte Adressierungsbeziehung*. Alle anderen Adressierungsbeziehungen in e_4 sind *bedingende Adressierungsbeziehungen*.

Alle Adressierungsbeziehungen in wa_6 bis wa_{15} (vgl. Abb. 17) sind dagegen *unbedingte Adressierungsbeziehungen*, weil keine ausgangssprachlichen und/oder zielsprachlichen Bedingungen dafür angegeben sind, daß die Äquivalentbeziehungen korrekt sind. Auch in einsprachigen Wörterbuchartikeln — das sei hier ergänzend zu den Ausführungen in 2. erwähnt — können bedingte und bedingende Adressierungsbeziehungen auftreten. Dies sei nachfolgend kurz anhand des HWDG-Artikels wa_{17} in Abb. 20 erläutert.

Abb. 20: wa_{17} aus dem HWDG

Moos, das; -es, -e 1. *kleine, in großer, dicht stehender Menge, vorzugsweise an schattigen, feuchten Stellen wachsende immergrüne, sich durch Sporen vermehrende Pflanze: weiches, grünes M.; Moose und Flechten; sich im Wald ins M., auf das M. setzen – 2. /o. Pl./ u m g. Geld (1): da muß doch eine Masse M. herausspringen*

Im zweiten semantischen Subkommentar von wa_{17} gelten u.a. folgende Adressierungsbeziehungen:

- (i) *Geld* ist lemmatisch an **Moos** adressiert.
- (ii) „o. Pl.“ ist lemmatisch an **Moos** adressiert.
- (iii) „u m g.“ ist lemmatisch an **Moos** adressiert.

Die Bedingungsverhältnisse werden anhand folgender Aussagen deutlich:

- (a) Wenn *Moos* soviel bedeutet wie *Geld*, dann wird *Moos* ohne Plural und umgangssprachlich verwendet.
- (b) Wenn *Geld* an **Moos** adressiert ist, dann sind „o. Pl.“ und „u m g.“ an **Moos** adressiert.

Die Adressierungsbeziehung (i) ist daher bedingend; die Adressierungsbeziehungen (ii) und (iii) sind bedingt.⁷

Ein zweiter Sachverhalt, warum Adressierungsbeziehungen in zweisprachigen Wörterbüchern komplex sind, besteht darin, daß es zu einem ausgangssprachlichen Lemmazeichen mehrere zielsprachliche Äquivalente geben kann, so daß mit Äquivalentunterscheidungsangaben ihre Bedeutungsunterschiede angegeben werden müssen. Wie die Adressen von Äquivalentunterscheidungsangaben sind, ist vom Benutzer häufig schwer zu ermitteln (vgl. Meyer und Wiegand 2000). Auch eine saubere (vgl. wa_4 , 2.3!) theoretische

Lösung macht große Schwierigkeiten. Die folgende partielle Analyse anhand von wa_{18} in Abb. 21 ist daher als Versuch einzuschätzen (vgl. auch Wiegand 2000e mit weiteren Versuchen).

Abb. 21: wa_{18} aus Slabý et al. 1994

blau adj *azul* || [hell-] *celeste*, <poet> *cerúleo*
 || [Heraldik] *azur* || [Auge &] *garzo* || [vor Kälte]
livido, *amorado* || [Traube] *tinto* || <fig>
 [betrunken] *borracho* || \diamond ~ anlaufen lassen
 [Stahl] *pavonar* || ~ färben *teñir de azul* od *antil*,
azular || ~ unterlaufene Augen *ojos* mpl *lividos*
 || [Augenringe] *ojerds* fpl

Die Äquivalentunterscheidungsangaben (ÄUntA) in wa_{18} sind daran erkennbar, daß sie in eckigen Klammern stehen. Äquivalentunterscheidungsangaben können unterschiedlich realisiert sein, z.B. als Fachgebietsangabe (wie „Heraldik“) oder als Bezugsobjektangabe (wie „Traube“). Meistens sind sie Angaben aus anderen Angabeklassen in äquivalentunterscheidender Funktion oder sie werden vom Lexikographen ad hoc gebildet. In wa_{18} unterscheiden sie für den deutschen Benutzer n (mit $n \geq 2$) in Äquivalentangaben erwähnte spanische Äquivalente zu einem deutschen Lemmazeichen.

Ich argumentiere nachfolgend anhand folgenden Ausschnittes aus wa_{18} :

e_5 : blau [...] [Traube] *tinto* [...]

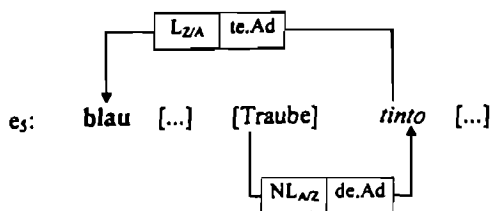
Wenn „Traube“ die adressierte Angabe ist, gibt es für die hier fraglichen Adressierungsbeziehungen in e_5 drei im Prinzip mögliche Fälle:

- (1) „Traube“ ist nichtlemmatisch nach rechts an die Wortäquivalentangabe *tinto* adressiert, so daß eine nichtlemmatische Rechtsadressierung von der Ausgangs- in die Zielsprache vorliegt.
- (2) „Traube“ ist lemmatisch nach links an die Lemmzeichengestaltangabe blau adressiert, so daß eine lemmatische ausgangsspracheninterne (= A-interne) Linksadressierung gegeben ist.
- (3) „Traube“ ist sowohl an blau als auch an *tinto* adressiert. In diesem Falle hätte man eine „Doppeladressierung“ nach (1) und (2).

Eine Veranschaulichung der Adressierungsbeziehung in e_5 nach Fall (1) findet sich in Abb. 22.

Mit Bezug auf den Wörterbuchgegenstandsbereich von Slabý et al. (1994) (also Deutsch und Spanisch) kann e_5 (Fall 1) wie folgt interpretiert werden: Wenn *tinto* auf das Bezugsobjekt Traube(n) bezogen ist, dann ist das Wortäquivalent zu *tinto* das Lemmazeichen blau. Dem entspricht auf der Ebene der Wörterbuchform: Wenn „Traube“ an *tinto* adressiert ist, dann ist *tinto* an blau

Abb. 22: Erste Veranschaulichung zu den Adressierungsbeziehungen in e_5 (Fall 1)

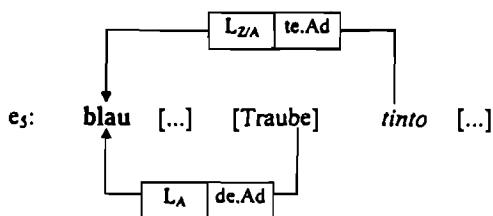


Abkürzungen: **te.Ad** = bedingte Adressierung; **de.Ad** = bedingende Adressierung; $L_{Z/A}$ = von der Zielsprache (Z) nach der Ausgangssprache (A) lemmatisch adressiert; $NL_{A/Z}$ = von A nach Z nichtlemmatisch adressiert. Erläuterungen: „ $x \xrightarrow{u|v} y$ “ bedeutet soviel wie x ist adressiert an y , wobei „ u “ und „ v “ Variablen für Bezeichnungen von Adressierungsarten sind.

adressiert. Oder anders ausgedrückt: Wenn **tinto** Adresse von „Traube“ ist, dann ist **blau** Adresse von **tinto**.

Eine Veranschaulichung der Adressierungsbeziehung in e_5 nach Fall (2) findet sich in Abb. 23.

Abb. 23: Zweite Veranschaulichung zu den Adressierungsbeziehungen in e_5 (Fall 2)



Im Hinblick auf den Wörterbuchgegenstandsbereich kann e_5 (Fall 2) wie folgt gedeutet werden: Wenn **blau** auf das Bezugsobjekt Traube(n) bezogen ist, dann ist **tinto** ein Wortäquivalent zu **blau**. Dem entspricht auf der Ebene der Wörterbuchform: Wenn „Traube“ an **blau** adressiert ist, dann ist auch **tinto** an **blau** adressiert. Oder anders gesagt: Wenn **blau** die Adresse von „Traube“ ist, dann ist **blau** auch die Adresse von **tinto**.

Der 3. Fall wäre eine Kombination der beiden ersten Fälle und seine Interpretation entsprechend wenig sinnvoll.

Damit haben wir nun das große Rätsel, welche der beiden Interpretationen nach welchem Kriterium als die angemessene gelten kann. Das Kriterium kann nur der genuine Zweck eines zweisprachigen Wörterbuchs sein (vgl. Wiegand 1996: 42ff.) und damit die für das Wörterbuch vorgesehenen Wörterbuchfunk-

tionen (i.S.v. Tarp 1994). In Slabý et al. (1994) findet man dazu allerdings nichts Nennenswertes. Daher argumentiere ich im folgenden zwar anhand von wa_{18} , aber ohne Bezug zu den Umtexten von Slabý et al. (1994); dies ist deswegen vertretbar, weil es hier um systematische Zusammenhänge geht und nicht um ein bestimmtes zweisprachiges Wörterbuch.⁸

Wir betrachten im folgenden drei Fälle.

Fall a: wa_{18} steht in einem Wörterbuch Wb_1 , für das gilt:

- (i) Wb_1 ist monoakzessiv.
- (ii) Wb_1 ist für deutsche Benutzer konzipiert.
- (iii) Wb_1 ist ein sog. aktives Wörterbuch, so daß die Produktions- und die Hinübersetzungsfunktion als primäre Wörterbuchfunktionen gelten.

Angenommen, ein Deutscher möchte den Satz

(7) Blaue Trauben passen gut zum Käse

ins Spanische übersetzen, weiß aber nicht, was *blau* auf Spanisch heißt und schlägt in Wb_1 nach, dann muß er e_5 lesen wie: Wenn *blau* auf Trauben bezogen ist, dann heißt es auf Spanisch *tinto*. Damit gilt für Fall a die zweite Interpretation der Adressierungsbeziehungen.

Fall b: wa_{18} steht in einem Wörterbuch Wb_2 , für das gilt:

- (i) Wb_2 ist monoakzessiv.
- (ii) Wb_2 ist für spanische Benutzer konzipiert.
- (iii) Wb_2 ist ein sog. passives Wörterbuch, so daß die Rezeptions- und die Herübersetzungsfunktion als primäre Wörterbuchfunktionen gelten.

Will ein Spanier (7) ins Deutsche übersetzen und weiß nicht, was *blau* bedeutet, dann gilt die gleiche Interpretation wie in Fall a. (Es ist klar, daß in einem Wörterbuch für spanische Benutzer die Äquivalentunterscheidungsangabe in Spanisch stehen müßte.)

Fall c: wa_{18} steht in einem Wörterbuch Wb_3 , für das gilt:

- (i) Wb_3 ist extern polyakzessiv (i.S.v. Wiegand 1996: 47) und weist ein span.-dt. alphabetisches Äquivalentregister auf mit Registereinträgen wie „*tinto blau*“.
- (ii) Wb_3 ist primär für deutsche und sekundär für spanische Benutzer konzipiert.
- (iii) Wb_3 ist in erster Linie ein aktives Wörterbuch für Deutsche (wie Wb_1 , (iii)) und in zweiter Linie ein aktives Wörterbuch für Spanier.

Will ein Spanier in einem spanischen Text, in dem *tinto* auf das bezogen ist, was Deutsche Trauben nennen, ins Deutsche übersetzen und kennt kein deut-

sches Äquivalent, dann kann er zunächst im Register von Wb_3 nachschlagen und dann unter **blau**. Für diesen Fall gilt offensichtlich die erste Interpretation.

In den Benutzungshinweisen eines zweisprachigen Wörterbuchs vom Typ Wb_1 muß demnach ein Wörterbucheintrag wie e_3 wie folgt erläutert werden:

Wenn *Blau* sich auf Trauben bezieht, dann heißt es auf Spanisch *tinto*. In einem Wörterbuch vom Typ Wb_2 muß das gleiche auf Spanisch stehen. Schließlich muß in den Benutzungshinweisen eines Wörterbuchs vom Typ Wb_3 für die deutschen Benutzer das gleiche stehen wie in Wb_1 , und für die spanischen Benutzer, die über das erweiterte Äquivalentregister zugreifen möchten, muß es (auf Spanisch) heißen: Wenn *tinto* auf Trauben bezogen ist, heißt es auch Deutsch *blau*.

Die gerade gegebenen pauschalen Hinweise zu den Benutzungshinweisen in zweisprachigen Wörterbüchern sind zu verstehen als ein Plädoyer, die Adressierungsbeziehungen anhand von Beispielen für Wörterbucheinträge und durch die Wiedergabe der *wenn-dann*-Beziehungen zu erklären. Der derzeitige Zustand, daß in vielen zweisprachigen Wörterbüchern die Adressierungsbeziehungen meistens nicht erläutert werden, ist — so sollte nach den Ausführungen in diesem Beitrag wohl deutlich geworden sein — nicht im Sinne der Benutzer.

4. Ausblick

Hochverdichtete und stark standardisierte Wörterbuchartikel in ein- und zweisprachigen Wörterbüchern stellen keine natürlichsprachlichen Texte dar (vgl. auch näher Wiegand 2000f.). Sie bilden vielmehr ein hochkomplexes Ineinander von formaler Syntax und der Semantik derjenigen Sprach(en), die den Wörterbuchgegenstandsbereich bilden. Die Syntax hat zwei Aufgaben: sie ordnet die Konstituenten (Angaben) linear in eine bestimmte Reihenfolge und sie stellt Beziehungen zwischen den Konstituenten her. Letzteres wird als Adressierung analysiert. Da immer vorausgesetzt werden muß, daß der Benutzer das, was er sucht, nicht kennt und weiterhin, daß er ganze Teile der lexikographisch bearbeiteten Sprache nicht kennt, so daß er die Konstituenten nicht aufgrund ihrer Bedeutungen aufeinander beziehen kann, muß der Artikeltext so gestaltet sein, daß man nicht nur erkennen kann, welche Artikelpositionen aufeinander folgen und beim suchenden Lesen gerade gegeben ist, sondern es muß anhand von Eigenschaften der Artikelform auch jeweils klar sein, an welche Adresse eine Angabe adressiert ist, weil sonst nicht sichergestellt ist, daß ein Benutzer lexikographische Informationen richtig erschließen kann. Dies gilt besonders für allgemeine zweisprachige Wörterbücher. Das heißt, daß in den Benutzungshinweisen nicht nur die mikrostrukturelle Konstituenz von Wörterbuchartikeln, sondern auch die Angabestruktur und damit die Adressierung erläutert werden muß. Es ist daher ein Desiderat, herauszufinden, wie dies am einfachsten geschehen kann.

5. Endnoten

1. Es ist geplant, in dem mehrsprachigen *Wörterbuch zur Lexikographie und Wörterbuchforschung/Dictionary of Lexicography and Dictionary Research* (das voraussichtlich herausgegeben wird von Rufus H. Gouws, Henning Bergenholtz, Matthias Kammerer, Angelika Storrer, Sven Tarp, Herbert Ernst Wiegand und Werner Wolski) die metalexikographische Terminologie und lexikographische Werkstattsprache möglichst exhaustiv zu bearbeiten. Die derzeit noch vorläufige Lemmakandidatenliste enthält für den Bereich der Adressierung folgende Lemmakandidaten: *Adresse; Adresse, artikelinterne/artikelexterne/nichtstandardisierte/standardisierte/wörterbuchexterne; Adressenfundort; Adressensymbol; Adressierung; Adressierung, bedingte/bedingende/unbedingte/ infralemmatische/lemmatische/ nichtlemmatische/sublemmatische/nichtstandardisierte/standardisierte; Adressierungsrelation; Adressierungsstruktur, Außentextadresse, Nahadressierung, Fernadressierung, Glossat, Glossatadresse, Präglossat, Binnenglossat, Postglossat, Umtextadresse, Verweisadresse, Verweisaußenadresse*. Die meisten dieser Termini (und einige weitere) werden nachfolgend erläutert.
2. Die zentralen Teile einer Theorie der Wörterbuchform liegen inzwischen vor; vgl. u.a. Wiegand (1988, 1989, 1989a, 1989b, 1990, 1991, 1993 [1994], 1995, 1995a, 1996, 1996a, 1996b, 1996c, 1997, 1998, 1998a, 1998b, 1999, 1999a, 2000, 2000b, 2000d, 2000e); Bergenholtz, Tarp und Wiegand (1999); Kammerer und Wiegand (1998); Konerding und Wiegand (1995); Meyer und Wiegand (2000); vgl. auch: http://www.uni-heidelberg.de/institute/fak9/gs/sprache2/hew_publ.htm.
3. Vgl. die in Anm. 2 genannte Literatur.
4. Für eine durchgehende Erläuterung aller Segmentationssschritte einer funktional-positionalen Segmentation vgl. z.B. Wiegand 2000: 235ff.
5. Zu weiteren Eigenschaften, die den Sonderstatus der Lemmazeichengestaltangabe begründen, vgl. Wiegand 2000e.
6. Mit den Glossaten sind eine Reihe von Segmentationsproblemen verbunden. Diese lassen sich wahrscheinlich lösen, wenn das in Meyer und Wiegand (2000: 147) angedeutete Konzept der gebundenen Angaben erweitert und präzisiert wird, was hier nicht möglich ist.
7. Mit der Festlegung der Bedingungsverhältnisse sind sprachtheoretische Voraussetzungen gegeben, worauf hier nicht näher eingegangen werden kann.
8. Zu Slabý et al. (1994) kann im übrigen Meyer und Wiegand (2000: 114ff.) verglichen werden.

6. Literatur

- Bergenholtz, H. und S. Tarp (Red.). 1995. *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Mit Beiträgen von G. Duvá, A.-L. Laursen, S. Nielsen, O. Norling-Christensen und J. Pedersen. Amsterdam/Philadelphia: John Benjamins.
- Bergenholtz, H., I. Contell et al. 1997. *Nordisk Leksikografisk Ordbok*. Mit Beiträgen von H. Haraldsson, H.K. Mikkelsen und J. Sivula. Skrifter utgitt av Nordisk forening for leksikografi. skrift nr. 4. Oslo: Universitetsforlaget AS.
- Bergenholtz, H., S. Tarp und H.E. Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. *HSK 14.2*: 1762-1832.

- Burkhanov, I. 1998. *Lexicography. A Dictionary of Basis Terminology*. Rzeszów: Wyzszej szkoły pedagogicznej.
- BW** = *Brockhaus-Wahrig: Deutsches Wörterbuch in sechs Bänden*. 1980-1984. Hrsg. v. Gerhard Wahrig†, Hildegard Krämer und Harald Zimmermann. 1. Bd.: A-BT 1980. 2. Bd.: BU- FZ 1981. 3. Bd.: G-JZ 1981. 4. Bd.: K-OZ 1982. 5. Bd.: P-STD 1983. 6. Bd.: STE-ZZ 1984. Wiesbaden: Brockhaus.
- Campe-wdS** = Joachim Heinrich Campe. 1969. *Wörterbuch der Deutschen Sprache*. I. A-E. Mit einer Einführung und Bibliographie von Helmut Henne. Documenta Linguistica. Reihe II. Wörterbücher des 17. und 18. Jahrhunderts. Hildesheim/New York. [Reprogr. Nachdruck der Ausgabe Braunschweig 1807].
- Duden-Oxford 1990** = *Duden-Oxford. Großwörterbuch Englisch. Englisch-Deutsch/Deutsch-Englisch*. Hrsg. v. der Dudenredaktion und Oxford University Press. Red. Leitung: Werner Scholze-Stubenrecht und John Sykes. Mannheim [...]: Dudenverlag 1990.
- Glück, H.** (Red.). 2000. *Metzler Lexikon Sprache*. 2. überarb. und erw. Aufl. mit 70 Abbildungen, davon 17 Karten. Stuttgart/Weimar: J.B. Metzler.
- Hartmann, R.R.K. und G. James**. 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- HSK 5.1** = *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Erster Teilbd. 1989. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand und Ladislav Zgusta. Handbücher zur Sprach- und Kommunikationswissenschaft 5.1. Berlin/New York: De Gruyter.
- HSK 5.2** = *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Zweiter Teilbd. 1990. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand und Ladislav Zgusta. Handbücher zur Sprach- und Kommunikationswissenschaft 5.2. Berlin/New York: De Gruyter.
- HSK 5.3** = *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Dritter Teilbd. 1991. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand und Ladislav Zgusta. Handbücher zur Sprach- und Kommunikationswissenschaft 5.3. Berlin/New York: De Gruyter.
- HSK 14.2** = *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologie-wissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research*. 2. Halbbd. 1999. Hrsg. v. Lothar Hoffmann, Hartwig Kalverkämper und Herbert Ernst Wiegand. In Verbindung mit Christian Galinski und Werner Hüllen. Handbücher zur Sprach- und Kommunikationswissenschaft 14.2. Berlin/New York: De Gruyter.
- HWdG** = *Handwörterbuch der deutschen Gegenwartssprache. In zwei Bänden*. 1984. Von einem Autorenkollektiv unter der Leitung v. Günter Kempcke [...]. Bd. 1: A-K. Bd. 2: L-Z. Berlin: Akademie Verlag.
- Kammerer, M. und H.E. Wiegand**. 1998. Über die textuelle Rahmenstruktur von Printwörterbüchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-238.
- Kluge, F.** 1995. *Etymologisches Wörterbuch der deutschen Sprache*. Bearb. v. Elmar Seebold. 23. erw. Aufl. Berlin/New York: De Gruyter.
- Küen, E.** 1993. *Deutsch-Niederländisches Lernwörterbuch*. Hamburg: Buske.
- Landau, S.I.** 1984. *Dictionaries. The Art and Craft of Lexicography*. New York: Charles Scribner's Sons.

- Martínez de Sousa, J. 1995. *Diccionario de lexicografía práctica*. Barcelona: Bibliograf, S.A.
- Meyer, M. und H.E. Wiegand. 2000. Gemischt-semiintegrierte Mikrostrukturen für deutsch-spanische Printwörterbücher. Wiegand, H. E. (Red.). 2000. *Studien zur zweisprachigen Lexikographie mit Deutsch V*: 87-171. Germanistische Linguistik 151-152. Hildesheim/New York: Olms.
- Pan Zaiping und H.E. Wiegand. 1995. Über die Musterartikel für das Große Deutsch-Chinesische Wörterbuch. Zugleich ein Beitrag zu einer Theorie zweisprachiger lexikographischer Texte. Wiegand, H. E. (Red.). 1995. *Studien zur zweisprachigen Lexikographie mit Deutsch II*: 63-190. Germanistische Linguistik 127-128. Hildesheim/New York: Olms.
- Prinsloo, D.J. und B.P. Sathekge. 1996. *New Sepedi Dictionary. English-Sepedi (Northern Sotho); Sepedi (Northern Sotho)-English*. Pietermaritzburg/Kapstadt/Randburg: Shuter & Shooter.
- Rey-Debove, J. 1971. *Étude linguistique et sémiotique des dictionnaires français contemporains*. Approaches to Semiotics 13. Den Haag/Paris: Mouton.
- Slabý, R.J., R. Grossmann und C. Illig. 1994. *Wörterbuch der spanischen und deutschen Sprache*. 4. Aufl. Wiesbaden: Horrasowitz.
- Svensén, B. 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Aus dem Schwedischen übers. v. John Sykes und Kerstin Schofield. Oxford/New York: Oxford University Press.
- Tarp, S. 1994. Funktionen in Fachwörterbüchern. Schaedter, B. und H. Bergenholtz (Red.). *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 229-246. Forum für Fachsprachen-Forschung 23. Tübingen: Narr.
- Wiegand, H.E. 1988. Wörterbuchartikel als Text. Harras, G. (Red.). 1988. *Das Wörterbuch. Artikel und Verweisstrukturen. Jahrbuch 1987 des Instituts für deutsche Sprache*: 30-120. Sprache der Gegenwart LXXIV. Düsseldorf: Schwann. [Auch in: Wiegand, H.E. 2000a: 877-950].
- Wiegand, H.E. 1989. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: alphabetische Anordnungsformen und ihre Probleme. *HSK 5.1*: 371-409.
- Wiegand, H.E. 1989a. Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. *HSK 5.1*: 409-462.
- Wiegand, H.E. 1989b. Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. *HSK 5.1*: 462-501.
- Wiegand, H.E. 1990. Printed Dictionaries and their Parts as Texts. An Overview of More Recent Research as an Introduction. *Lexicographica 6*: 1-126. [Auch in: Wiegand, H.E. 2000a: 951-1062].
- Wiegand, H.E. 1991. Über die Strukturen der Artikeltexte im Frühneuhochdeutschen Wörterbuch. Zugleich ein Versuch zur Weiterentwicklung einer Theorie lexikographischer Texte. Goebel, U. und O. Reichmann (Red.). 1991. *Historical Lexicography of the German Language. Vol. 2*: 341-673. Studies in Russian and German 3. Lewiston/Queenston/Lampeter: Edwin Mellen Press.
- Wiegand, H.E. 1993 [1994]. Kritische Lanze für Fackel-Redensartenwörterbuch. Bericht und Diskussion zu einem Workshop in der Österreichischen Akademie der Wissenschaften am 14.2.1994. *Lexicographica 9*: 230-271. [Auch in: Wiegand, H.E. 2000a: 1063-1105].
- Wiegand, H.E. 1995. Lexikographische Texte in einsprachigen Lernerwörterbüchern. Kritische Überlegungen anlässlich des Erscheinens von Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Popp, H. (Red.). 1995. *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag*: 463-499. München: Iudicium. [Auch in: Wiegand, H.E. 2000a: 1128-1162].

- Wiegand, H.E. 1995a. Deutsch-Türkmenisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Türkmenischen Staatlichen Magtymguly-Universität in Aschghabat. *Lexicographica* 10: 249-300.
- Wiegand, H.E. 1996. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. Wiegand, H.E. (Red.) 1996. *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 11-43. *Lexicographica. Series Maior* 76. Tübingen: Max Nieweyer. [Auch in: Wiegand, H.E. 2000a: 1193-1277].
- Wiegand, H.E. 1996a. A Theory of Lexicographic Texts. An Overview. *South African Journal of Linguistics* 14(4): 134-149.
- Wiegand, H.E. 1996b. Deutsch-Uzbekisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Staatlichen Usbekischen Weltsprachen-Universität in Taschkent. *Lexicographica* 12: 190-254. [Auch in: Wiegand, H.E. 2000a: 1339-1409].
- Wiegand, H.E. 1996c. Über die Mediostrukturen bei gedruckten Wörterbüchern. Zettersten, A. and V. Hjørnager Pedersen (Red.). 1996. *Symposium on Lexicography VII. Proceedings of the Seventh Symposium on Lexicography May 5-6, 1994 at the University of Copenhagen*: 11-43. *Lexicographica. Series Maior* 76. Tübingen: Niemeyer. [Auch in: Wiegand, H.E. 2000a: 1163-1192].
- Wiegand, H.E. 1997. Printed Language Dictionaries and their Standardization: Notes on the Progress toward a General Theory of Lexicography. Hock, H.H. (Red.). 1997. *Historical, Indo-European, and Lexicographical Studies. A Festschrift for Ladislav Zgusta on the Occasion of his 70th Birthday*: 319-380. *Trends in Linguistics. Studies and Monographs* 90. Berlin: Mouton de Gruyter.
- Wiegand, H.E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. 1. Teilbd. Mit 159 Abbildungen im Text. Berlin/New York: De Gruyter.
- Wiegand, H.E. 1998a. Lexikographische Textverdichtung. Entwurf zu einer vollständigen Konzeption. Zettersten, A. und V. Hjørnager Pedersen (Red.). 1998. *Symposium on Lexicography VIII. Proceedings of the Eighth Symposium on Lexicography May 2-4, 1996 at the University of Copenhagen*: 1-35. *Lexicographica. Series Maior* 90. Tübingen: Niemeyer. [Auch in: Wiegand, H.E. 2000a: 1454-1489].
- Wiegand, H.E. 1998b. Altes und Neues zur Makrostruktur alphabetischer Printwörterbücher. Wiegand, H.E. (Red.). 1998. *Wörterbücher in der Diskussion III. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 348-372. *Lexicographica. Series Maior* 84. Tübingen: Niemeyer. [Auch in: Wiegand, H.E. 2000a: 1428-1453].
- Wiegand, H.E. 1999. Artikel einsprachiger Lernerwörterbücher, Textgestaltwahrnehmung und Suchbereichstrukturen. Plädoyer für übersichtliche Printwörterbücher im Zeitalter der Neuen Medien. Skibitzki, B. und B. Wotjak (Red.). 1999. *Linguistik und Deutsch als Fremdsprache. Festschrift für Gerhard Helbig zum 70. Geburtstag*: 259-281. Tübingen: Niemeyer.
- Wiegand, H.E. 1999a. *Semantics and Lexicography. Selected Studies (1976-1996)*. Hrsg. v. Antje Immen und Werner Wolski. *Lexicographica. Series Maior* 87. Tübingen: Niemeyer.
- Wiegand, H.E. 2000. Über Suchbereiche, Suchzonen und ihre textuellen Strukturen in Printwörterbüchern. Ein Beitrag zur Theorie der Wörterbuchform. Wiegand, H.E. (Red.). 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 233-301. *Lexicographica. Series Maior* 1000. Tübingen: Niemeyer.

- Wiegand, H.E. 2000a. *Kleine Schriften. Eine Auswahl aus den Jahren 1970 bis 1999 in zwei Bänden. Bd. 1: 1970-1988; Bd. 2: 1988-1999.* Hrsg. v. Matthias Kammerer und Werner Wolski. Berlin/New York: De Gruyter.
- Wiegand, H.E. 2000b. Sprachkontaktwörterbücher: Typen, Funktionen, Strukturen. Petkov, P. und H.E. Wiegand (Red.). *Theoretische und praktische Probleme der Lexikographie. 1. Internationales Kolloquium zur Wörterbuchforschung am Institut „Germanicum“ der St. Kliment-Ohridski-Universität in Sofia.* Germanistische Linguistik. Hildesheim/New York: Olms.
- Wiegand, H.E. 2000c. Fremdwörterbücher und Sprachwirklichkeit. *Jahrbuch des Instituts für deutsche Sprache.* Erscheint.
- Wiegand, H.E. 2000d. Über tabellarische Wörterverzeichnisse und deren Wörterbuchartikel. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica.* Erscheint.
- Wiegand, H.E. 2000e. Probleme der Adressierung in der zweisprachigen Lexikographie. Vortrag, gehalten auf dem III. Ost-West-Kolloquium im Februar 2000. Erscheint.
- Wiegand, H.E. 2000f. Wissen, Wissensrepräsentation und Printwörterbücher. Heid, Ulrich, Stefan Evert, Egbert Lehmann und Christian Rohrer (Red.). 2000. *Proceedings of the Ninth EURALEX International Congress, Euralex 2000, Stuttgart, Germany, August 8th-12th, 2000. Vol. 1:* 15-38. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.
- Wolski, W. 1986. *Partikellexikographie. Ein Beitrag zur praktischen Lexikologie.* Mit einer englischen Zusammenfassung. *Lexicographica Series Maior 14.* Tübingen: Niemeyer.
- Wolski, W. 1989. Formen der Textverdichtung im allgemeinen einsprachigen Wörterbuch. *HSK 5.1:* 956-967.
- Wolski, W. 1991. Formen der Textverdichtung im zweisprachigen Wörterbuch. *HSK 5.3:* 2837-2854.

The Shona Corpus and the Problem of Tagging*

Emmanuel Chabata, *African Languages Research Institute, Department of African Languages and Literature, University of Zimbabwe, Harare, Zimbabwe*

Abstract: In this paper the writer examines problems the African Languages Lexical (ALLEX) Project (at present the African Languages Research Institute (ALRI)) encountered while tagging the Shona corpus. The problems to be highlighted include general problems which apply to more than one language as well as problems peculiar to Shona. The paper was inspired by the challenges the writer encountered when he took part in building the Shona corpus. An analysis of the problems that most corpus builders face shows that more problems are likely to be encountered when dealing with spoken corpora than with written corpora. The paper demonstrates that tagging is an important component of corpus building as it makes it easier for a researcher to extract relevant data. To utilise the benefits of a tagged corpus, the tagging should be thorough and accurate. Well-informed decisions form an integral part of the tagging process since the utility of a tagged corpus depends largely on the input of the tagging process. This paper shows the need to take the tagging process seriously.

Keywords: ALEX PROJECT, COMPUTER, CORPUS, ENCODING, FOREIGN WORD, LEMMATIZATION, LEXICOGRAPHY, MONITOR CORPUS, PART OF SPEECH, SCANNING, SHONA, SLANG, TAGGING, TRANSCRIPTION, WORD

Opsomming: Die Shonakorpus en die probleem van etikettering. In hierdie artikel ondersoek die outeur probleme wat die African Languages Lexical (ALLEX) Project (tans die African Languages Research Institute (ALRI)) teëgekomp het terwyl die Shonakorpus geëtiketteer is. Die probleme wat bespreek word, sluit algemene probleme in wat van toepassing is op meer as een taal, sowel as spesifieke probleme wat eie aan Shona is. Die artikel het sy ontstaan in die uitdagings wat die outeur teëgekomp het terwyl hy deel gehad het aan die opbou van die Shonakorpus. 'n Ontleding van die probleme waarvoor die meeste korpusbouers te staan kom, toon dat daar waarskynlik meer probleme teëgekomp word wanneer daar met gesproke korpora as met geskrewe korpora gewerk word. Die artikel toon dat etikettering 'n belangrike komponent van korpusbou is, aangesien dit dit vir die navorser makliker maak om relevante data te onttrek. Om die voordele van korpusetikettering te realiseer, moet die etikettering deeglik en akkuraat wees. Ingevolge besluite vorm 'n integrale deel van die etiketteringsproses aangesien die bruikbaarheid van 'n geëtiketteerde korpus hoofsaaklik afhang van die inset tydens die etiketteringsproses. Hierdie artikel toon die noodsaaklikheid om die etiketteringsproses ernstig op te neem.

* This paper was presented at the Fourth International Conference of the African Association for Lexicography, held at the University of Pretoria, Pretoria, 5-6 July 1999.

Keywords: ALLEXPROJEK, REKENAAR, KORPUS, ENKODERING, VREEMDE WOORD, LEMMATISERING, LEKSIKOGRAFIE, MONITORKORPUS, WOORDSOORT, SKANDERING, SHONA, SLENG, ETIKETTERING, TRANSKRIPSIE, WOORD

1. Introduction

Recent linguistic researches have shown corpora to be important as a source of data for linguistic analyses. As Kennedy (1998: 88) for example observes, a linguistic corpus, in whatever form, is important as a basis for more accurate and reliable descriptions of how languages are structured and used. Thus, they are a source of evidence for linguistic descriptions and other researches that have to do with exploring language use. Renouf (1987: 1) has defined a corpus as "a collection of texts, of the written or spoken word, which is stored and processed on computer for the purposes of linguistic research". As implied in this definition, the term corpora is synonymous to machine-readable and computer-processed data. This does not, however, play down traditional corpora where texts were put on slips of paper and where relevant information could only be accessed manually.

Corpora may exist in two forms, namely, nontagged/unannotated or tagged/annotated. The difference between these two forms is that an unannotated corpus is just a plain text in its raw form, whilst a tagged corpus is enhanced with different kinds of information attached to each text or to items in a text. A close look at these two forms would probably show that it is more advantageous to work with an annotated corpus than with a plain one. McEnery (1996: 24) observes that annotating considerably increases the utility of a corpus. This is because information that may be implicit in plain text is made explicit through concrete annotation. In this case annotation makes it quicker and easier to retrieve and analyse relevant data from the corpus.

The term "tagging" has received different interpretations from different scholars. Svartvik (1982: 92) defines it as "the assignment of a lexical-grammatical description to linguistic units in the transcription from the audio tapes". Whilst this definition captures the idea that tagging is a way of marking those linguistic features that describe linguistic units, it tends to restrict the tagging process to spoken corpora, that is, to corpora built from oral material recorded on audio tapes and then transcribed. However, material in a corpus can also come from other different sources which do not necessarily involve transcription. This material can also be marked with codes that indicate various linguistic features. As a result we may need a definition that also incorporates corpus material collected from books, magazines, newspapers as well as material in electronic form. We can thus work with a more general and inclusive definition by McEnery (1996: 36), who defines tagging as "the attachment of specific codes to words in order to indicate particular features". The choice of linguistic features to be marked, depends on the kinds of studies for which the researcher wants to use the corpus. An example of the kinds of linguistic

information that can be attached to items in a corpus, are shown in the sentence **Amai vanoda sadza** (Mother likes sadza) which can be tagged as follows:

```
<subject>Amai</subject> <subject concord><tense>va</tense></subject concord><aspect>no</aspect><verb stem>da</verb stem>
<object>sadza</object>.
```

The tagging in this short sentence is aimed at showing the different roles that each item plays in the sentence. **Amai** plays the role of the subject, **v-** is the subject concord, **-a-** is the present tense marker, **-no-** marks aspect, **-da** is the verb stem and **sadza** is the object. It is, however, important to note that these are not the only kinds of information with which each of these items can be marked. One can also add information about word category, noun class and any other features peculiar to each of the items. Using the codes presented in this example, we could, for instance, extract all subjects, verb stems or objects that are marked in the corpus. This is done by instructing the computer to extract all data with a specific code.

The process of tagging can be compared to the process of putting clothes into a wardrobe where one does not just dump all of them into one pile. Instead, one would put shirts in their own compartment, trousers in another and socks in yet another. This is not a purposeless arrangement. Instead, this is done so that one can have easy access to a particular type of clothing for which one is looking. As already noted, in corpus building, tagging involves the marking of a plain text. Specific features of linguistic units, for example, word category and grammatical function, would be indicated through the use of specific codes. As with clothes in a wardrobe, the codes place linguistic units with similar features into specific compartments which the researcher has created. Using particular codes, a researcher can easily retrieve only the information which is relevant to his/her study.

Tagged corpora have been found to be more and more useful in language-related studies such as lexicography, dialectology, semantics, syntax, psycholinguistics, sociolinguistics and code-switching and code-mixing. These disciplines either focus on how language is structured or how it is used, and conclusive evidence for any controversial issues in these areas can only come from instances where language is found in use. A well-collected corpus should be a true reflection of the everyday use of language, and it should be a useful source of linguistic evidence. This is especially so in cases where the corpus or corpora have been built from texts collected from ordinary speech. Ordinary speech is taken here to refer to conversations that take place while people are interacting in the manner they do in their daily life.

It is, however, important to note that a tagged corpus is most useful when the tags are accurate. Producing an accurately tagged corpus is not an easy task; it has a number of challenges. This paper will highlight some of the constraints by looking at the tagging of the Shona corpus. However, before looking

at problems, we look at the Shona corpus, that is, how and why it was built. This would make it easier to understand and appreciate some of the challenges, especially those that have already been encountered by the ALLEX team.

2. The Shona corpus

The Shona corpus, with a current size of about two and a half million words, is a product of the African Languages Lexical (ALLEX) Project (at present the African Languages Research Institute (ALRI)). The ALLEX Project is housed in the Department of African Languages and Literature at the University of Zimbabwe. The Project was launched in 1992 and its major objective is to publish reference works that enhance the development of the indigenous languages of Zimbabwe. Some of the Project's envisioned publications include both monolingual and bilingual dictionaries. The Project also hopes to publish glossaries, for example, glossaries of musical, linguistic and literary terms, as well as specialized glossaries for ZimSign (Zimbabwean sign language with glossaries in Shona, Ndebele and English), science and technology (Chimhundu 1994: 21). All these activities need evidence of language use. As a result, the ALLEX Project made the building of corpora one of its priorities. In this case the corpora provide instances in which particular word forms are used for purposes of headword selection, defining and creation of usage examples.

To date, the ALLEX Project has published *Duramazwi ReChiShona*, the first Shona monolingual dictionary, and is currently working on two dictionaries, the Advanced Shona Dictionary (ASD) and the General Ndebele Dictionary (GND). Whilst the production of *Duramazwi ReChiShona* was corpus-aided, that is, its compilation was assisted with material from the corpus, that of the ASD and the GND is corpus-based. This means that headwords, senses and other relevant linguistic information required in the compilation of these dictionaries come from corpora in the two languages, Shona and Ndebele.

The research done by the ALLEX Project is based not only on theoretical linguistics, but also on linguistic data which shows how speakers use language. The aim is to publish reference works that reflect and represent actual language usage, not usage that is only theoretically possible. It is because of the nature of this research by the Project that the building of corpora became necessary.

The building of the Shona corpus started in 1992 when the ALLEX Project sent out student research assistants (undergraduate Shona students) to conduct interviews in all the districts in Zimbabwe where Shona is spoken. The interviews were on various socioeconomic, cultural, religious and political issues and comprised a representative sample of Shona discourse as it is used by the total population of the language's speakers. They captured vocabulary used in all varieties of Shona by people in different age and social groups.

For the collection of data in the field, each research assistant was provided with a cassette recorder and about 20 audio tapes. In order to obtain material from a variety of text types, context-governed material was collected. Some of

these contexts included public speeches, church sermons, school lessons, lectures, individual interviews on people's life experiences, narratives of historical events and descriptions of major social events. In the process of systematically collecting this oral material, details on the context of discourse, for example, date of interview, physical location, topic, gender, age, education and social status of participants, the setting and other relevant details were recorded. Extralinguistic features such as hesitations, repetitions, shouts, coughs and whispers were also recorded and marked. The collection of the oral material resulted in 750 audio tapes of spoken Shona, which constitutes about 70% of the current Shona corpus.

After recording, the research assistants transcribed the material. Since the interviews were not edited, the transcriptions represent speech as it was recorded. The decision not to edit the interviews was taken to ensure that the oral material in the corpus was as natural as possible. The transcribed material was encoded or keyed into the computer, proofread and tagged. The tags include those for the header to identify each interview as a unit different from the others, and body tags which mark each speaker's utterances and the sentences that make them up. This oral material can now be accessed in three different modes, that is, as speech on audio tapes, as transcriptions on paper and in electronic form on the computer. In the corpus each interview constitutes a corpus text.

In addition to oral material, there is also written material which constitute the remainder of the corpus. This material came from a variety of sources such as novels, poems, drama and school text books. It was scanned, proofread, tagged and then stored in electronic form. The sources of written material include newspapers published in Shona (such as *Kwayedza*), magazines, pamphlets, advertisements, evidence taken down in magistrate courts (where it is given in Shona), material from the Zimbabwe parliament (mainly translations of parliamentary sessions) and the Zimbabwe Broadcasting Corporation (where radio stations 2 and 4 mainly broadcast in Shona and Ndebele). Each publication would then constitute a corpus text. Emphasis is now on collecting material already in electronic form, which is being done through negotiations with publishing houses that publish works written in Shona, since material in such a form minimise the laborious task of scanning and proofreading. The Project's ultimate goal is to develop and maintain a monitor corpus of spoken and written uses of Shona.

The Shona corpus can be described as a general-purpose corpus. Its use is not restricted to a specific type of linguistic research. As noted earlier, the Project's aim is that the corpus should be used as a source of data for a variety of linguistic researches. The corpus can also be described as a monitor corpus. According to McEnery (1996: 22), a monitor corpus is open-ended. Texts are constantly added to it so that it gets bigger and bigger as more samples are added. A monitor corpus is important for the ALLEX Project, which specialises in dictionary making. In fact, monitor corpora, according to McEnery (1996:

22), "are primarily important in lexicographic work for they enable lexicographers to trawl a stream of new texts looking for occurrence of new words or for changing meanings of old words".

3. The tagging process and the challenges encountered

The importance of tagging a corpus has already been alluded to. Notwithstanding its significant role in the building of corpora, tagging has its challenges. Besides the need for a thorough understanding of the structure of the language with which the researcher is working, there are problems inherent in the corpus building process itself. In this section of the paper, the writer pays attention to problems associated with tagging, including those that have already been encountered by the ALLEX team. It is important to note that when discussing the Shona corpus as it now exists, two levels of tagging are involved. The first is the tagging that deals with marking each text as a unit as well as marking individual sentences in each speaker's utterance. This is what the ALLEX team has been doing. The second level is the grammatical or morphological tagging for which the team is developing programmes at present. This section deals with problems that have already been encountered at the first level of tagging as well as those that are being encountered in the process of developing the programmes that should cater for morphological tagging.

Some tagging problems are old problems that linguists have already encountered even before the establishment of corpus linguistics as a linguistic discipline or the acceptance of a corpus as a standard tool for linguistic research. These are problems that emanate from the way languages are structured. One such problem is the definition of the word "word". The notion of what a word is, has a number of possible interpretations, some of which are beyond the scope of this paper. In fact, what may be considered a word by one researcher may not constitute a word when viewed from another angle. The definition of a word is controversial and up to date linguists have not yet agreed as to what a word is or what it is not. In corpus building, the definition of a word is especially important when one is doing part of speech tagging, that is, when marking items with codes that indicate word-class category. However, in the process of marking it is sometimes difficult to distinguish a word from a phrase. At times it is difficult to establish a dividing line between these two. To illustrate this point we can take a Shona example, *Ndakazomuona* (I eventually saw him). Graphologically, this construction can be thought of as a word. However, a deeper analysis of this example shows that it is an inflected verb phrase consisting of:

Nda- (subject concord and tense),
-ka- (aspect),
-zo- (auxiliary),
-mu- (object concord),

-on- (verb radical), and
-a (terminal vowel).

Although they occur together as one form, each morpheme has a grammatical function which is realised at phrasal level. Given this information, it may not be enough to look at a word just as a group of letters written together and separated from the others by spaces.

The definition of what constitutes a word also poses a problem for word division. There are cases where it is not clear where and when to separate words in a sentence. This is especially so in cases where there is no clear definition of a word. To show this problem we can take the Shona form of the sentence "I was already going". The Shona translation of this sentence can be written in two forms, both of which seem to be acceptable, that is, (a) **Ndakange ndaa kutoenda** or (b) **Ndakange ndaakutoenda**. The problem here is whether **ndaa** and **kutoenda** should be written as one word form or as two separate units. This situation is a challenge to a corpus builder whose decision on such issues are reflected through the codes that he/she attaches to each unit.

Linked to the issue of word division are problems that emanate from the process of reducing spoken Shona to the written form. These problems are shared by the transcriber and encoder before they are passed on to the tagger, and they arise from the difficulties involved in the process of listening to and transcribing a spoken text in a way that is representative of the spoken language. In fact, it is difficult to have hundred percent accuracy when transcribing a spoken text. The whole problem according to McEnery (1996: 35) emanates from the fact that "in speech there is no explicit punctuation and any attempt at breaking down spoken language into sentences and phrases is an act of interpretation on the part of the corpus builder". Along with the unsettled problem of what a word is, the reduction of speech to writing posed a big challenge for those who worked with data collected through oral interviews. This is especially so given the fact that oral material in a corpus, when tagged, should represent original speech. The challenge is to come up with decisions that ensure that the resultant corpus captures the way the Shona language is structured or used both in the spoken and written modes.

For most linguistic research, part of speech or syntactic tagging is important. As noted earlier, this tagging process involves the assignment of codes to each word in a text as a way of labelling the word-class category to which it belongs in a particular context. Word-class tagging is especially important for researchers working with data restricted to a specific word-category, for example, noun, verb, adjective. If the tagging is done accurately, it should be possible for the researcher to capture only those words in which he/she is interested. However, in Shona, like in many other languages, the assignment of word-class tags is not always an easy task. There are times when the tagger comes across words which do not fit into any of the conventional word categories of the language one is working with. In Shona, for example, there are a

substantial number of words that pose classificatory problems. In Chimhundu (1996)'s *Duramazwi ReChiShona*, most of these words that could not be fitted into any of the conventional word categories of Shona ended up being put into the *kanu* (interjective) category. Examples of such words include the following:

- ahiwe** (word used when giving someone a warning against his/her behaviour),
- bodo** (word used when denying something),
- chagwa** (word used for claiming something that has fallen),
- mazvita** (word used for thanking), and
- diko** (word used to show that one agrees to something).

Despite the fact that these words were classified under the same word-category, they are different in a number of respects. For example, whilst *ahiwe* can carry an exclamation mark as a sign of strong interjection, *bodo* cannot. The problem with such a classification is that words are tagged with similar codes irrespective of the fact that they have different linguistic features if looked at closely. The adoption of such an approach when tagging would not help much in giving each of these words its distinctive features through the codes.

Related to the problem of word class assignment is the problem of lemmatization. Lemmatization involves the reduction of words in a corpus to their respective lexemes, that is, to the forms that one would look up if one were looking for the words in a dictionary. Besides ensuring that related words, for example, homographs are not classified under the same lemma, lemmatization also shows how derived word forms evolve from their respective base forms. A lemmatized text is therefore important to lexicographers, especially during the process of headword selection. If successfully done, it resolves the problem that most lexicographers face: that of determining what should be and should not be a headword. However, lemmatization is not an easy task. In Shona, for example, it is sometimes difficult to trace extended verb stems to their base forms. This is because sometimes when a verb is extended, it can become lexicalised, that is, it acquires new meanings that may not be related to the meaning of the base form, thus qualifying to stand on its own as a separate lexeme. To illustrate this point, we can take an example of the verb stem *-gadza* (1. make someone sit 2. put something on a fire 3. put someone in a position of responsibility) which is derived from *-gara* (sit) by the causative verbal extension *-dz-*. As we can see from the translations of the two forms, *-gadza* has acquired, besides its first sense, other new meanings which are not easily traceable to *-gara*. Also, in the everyday use of the verb *-gadza* no connection is made with *-gara*. The problem arises when one tries to link *-gadza* to *-gara* since, besides the literal meaning that *-gadza* has by virtue of adding the causative *-dz-* to *-gara*, it acquires more senses that have nothing to do with sitting as it is expressed in *-gara*.

Another difficult task with lemmatizing a text is the fact that most of the work has to be done manually. Although automatic lemmatization is possible, it is sometimes difficult to instruct a machine to make accurate decisions about tags, especially given the complexity of natural language (Biber 1998: 262). As noted by Kennedy (1998: 208), reliable automatic lemmatization also depends on reliable grammatical tagging. However, as we have already noted above, coming up with reliable grammatical tagging is a huge task.

The problem of reducing words in a corpus to their base forms is related to that of identifying the semantic input each morpheme makes to a construction. This is a tagging process that involves marking of semantic relationships between items in the text, for example, the relationship between agents and patients of particular actions. The tagging can also be used to show grammatical functions such as subjects and objects. This kind of tagging would be useful for those who would want to study the grammar of a particular language. Though important, the exercise has its constraints. Sometimes, a single morpheme is used to perform more than one grammatical function at the same time. In Shona, for example, the subject marker can also function as the tense marker in the same construction. To illustrate this we can take an example such as *Aenda mangwanani* (He went (today) in the morning). In this sentence, the *a-* in *aenda* is a marker for the subject as well as for recent past tense: *a-* cannot, therefore, be distinct in terms of its grammatical function. To try and show both functions may be very difficult.

Corpus builders also face challenges that are linked to language contact and language development. It is a known phenomenon that when two languages come into contact, there is a tendency to borrow vocabulary from one another. Words are adopted and adapted from one language to another. Since adoption and adaptation is a process, it is difficult to determine the status of some words at a particular time in their history. A word would enter a language as foreign, and with time it may settle by changing its linguistic characteristics in order to fit into the new linguistic environment. Once the word has settled, it becomes part and parcel of the new language and is no longer perceived as foreign. However, if it keeps its form and pronunciation in the new environment, it remains foreign and should be marked so in a corpus. To indicate that words are foreign is important for linguistic studies such as code-mixing and code-switching. Marking them as foreign, in this case, would make it possible for the researcher to easily access those sentences or contexts where these words are found without going through the whole text. However, we have just noted that the process to fit into the new language is slow. The slowness of this change poses the problem of determining whether, at a certain time, a word has remained foreign or whether it is still in the process of change or whether it has already settled.

In the Shona corpus, the problem is encountered when dealing with words that are borrowed from either English or Ndebele and are attached to Shona word forms. This can be seen in the following constructions: *kureceiver* (to

receive), *kumixa* (to mix) *kuaccepta* (to accept) and *Ndakaapplaya* (I applied). The problem is to represent the English forms as foreign in their respective contexts. One option of solving the problem would be to remorphologise and rephonologise such forms. However, the problem with such an option is to ensure that corpus users would only recognise easily the original form of the words in spoken language. For example, *Ndakaapplaya* may become *Ndakaapuraya* in which case, it no longer "looks like" English. Although this may be a way of trying to deal with the problem, it may not capture the manner in which most Shona people use or recognise this construction. Furthermore, remorphologising and rephonologising would also mean that the spoken text is edited. This in turn contradicts the principle that spoken corpora should mirror speech. The other option would be to tag the English parts as foreign. In this case, a construction like *kumixa*, would be tagged as follows: *ku<foreign>mix</foreign>a*. Straightforward cases like this one are not a problem. The problem comes when one is trying to tag 'apply' as a foreign word in the construction, *Ndakaapplaya* where a Shona vowel -a- has been introduced between consonants *l* and *y*. It is difficult to pick out and mark the English form without misrepresenting the way it is pronounced by most Shona people.

Related to the issue of foreign words is the problem of handling slang words. Some slang words come into a language for a short time and immediately fall out of use again, whilst others stabilise and end up being conventionalised in the language. This means that what may be regarded as slang at one point in the history of a language, may not be slang at another time. This poses the problem of determining whether a word is still slang or has settled in the language.

4. Conclusion

In this paper, we have seen some of the challenges that corpus builders face when tagging. Although examples were drawn from Shona alone, the problems discussed here can apply to most languages, particularly those in the Bantu family. What this paper has shown is that some problems come from the nature of the language under consideration, the language's interaction with other languages and also from the way the corpus material is collected.

This paper has shown that any aspect in a corpus can be tagged, depending on what one wants to study. This paper has given few examples of linguistic information that can be tagged, but has rather highlighted some problems that may be encountered in the process of tagging. Tagging is an important but big task in the building of corpora and should therefore be taken seriously.

References

- Biber, D. et al. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Chimhundu, H. (Ed.). 1994. *The ALLEX Project: Second Progress Report*. Harare: University of Zimbabwe.
- Chimhundu, H. (Ed.). 1996. *Duramazwi ReChiShona*. Harare: College Press.
- Kennedy, G. 1998. *An Introduction to Corpus Linguistics*. London / New York: Longman.
- McEnery, T. and A. Wilson (Eds.). 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Renouf, A. 1987. Corpus Development. J.M. Sinclair (Ed.). *Looking Up: An Account of the COBUILD Project in Lexical Computing*. London: Collins.
- Svartvik J. et al. 1982. Tagging the London-Lund corpus of Spoken English. Johansson, S. (Ed.). *Computer Corpora in English Language Research*. Bergen: Norwegian Computing Centre for the Humanities.

Should LSP Dictionaries also Include Professional Jargon and Slang?*

Rosemarie Gläser, *Foreign Language Centre, University of Leipzig, Leipzig, Germany*

Abstract: The paper discusses the question whether an LSP dictionary should take the sociolinguistic stratification of the entire speech community in a particular domain of activity into consideration, and as a consequence, also encompass vocabulary on the lower stylistic levels, such as professional jargon and occupational slang in addition to technical terminology and nomenclature. The author's contention is that by including such vocabulary, an LSP dictionary will be more comprehensive and closer to the communicative reality. Moreover, it will meet the more specific requirements of translators and interpreters. To substantiate this claim, the author examined two monolingual English dictionaries and five bilingual English-German (German-English) dictionaries published between 1973 and 2000, which have successfully incorporated professional jargon and slang. The subject areas covered range from law, economics, accountancy and business administration to geology, geophysics, geotectonics, and to automotive engineering. Some lexicographers have used style markers or usage labels (such as *informal, colloquial, slang, vulgar, umgangssprachlich, förmlich, gehoben, finance, motoring, press*) for vocabulary bearing stylistic or expressive connotations. As a rule, these style markers have only been given to key words of the source language so that they may be inferred for the target language equivalent. Ideally, as one dictionary author has proved, style markers should also be attached to target language equivalents. However, an LSP lexicographer must be aware of the fact that informal and slang words may be short-lived and be substituted by other colloquial and colourful words from oral communication — and that slang words may gradually lose their connotations and become stylistically neutral words and even terms.

Keywords: LSP DICTIONARY, TECHNICAL TERM, ITEM OF NOMENCLATURE, TRIVIAL (POPULAR) NAME, PROFESSIONAL JARGON, OCCUPATIONAL SLANG, TRADE NAME, CONNOTATION, STYLISTIC LEVEL/LEVEL OF STYLE, STYLE MARKER, USAGE LABEL, MONOLINGUAL/BILINGUAL DICTIONARY, SOCIOLINGUISTIC STRATIFICATION OF THE SPEECH COMMUNITY.

Zusammenfassung: Sollen Fachwörterbücher auch Berufsjargon und Slang einbeziehen? Der Artikel erörtert das Problem, ob ein Fachwörterbuch auch die soziolinguistische Schichtung der gesamten Kommunikationsgemeinschaft eines Fachgebietes berücksichtigen

* This paper was presented at the Fifth International Conference of the African Association for Lexicography, held at the University of Stellenbosch, Stellenbosch, 3-5 July 2000.

und folglich neben der Terminologie und Nomenklatur auch den Wortschatz der unteren Stilebenen, wie Berufsjargon und -slang, einbeziehen sollte. Die Autorin geht von der These aus, daß durch die Einbeziehung dieses Wortschatzes ein Fachwörterbuch an Umfang gewinnt, der kommunikativen Wirklichkeit stärker angenähert ist und speziellen Anforderungen der Übersetzer und Dolmetscher eher gerecht wird. Zur Stützung dieser These hat die Autorin zwei einsprachige englische und fünf zweisprachige englisch-deutsche (deutsch-englische) Fachwörterbücher untersucht, die zwischen 1973 und 2000 erschienen sind und den Vorteil haben, daß sie auch Berufsjargon und -slang berücksichtigen. Die hier vertretenen Fachgebiete reichen von der Rechts-, Wirtschafts- und Börsensprache zur Sprache der Geowissenschaften und der Kraftfahrzeugtechnik. Einige Lexikographen verwenden Stilmarkierungen und Sachgebietsindikatoren (z.B. *informal, colloquial, slang, vulgar, umgangssprachlich, förmlich, gehoben, finance, motoring, press*) für Wörter mit stilistischen oder expressiven Konnotationen. In der Regel sind diese Stilmarkierungen nur bei den Ausgangssprachlichen Wörtern verzeichnet, so daß sie bei dem Zielsprachlichen Äquivalent erschlossen werden müssen. Im Idealfalle sollten die Stilmarkierungen auch bei den Zielsprachlichen Äquivalenten angegeben werden, wofür einer der Wörterbuchautoren den Beweis geliefert hat. Jedoch muß sich der Fachwörterbuchautor stets der Tatsache bewußt sein, daß Jargon- und Slangwörter kurzlebig sein und durch andere umgangssprachliche und plastische Wörter aus dem mündlichen Sprachgebrauch ersetzt werden können — ebenso wie Slangwörter allmählich ihre Konnotationen verlieren, stilistisch neutrale Bezeichnungen und sogar Termini werden können.

Stichwörter: FACHWÖRTERBUCH, TERMINUS, NOMENKLATURZEICHEN, TRIVIAL-NAME, BERUFSJARGONISMEN, BERUFSSLANG, WARENNAME, KONNOTATION, STILEBENE, STILMARKIERUNG, SACHGEBIETSINDIKATOR, EINSPRACHIGES / ZWEISPRACHIGES WÖRTERBUCH, SOZIOLINGUISTISCHE SCHICHTUNG DER KOMMUNIKATIONSGEMEINSCHAFT.

1. Introduction

In this paper I will discuss the question to what extent a specialist dictionary should reflect the social stratification of an LSP in such a way that it also includes words and phrases from the lower stylistic levels. This general problem touches monolingual and bilingual dictionaries alike, but its solution depends on a number of presuppositions. The first is the individual lexicographer's theoretical understanding of the range of a technical vocabulary and of the social differentiation of LSP communication. In other words, the option is

- whether technical vocabulary is limited to terminology and nomenclature as lexical units of written communication that have a clear-cut meaning and are stylistically neutral, or
- whether technical vocabulary is extended to words and phrases of spoken discourse in which they occur as professional jargon and occupational slang, as popular names for items of nomenclature, and also as trade names.

The lexicographer's primary decision on the range of vocabulary covered in an LSP dictionary will be influenced by a number of factors. These include

- (a) the composition of the word-stock in a particular domain,
- (b) the intention of the LSP dictionary: whether it will describe and explain current usage or prescribe and stipulate preferable vocabulary, and
- (c) the target group of the prospective dictionary users, which may range from specialists on a high theoretical, academic level to technicians, laboratory staff, and factory workers — and to people in the applied field in the broadest sense.

My paper is based on the assumption that the inclusion of technical jargon and slang in an LSP dictionary will always be an asset to such a reference book and provide an additional benefit to the user, especially to translators and interpreters.

As a rule, professional jargon and slang are typical of oral communication, but they may occasionally enter written texts, especially in journalism. They are colourful metaphorical or metonymical expressions. As they spring up in everyday practical work, their appearance may be drastic and "shirt-sleeved". In this respect such informal or slangy LSP words confirm and underpin an old dictum, a "winged word" about slang which reads: "Slang is language that rolls up its sleeves, spits in its hands and sets to work."

2. The stratification of LSP vocabulary

The specialist vocabulary used in a particular domain of social activity is not monolithic, but structured according to

- (a) the level of abstraction (and cognition) which determines the meaning of technical terms,
- (b) the degree of the communicative partner's expertise, and
- (c) the character of work done by those people who work in a special domain, ranging from research and development, to designing and planning, and finally to the sphere of production and distribution.

In terms of a vertical stratification of LSP (cf. Sager, Dungworth and McDonald 1980), its vocabulary includes

(a) technical terms

As a rule, technical terms designate a defined concept. Thus, they have a nominating and a defining function. They are expected to be precise, unequivocal, unemotional, stylistically neutral, without evaluative connotations, and suitable for patterns of word-formation.

(b) items of nomenclature

Nomenclatures are taxonomies and are understood as a specific field of terminology. These lexical units only have a nominating and classifying, but no defining function. They designate individual objects which as a whole constitute a scientific system, for example in biology (the Linnaean system of plants and animals), in chemistry (the periodical system of elements, or the IUPAC nomenclature of compounds), in medicine (the nomenclature of bones, muscles, nerves, glands, inner organs, diseases and symptoms, etc.). The individual items of nomenclature cannot be named at random, but must follow international word-formation rules.

(c) trivial (popular) names

These are colloquial synonyms of systematic names and are circulating in everyday language. Moreover, they are much older than systematic names, may count as prescientific, and designate minerals, plants, animals, diseases, chiefly in the spoken language. Trivial names are often listed in LSP dictionaries, depending on the lexicographer's broader or narrower concept of technical discourse. They are always included in traditional encyclopedic LSP dictionaries, such as *Langenscheidts Grosswörterbuch der englischen und deutschen Sprache*, *Der kleine Muret-Sanders*.

(d) jargon words (professional jargon, indoor jargon) and slang words

These informal words are typical of oral LSP communication in an everyday working environment; they do not designate a concept classified in a system or coined according to international patterns. Professional jargon is down-to-earth, drastic and shirt-sleeved. It may include colourful idiomatic expressions.

(e) trade names

Names for commodities and services are arbitrary and often short-lived. They label specific products in the sphere of production and consumption and have no conceptual essence. Trade names or trade marks are in fact proper names. A number of them are registered trademarks and therefore an essential part of LSP vocabulary.

3. Materials and Method

My paper is based on a linguistic analysis of two monolingual and five bilingual LSP dictionaries whose authors have successfully incorporated technical jargon and slang and, with few exceptions, have also used style markers. These include the labels informal, colloquial, slang, vulgar, rare, and obsolete. The subject areas covered in this comparative survey range from law (Peter H. Collin; Alfred Romain, Hans Anton Bader and B. Sharon Byrd); economics, accountancy business (Hermann Fink; Theo van Bernem; Wilhelm Schäfer) to

geophysics, geotechnology and geotectonics (H.-J. Behr and W. Bachmann), and to automotive engineering (Peter Axel Schmitt).

4. Corpus Analysis

4.1. Monolingual dictionaries

4.1.1. Peter H. Collin (1994² [1986]): Dictionary of Law.

This dictionary contains about 7 000 lexical items (classified as "terms") from the legal language. The dictionary implicitly mirrors the social, communicative composition and stratification of the community that constitutes the "legal profession". There are numerous Latin words and phrases on the formal stylistic level, such as *audi et alteram partem* ("hear the other side"), *de jure - de facto* ("by right - in fact"), *casus belli* ("case of war"), *donatio mortis causa* ("gift because of death"), abbreviations (*c.o.d.* = "cash on delivery"), French expressions like *carte blanche*, and words which the author has labelled as "informal". These are typical examples of professional jargon. To quote a few expressions:

- graft* - "corruption of officials",
- quickie* - "divorce which is processed rapidly through the court by the use of the special procedure",
- devil* - "barrister to whom another barrister passes work because he is too busy",
- dove* - "person who prefers diplomacy and tries to achieve peace",
- hawk* - "person who believes in threatening the use of armed force as a means of settling problems between countries",
- kangaroo* - "system used when discussing a bill where some clauses are not discussed at all, but simply voted on, with the discussion then moving on to the next item; this is illustrated by the compound *kangaroo court*".

A remarkable feature of this dictionary is the listing of number of words classified as "(prison) slang". They are either simplex or compound lexemes, native (Germanic) words, and chiefly metaphors. They designate persons involved in criminal activities. Other words are related to things. Compare the following examples:

- grass* - "criminal who gives information to the police about other criminals",
- bellman* - "criminal who specialises in stopping alarm signals",
- cracksman* - "criminal who specialises in breaking safes",
- baron* - "prisoner who has power over other prisoners because he sells tobacco and runs rackets in prison",

dip – "pickpocket",

screw – "prison warden" (American English for "warden"),

bracelets – "handcuffs" (a metonymical expression).

The slang words in a legal dictionary are an essential supplement because they shed some light on down-to-earth communication among prisoners and indoor staff.

4.1.2. Hermann Fink (1999⁶): Econoterms. A Glossary of Economic Terms. With a Supplement EconoSlang

Hermann Fink's glossary, which has become a household reference book for students, translators and subject specialists, includes an independent wordlist under the heading "EconoSlang". The author's justification for including this stratum of nonterminological words into his dictionary of economics and business administration is of particular sociolinguistic interest.

In the introduction (Fink 1999⁶: 279) we read:

EconoSlang presents a collection of approximately 1 300 colloquial terms and expressions typical of American business jargon frequently used in trade and industry and in the press.

Most of the entries are slang, indeed, some are considered vulgar or obscene. Other entries simply reflect the linguistic creativity of the American language — the tendency to play on words and use metaphorical, colourful, and often exaggerated expressions especially in business terminology ... Consequently, EconoSlang should be regarded as a guide reference tool and as an introduction to the typical American business jargon that economics students and others interested or working in the field of business and economics will eventually encounter.

Examples from the wordlist of EconoSlang confirm the reader's observation that most expressions have a figurative meaning. The vocabulary includes metaphorical simplex words, compounds and idiomatic phrases. Vulgar words are rather marginal cases.

Simplex words:

grease – "bribe",

cannibalize – "(to) use assets of one part of a company to support another",

trafficking – "(to) deal in illegal merchandise (e.g. drugs)".

Compounds:

bullshit (vulgar) – "false or faulty information, nonsense",

bellyache – "complaint" (e.g. the union has a bellyache),

play-money – "money one can afford to lose; money used in games, foreign currencies (derogatory)".

Idiomatic phrases:

hot money – "stolen money",

to earmark (funds) – "(to) reserve funds for special uses",

to bust one's britches (vulgar) – "to work very hard" (*britches* is the American spelling of *breeches*),

to piss on (vulgar) – "attempt to damage the reputation of a co-worker".

On the whole, Hermann Fink's glossary is unrivalled in its scope of current vocabulary on the lower levels of style, which is, nevertheless, current usage of American journalists in economic reports.

4.2. Bilingual Dictionaries

4.2.1. Alfred Romain, Hans Anton Bader and B. Sharon Byrd (2000⁵): *Wörterbuch der Rechts- und Wirtschaftssprache. Teil 1. Englisch-Deutsch.*

In this English-German Dictionary of Law and Economics, the authors have included regional and social varieties of LSP use, such as specific German words used in Austria, Americanisms, on the one hand, and slang words, largely described as "Fachjargon", on the other. The following examples are most illustrative:

cop – (sl.) "Polizist"

cop – (verb) (US sl.) "wegnehmen, stehlen, gewinnen"

cook – (verb) (sl.) "(Bilanz, Rechnung) verschleiern, fälschen"

lifer – (sl.) "Lebenslänglicher, zu 'lebenslänglich' Verurteilter, lebenslange Freiheitsstrafe".

4.2.2. Theodor van Bernem (1994³): *Wirtschaftsenglisch-Wörterbuch. Englisch-Deutsch. Deutsch-Englisch.*

Although Van Bernem does not apply usage labels in his dictionary of business English, a considerable number of words and phrases are qualified as *informal*, *colloquial* or even *slang* by their German equivalents. In this respect, the dictionary user is offered word equivalents on the same level of style. This vocabulary includes new coinages from business journals and journalists' jargon which have not yet gained a wide circulation. For this reason, the author could in a number of cases only suggest the German equivalent of an English neologism

as a context-sensitive translation. But the author was able to offer the German user an acceptable equivalent with approximate connotations and to trace business language in its making.

Another striking feature is that quite a number of the business expressions on the level of jargon or slang have negative conceptual associations and derogatory expressive connotations. They may be allocated to semantic fields related to persons, actions or situations. Many of them are animal metaphors.

Colloquial and slang expressions for persons are:

moonlighter – "Schwarzarbeiter",
toady – "Speichellecker",
sweater – "Schinder" (a person who toils),
standby – "Springer, Stütze, Reserve-",
couch potato – "fernsehsüchtiger Fast-Food-Freak".

Compound nouns based on a metaphorical element may designate situations. The conceptual content of the expressions equally contains a negative value judgement. Examples are:

window dressing – "Bilanzverschleierung, Bilanzkosmetik",
rat race – "Postenjägerei, harte Berufskonkurrenz",
sunset industry – "niedergehender, sinkender Wirtschaftszweig" (as opposed to *sunrise industry* – "aufstrebender Wirtschaftszweig"),
peanut economics – "wirtschaftliches Intrigenspiel".

Similar negative associations are connected with string compounds of the following type which also have a metaphorical constituent:

meat-axe reduction – "pauschale Kürzung",
yellow-dog contract – "Arbeitsvertrag, der den Beitritt zur Gewerkschaft verbietet",
blind-alley job/dead-end job – "Beruf ohne Zukunft",
bait and switch technique – "Lockvogeltaktik",
beggar-my-neighbour policy – "Leistungsbilanzüberschußpolitik".

The phrase *pop-and-mom store* "Tante-Emma-Laden" may be associated with jocular connotations, and is colloquial.

The dictionary presents a number of metaphorical verb phrases which also correspond in English and German on the stylistic level, such as.

to jam on the credit brake – "Kreditbremse anziehen",
to work on the knocker – "Klinken putzen",
to eke out a living – "sich kümmerlich durchschlagen".

Quite a number of simplex and phrasal verbs listed by the author have derogatory connotations, such as *to sweep off* – "Geld einstreichen", and *to sap* –

"jemanden fertigmachen".

Colloquial and/or slangy connotations are also shared by several expressions consisting of an adjective plus noun in English and a compound in German. The metaphorical element is present in the source and the target language. Examples for illustration are:

filthy lucre – "schmutziges Geld",
messy job – "Drecksarbeit",
sweated money – "Hungerlohn",
slaughtered price – "Schleuderpreis".

These expressions designate an unpleasant situation or state of affairs by means of an adjective which associates a negative evaluation.

Obviously, derogatory and pejorative connotations seem to abound in those economic words and phrases that have no terminological status, but are colloquial or slangy vogue words used by journalists. Drastic examples of this type also include:

milking – "Ausbeutung eines Unternehmens",
rag business – "Damenbekleidungsindustrie" (women's clothing industry),
shopaholic (analogous to *workaholic*) – "kaufwütig",
frothy market – "nervöser Markt",
fledgling recovery – "zögernde Erholung".

To sum up, Theo van Bernem's dictionary is rich in informal economic vocabulary. The author has captured the context-sensitive meaning of the English expression in a more generalised German equivalent. He has included many neologisms. On the whole, the German equivalents are on a similar stylistic level, even though the author does not offer usage labels.

A remarkable advantage of this dictionary is that it has been updated in several recent editions, and that university students have made a considerable contribution by their project work in gathering current vocabulary from American economic journals.

4.2.3. Wilhelm Schäfer (2000⁶): Wirtschaftswörterbuch. Band I Englisch-Deutsch. Band II Deutsch-Englisch.

The authors have included English and German expressions on the informal (colloquial) and slang levels of usage and accordingly apply style markers/usage labels, such as *infml.* (informal), *sl.* (slang). Their dictionary is introduced by a linguistic chapter entitled "Bemerkungen zur Lexikographie".

Examples of economic vocabulary on the nonneutral level (words and phrases that have expressive or stylistic connotations) are:

Kredithai (Fin., infml.) – "loan shark" (*Fin.* stands for *Finance*),
Kredithahn zudrehen (Fin., infml.) – "to cut off credit to",
Kreditbremse ziehen (infml.) – "to clamp down on credits"; (sl) – "to jam on the credit brake".

4.2.4. Hans-Jürgen Behr and Walther Bachmann (1973): *Technik-Wörterbuch Geowissenschaften. Englisch-Deutsch.*

This dictionary was a joint project of a subject expert and a language specialist (who was an LSP teacher and a technical translator and interpreter) at the Mining Academy of Freiberg, Saxony. The authors collected colloquial and slangy vocabulary used in the field of mineral deposits ("Lagenstättenkunde"), such as colourful metaphors and monosyllabic, clipped words. These are summarised under the umbrella term "slang" and identified in the dictionary entry by the usage label *sl.* Unfortunately, the German equivalent is not marked in the same fashion. Thus, the German native speaker and translator into his/her mother tongue will be faced with a stylistically neutral expression which is constituted by a compound or by a descriptive phrase. It may happen that the German equivalent does not have the same stylistic (slangy) connotations as the English word or phrase. In this case, the dictionary authors have placed emphasis on conceptual equivalence rather than on stylistic adequacy.

The slangy vocabulary is interspersed with animal metaphors. To give a few examples:

ant mound (sl.) – "Schlammvulkan",
camel back (sl.) – "Ausbucklungen des Firstgesteins",
doghouse (sl.) – "Bohrmeisterbude" (drilling workers' shed; *Bude* is colloquial in German too),
bat (sl.) – "Kohlenschiefer",
cat dirt (sl.) – 1. "Kohle mit Pyrit"; 2. "harter, feuerfester Ton",
cow-dung bomb (not specified, but probably slang) – "Schlackenkuchen",
cowstone (sl.) – "Grünsandsteinblock",
lamb and slack (sl.) – "minderwertige Kohle",
toadstone (sl.) – "Melaphyr, Basaltporphyr".

Quite a number of slang words are used to designate workers in the oil-drilling field. Thus, *roughneck* (sl.) and *roustabout* (sl.) refer to a "Bohrarbeiter" (a drilling and milling worker in mining), and *mud smeller* (sl.) to an "Erdölgeologe" (a geologist who does the probing).

There are also numerous slang words that designate minerals, structures of rocks, and divisions inside a mine or pit. To quote a few examples:

barnyard (sl.) – "interglazialer Bodenhorizont",
balmstone (sl.) – "hangendes Gestein",

kackle-meckle (sl.) – "sehr reines Bleierz in Cornwall",
storm stone (sl.) – "Donnerkeil, Belemnit".

Another example not labelled as slang, but probably informal (descriptive, jocular or pejorative), is the expression *niggerheads* "Kugelnkohle".

To sum up, the slang vocabulary used in geophysics, geotectonics, geotechnology and mining shares the fields of metaphors (that is, names of animals, parts of the human body) with other subject domains and also the vocabulary on the neutral stylistic level.

4.2.5. Peter Axel Schmitt (1992): PONS Fachwörterbuch der KFZ-Technik. Englisch-Deutsch. Deutsch-Englisch.

The author of the dictionary of automotive engineering, Peter A. Schmitt, has systematically included colloquial vocabulary which was gathered by a considerable number of technical translators (postgraduate students and professional translators at the University of Mainz-Germersheim). Having worked for many years in several car-producing companies as a technical translator and handbook writer and being familiar with the factory floor, the author unites the expertise of a subject specialist and that of a language expert.

The head word in the dictionary as the key entry has been chosen as the preferred term. It is the standardised designation and does not have any expressive or stylistic connotations. Synonyms are of subordinate importance, but they are adequately described.

The author has developed a scale of usage levels which include pragmatic information such as *rare*, *coll* (colloquial), *form* (formal), *mot* (motoring), *obs* (obsolete), *poet* (poetic), *ppsc* (popular science), *pract* (practical), *press* (press media usage), *elevated* and *norm* (normalized). This type of highly specified information is a remarkable advantage over other comparable dictionaries in the same field. The detailed pragmatic information complements the semantic information which contains a definition or an explanation of the term. The author's intention is to offer the dictionary user computerised subject-specific, linguistic and pragmatic data in each lexical entry. Apart from a subject field indicator, the definition or explanation of the lexical unit and the German equivalent(s), there are references to synonymy, related terms, regional varieties of usage (BE (British English), AE (American English), Aus (Australian English)) and separate compounds. A typical example of this kind of a lexical entry is the term *convertible*. The entry has the following key word display:

convertible, conv (mot) (*conv* stands for *convertible*)
topless automobile coll press
open-air automobile press
ragtop coll
droptop coll

• *offenes Auto n* (*n* stands for *Neutrum*)

- *Cabrio* ugs (umgangssprachlich)
- *Oben-ohne-Auto* ugs press
- *Cabriolet* (2) obs form geh (gehoben)
- *Kabriolet* n obs form

Schmitt's dictionary is both descriptive and stipulative, and, by providing pragmatic information on the technical term in the source and target language, has many advantages over those LSP dictionaries which neglect or deliberately exclude colloquial and slangy words and phrases.

5. Results and Discussion

The presentation of a few select monolingual and bilingual dictionaries of different subject areas has shown that informal, colloquial, jargon and slang words have a justified place in an LSP dictionary because they also have firm footing in current LSP usage. Such words and phrases are a realistic picture of the discourse community of a special domain and its subfields. Moreover, they have sociolinguistic implications regarding the language user on a particular social stratum. In this respect, LSP dictionaries should collect, systematize and explain the word-stock (words and phrases of a language for special purposes) in all its dimensions of a functional variety of usage or a trade dialect. In the light of a vertical stratification of LSP communication, the vocabulary listed in an LSP dictionary should exceed that of pure terminology and nomenclature. The problem for the lexicographer, however, lies in collecting informal and slang terms. As a rule, statistically representative groups of informants are hardly available. Thus, the source of slang terms in economics are slipshod articles written by American journalists in newspapers and magazines. On the other hand, the geologist working on an oilfield abroad will become familiar with English slang words in the process of learning by doing. And, in a different context, the authors of legal dictionaries may have consultations with persons who are in charge of prison methods and familiar with the behaviour and the language of prisoners.

In the last instance, jargon and slang words will continue to enrich all sorts of technical vocabularies. An LSP lexicographer must, nevertheless, bear in mind that informal and slang words may be short-lived and be ousted by other colourful vocabulary often coined by the younger generation of special language users.

Literature (select bibliography)

- Behr, Hans-Jürgen and Walther Bachmann. 1973. *Technik-Wörterbuch Geowissenschaften. Englisch-Deutsch*. Edited by Adolf Watznauer. Berlin: VEB Verlag Technik.
- Collin, P.H. 1994¹ [1986]. *Dictionary of Law*. Teddington, Middlesex: Peter Collin Publishing Ltd.

- Fink, Hermann.** 1999⁶. *EconoTerms. A Glossary of Economic Terms. With a Supplement EconoSlang.* In collaboration with Paul Schons and Wolfgang Rothfritz. Sprachen und Kulturen. München/Wien: R. Oldenbourg Verlag.
- Fluck, Hans-Rüdiger.** 1996⁵. *Fachsprachen. Einführung und Bibliographie.* Tübingen: A. Francke Verlag.
- Gläser, Rosemarie.** 1978. *Fachstile des Englischen.* Leipzig: VEB Verlag Enzyklopädie.
- Messinger, Heinz und der Langenscheidt-Redaktion.** 1991⁶. *Langenscheidts Grosswörterbuch der englischen und deutschen Sprache. Der kleine Muret-Sanders. Deutsch-Englisch.* Berlin/Munich: Langenscheidt.
- Messinger, Heinz, Helmut Willmann und der Langenscheidt-Redaktion.** 1991⁵. *Langenscheidts Grosswörterbuch der englischen und deutschen Sprache. Der kleine Muret-Sanders. Deutsch-Englisch.* Berlin/Munich: Langenscheidt.
- Romain, Alfred, Hans Anton Bader and B. Sharon Byrd.** 2000⁵. *Wörterbuch der Rechts- und Wirtschaftssprache. Teil 1. Englisch-Deutsch.* München: C.H. Beck'sche Verlagsbuchhandlung/Basel/Mainz/Wien: Helbing und Lichtenhahn.
- Sager, Juan S., David Dungworth and Peter F. McDonald.** 1980. *English Special Languages. Principles and Practice in Science and Technology.* Wiesbaden: Oscar Brandstetter Verlag KG.
- Schäfer, Wilhelm.** 2000⁶. *Wirtschaftswörterbuch. Band I Englisch-Deutsch. Band II Deutsch-Englisch.* Established by Wilhelm Schäfer†, edited by Michael Schäfer in collaboration with Gabriele Strake-Behrendt. München: Verlag Franz Vahlen.
- Schmitt, Peter Axel.** 1992. *PONS Fachwörterbuch der KFZ-Technik. Englisch-Deutsch. Deutsch-Englisch.* Stuttgart/Dresden: Ernst Klett Verlag für Wissen und Bildung.
- Van Bernem, Theodor.** 1994³. *Wirtschaftsenglisch-Wörterbuch. Englisch-Deutsch. Deutsch-Englisch.* München/Wien: R. Oldenbourg Verlag.

Words in a Cultural Context*

Kathryn Kavanagh, *Executive Director, Dictionary Unit for South African English, Rhodes University, Grahamstown, Republic of South Africa*

Abstract: Successful communication is dependent upon an adequate level of cultural as well as linguistic understanding. Language is itself part of culture and reflects social structures and attitudes. This paper examines the case for dictionaries to carry cultural as well as lexical information. It reviews the American tradition and purpose of including cultural material in dictionaries, and considers the nature and scope of cultural information contained in a number of modern American and European dictionaries. Various means of presenting cultural information are also examined. The paper concludes by considering the implications for lexicographers developing dictionaries in multicultural South Africa.

Keywords: COMMUNICATION, CROSS-CULTURAL COMMUNICATION, CULTURAL INFORMATION, CULTURE, CURRICULUM, DICTIONARIES FOR SPECIAL PURPOSES, DICTIONARY, DICTIONARY DESIGN, ENCYCLOPEDIA, ENCYCLOPEDIA DICTIONARIES, LANGUAGE, LEXICOGRAPHY, MULTICULTURAL SOCIETY, SOUTH AFRICA

Opsomming: *Woorde binne 'n kulturele konteks.* Suksesvolle kommunikasie berus op 'n voldoende vlak van kulturele sowel as taalkundige begrip. Taal self is deel van kultuur en weerspieël sosiale strukture en houdings. Hierdie artikel ondersoek die insluiting van sowel kulturele as leksikale inligting in woordeboeke. Dit beoordeel die Amerikaanse tradisie en doel daarmee om kulturele materiaal in woordeboeke in te sluit, en oorweeg die aard en omvang van kulturele inligting vervat in 'n aantal moderne Amerikaanse en Europese woordeboeke. Verskillende maniere om kulturele inligting aan te bied, word ook ondersoek. Die artikel sluit af deur die implikasies vir leksikograwe te oorweeg wat woordeboeke saamstel in multikulturele Suid-Afrika.

Slutelwoorde: KOMMUNIKASIE, INTERKULTURELE KOMMUNIKASIE, KULTURELE INLIGTING, KULTUUR, KURRIKULUM, WOORDEBOEKE MET SPESIALE OOGMERKE, WOORDEBOEK, WOORDEBOEKONTWERP, ENSIKLOPEDIA, ENSIKLOPEDIA SE WOORDEBOEKE, TAAL, LEKSIKOGRAFIE, MULTIKULTURELE SAMELEWING, SUID-AFRIKA

Cultural factors and successful communication

My approach in this paper is not that of an academic who has spent several years researching language and its cultural context, but as a practising lexicographer who has been forced to consider the benefits and the problems of

* This paper was presented as a keynote address at the Fifth International Conference of the African Association for Lexicography, held at the University of Stellenbosch, Stellenbosch, 3-5 July 2000.

including cultural information in dictionaries. As lexicographers, we need to examine and reassess from time to time what is needed or expected by the users of our products, so I am going to consider the content of dictionaries, not so much the headword list but rather the cultural content, to see whether it meets the needs, and potential needs, of dictionary users.

We communicate mainly with words, and in order to communicate successfully with others we need to know the generally accepted meanings of words in our language, and to be able to structure them coherently in a phrase or sentence. If people are to interpret us correctly, we need a shared understanding of meaning and a shared knowledge of some basic grammatical rules.

We also have to take into account other factors. We need, for instance, to have some idea of register and tone of voice: to know when to use formal or informal language, or slang, in appropriate contexts, and to be able to appreciate the language level and tone of others. We learn much of this in relation to our own language as we grow up. In dictionaries, some social information can be coded in "register labels" at entries for individual words. However, standard dictionary entries give little further information about good manners and social norms.

Linguistic interaction is regularly constrained by cultural factors, such as age, seniority, or gender. We need to be aware of social conventions: when to speak, how to talk to people older than us, how to be polite. In our own language and in our own cultural setting this comes naturally. In other settings we may feel insecure. But we do need to know what is considered rude by people of other cultures, so that we can avoid it. For instance, we need to know how directly we can ask for something. Is it all right to be direct, or should you get round to the subject slowly? Is it all right to say "I want some chocolate", or should you say "Please may I have some chocolate?". Most British people would agree that introducing a request with "Please may I have ...", rather than saying "I want ..." or "Give me ...", shows better manners, and is more likely to prompt a favourable response. In other languages, amongst other groups of people, it is perfectly acceptable to express a want directly. We also need to know whether it is polite ever to say "no"? And what happens if you do something wrong, or inconvenience someone? In some cultures it is polite both to apologise and to give an excuse. Multilingual societies, such as ours, present a broad range of behaviour patterns, several different frameworks of acceptability.

We need also to consider how people of different cultures talk about sensitive topics, such as death or sexual relationships. Indeed, is it appropriate to talk about such topics at all? British people, and probably many others, especially those who are creeping past midlife, may avoid talking about death, or, if they need to talk about someone who has died recently, may speak more softly and use expressions like *pass on* or *pass away* rather than use the bleak and callous-sounding *die*. They use a euphemism. This may be to show respect for the dead, but is sometimes also due to a superstition that talking about death may

hasten your own. There is a feeling that it may bring bad luck, that it may be "tempting fate". In such situations there needs to be cultural as well as lexical sensitivity. And what about the funeral arrangements? The type of occasion, and whether you are invited to a *wake* or perhaps an *after-tears party*, can say much about attitudes to death and people's ease with the subject.

In the case of relationships, some cultural groups may talk frankly about very personal matters, others may not feel it appropriate to discuss relationships at all. A misguided question may be considered extremely rude and put up barriers to further communication. We need to know about words and their meanings, but we also need to know about attitudes, manners, and social norms.

Choice of words, ways of approaching certain subjects, knowing which subjects to avoid with whom, and the level of detail appropriate to a situation are difficult for outsiders to know yet vital to communication at all but the most basic level. Correctness and intelligibility may rest on purely linguistic factors, but successful communication depends on an adequate knowledge of language and an understanding of what it is appropriate to say in particular situations. Gumperz (1970: 205) comments: "Whereas linguistic competence covers the speaker's ability to produce grammatically correct sentences, communicative competence describes his ability to select, from the totality of grammatically correct expressions available to him, forms which appropriately reflect the social norms governing behaviour in specific encounters."

Communication requires a degree of lexical and syntactic competence, and cultural competence. We cannot operate in a cultural vacuum. In South Africa, a country where people are being positively encouraged to learn another language and to communicate across cultures, the need to develop cultural as well as linguistic awareness should be emphasised. Successful communication involves background knowledge, shared information about context, traditions, and attitudes, shared images in the mind's eye. Problems of cross-cultural communication are not necessarily a misunderstanding of words, but a lack of understanding of the concepts behind words. A major source of information for many language learners is a bilingual dictionary. Authors and editors of bilingual dictionaries know just how difficult it is to find "semantic equivalents" even for languages of similar cultures. There is often a need to offer the user a "communicative equivalent", which takes into account cultural factors. Both Gouws (1996) and Kotzé (1999) have addressed these issues in volumes of *Lexikos*. I would suggest that many a dictionary user would like to know more about the cultural differences which cause such problems and the reasons for the selection of a particular communicative equivalent.

Successful communication would seem to involve more than a knowledge of semantics and syntax. As lexicographers, we need to consider the implications for dictionary compilation. Dictionaries clearly have a responsibility for imparting or enhancing knowledge of a language. They provide lists of words with definitions and other very specific categories of information. Generally,

they do not provide explanations of cultural behaviour. In the past year, in lexicographic seminars arranged by the Pan South African Language Board, there has been much emphasis on the macrostructure and microstructure of dictionaries. So far, there has been little or no discussion of any cultural dimension. The relationship between language and culture is widely recognised, and I believe now is the time in South Africa to review that relationship vis-à-vis dictionaries.

Culture and language

Let us have a closer look at what we mean by "culture", and especially the relationship of culture with language. What is culture? Culture is a word with various different senses. In one sense it may include the arts — the theatre, literature, music, and painting — which are the products of human intellectual achievement. In another sense it may refer to the whole gamut of traditional beliefs and practices, activities and way of life, of a particular group of people.

In connection with the first sense, which may be described as "achievement culture", or "culture with a capital C", we can think of famous South Africans like J.M. Coetzee, Zanemvula Mda, Miriam Makeba, Hugh Masekela, and many more. As a result of their achievements they have become part of a national South African culture and are becoming more widely known. We can think also of internationally famous people and works of art — Beethoven and his symphonies, Van Gogh and his paintings, Shakespeare and his poetry and plays. These people and their achievements are no longer associated only with an ethnic or national culture but have taken their place on the international stage. They are part of "high culture", which may sometimes be referred to as "civilisation". Civilisation originally embraced the development of society and the arts in Europe, but it has since absorbed cultural achievement from the rest of the world. In a paper presented to the English Academy, Laurence Wright (2000, in press) said of civilisation that it "comes into being through gathering together and comparing the character and results, the achievements of different cultures. It is inherently transcultural, to some extent metacultural".

A more challenging approach is to focus on the behavioural or social aspects of culture, the second sense of culture that I mentioned. This is sometimes referred to as "behavioural culture" or "little c" culture. The anthropologist Walter Goodenough (1957), quoted in Wardhaugh (1986: 211), concentrates on this aspect of culture, saying: "The culture of a society consists of the particular knowledge and beliefs that members of a community have which enable them to operate in a manner acceptable to the group, and in any role that the group finds acceptable". Gabriele Becher (1995: 294), quoted by Kotzé (1999), says on the meaning of culture that it is "the sum total of a social community's behaviour patterns, including the 'rules' of behaviour and its (material and immaterial) results". Becher also relates language to culture, saying that "lan-

guage is considered to be a cultural product, and at the same time, language may be considered part of the manifestation of culture-specific behaviour".

This sense of culture is more about ordinary people and their daily life, how they live, the places they go to, the institutions they have contact with, their social interactions, their attitudes towards the issues and situations that confront them, their beliefs and perceptions, and their priorities, i.e. "what makes them tick". "Little c" culture also includes aspects of life such as religious practices, education and employment, leisure activities, and sport. It is about people and their behaviour and values, and may be specific to a particular group of people or nationality. Knowledge of this type of culture is more relevant to language users in terms of communication, but it is also more difficult to pin down. It is more likely to be changing and developing, and has a range of manifestations dependent on age, gender, class, and education.

Language is important as a means of ethnic or cultural identity. An important function of a language is to enable speakers of that language to identify with a group or to separate themselves from it. We need to be aware that we all have "cultural baggage", ideas and patterns of behaviour that were absorbed or instilled into us as children, and we bring this consciously or unconsciously into our dealings and conversations with others. We need also to be aware that other people are doing the same.

The knowledge, beliefs, and practices of a particular society are reflected in its language. The vocabulary of a language can provide some evidence of what is considered culturally important, and there may be several words for the most familiar items and concepts. The American linguists Edward Sapir and Benjamin Lee Whorf (discussed in Kaschula and Anthonissen 1995, and Wardhaugh 1986) suggested that people's views of the world are conditioned largely by their mother tongue. Though the extent to which language itself determines and restricts the way in which speakers understand the world is in doubt, it seems clear that attitudes and language are closely interlinked. Social change produces linguistic change: new inventions and ideas bring new vocabulary, and in multicultural societies greater contact brings code-switching and code-mixing, borrowing, and assimilation.

Modern language-learning strategies promote familiarisation with the culture of users of the language. Course books include cultural information of all kinds, social, artistic, and institutional. English language course books, such as the *Headway* and *Hotline* series from Oxford University Press, and Patricia Pinnock's Xhosa course book, *Xhosa: A Cultural Grammar for Beginners*, help the learner become more aware of differences in lifestyle, heritage, and values. Such cultural frames of reference are especially important when the speaker's first language and the language being learned are linguistically and physically remote. In multicultural societies such as South Africa people have more opportunity to become sensitive to others' social practices through firsthand experience.

Linguistic and cultural items are often one and the same and cannot be easily differentiated. People from different cultures have varied mental images of apparently simple items like *house*, *taxi*, even *jelly*. What about more complex terms at the centre of human life, like *family*, *marriage*, *funeral*, *manners*, and *government*? We all know these words, but we may have different images in our minds, different ideas of their meaning. For each of us the connotations are different. Take the word *marriage*, for instance. Does the term for you mean only one wife, or is it perfectly normal and acceptable to have several wives concurrently? Should you have one wife till death do you part, or is divorce possible? If so, can you remarry? Does marriage imply a religious service, or is it a civil contract? What is a common law wife? a customary marriage? And what about associated behaviour patterns? For instance, do the partners in a marriage choose each other, or do their parents choose for them? What is the relative status of each of the partners? What are their expected codes of behaviour — are they expected to be faithful to each other? Is romance involved? Is marriage seen rather as an alliance between families? Does any payment, money or goods, change hands? Many more questions can be asked, but it seems to me that a cultural framework is necessary in order to interpret the meaning of a term like *marriage* satisfactorily.

Chaffey (1992) talks about the "degree of cultural loading" to refer to the amount of culture-specific knowledge a language user must have in order to be able to understand the full meaning of a lexical item. Social terms, like *marriage*, have a high degree of cultural loading.

Names also carry cultural information and have particular connotations. I am not thinking just of people's names but also the names of products, shops, companies, and institutions. They play an important part in the understanding of an expression but are rarely found in dictionaries. To understand simple sentences like "The bus dropped us at Sainsbury's" and "You can get them at Boots", you need to understand that these are names of shops, the former a supermarket, the latter a pharmacy chain. You could substitute "Spar" and "Clicks" in South Africa. Names of people and organisations also carry important connotations. If you hear a "household name", like Margaret Thatcher, Nelson Mandela, Wall Street, Microsoft, or de Beers, you do not just know that these are names of people, places, companies, and such like, but you also know something about them, their reputation, the cultural connotations behind the names. David Crystal (1997) says that "a cultural perspective is necessary ... to grasp the social significance of a linguistic feature ... or to follow the subject matter of daily conversation". Knowing the cultural connotations of a word, knowing the names of famous people and what they stand for, knowing about major events in a people's history, knowing its institutions, and even knowing the names of shops or brand names, are major factors in understanding a newspaper report or television programme, or just taking part in a casual conversation.

In South Africa we have eleven official languages. The claim of each language to equal status has little to do with linguistic merit, but is almost entirely

about cultural and social factors: equitable treatment of people from different cultures and linguistic backgrounds, respect for human rights, respect for language rights, a desire to preserve long-standing cultural and social traditions, and more. Language is inextricably bound to culture, but yet we often separate them and try to avoid including cultural information in dictionaries on the grounds that it is extraneous or, more pragmatically, that the space can be more than taken up by lexical information. It is impossible to include everything, but we can try to give a fuller cultural picture in many instances.

Dictionary or encyclopedia?

It may be thought that cultural information should be presented in an encyclopedia rather than a dictionary. What is a dictionary, what is an encyclopedia? The difference may appear to be obvious. Think of a dictionary such as the *Concise Oxford Dictionary*, which has entries for words such as *heritage*, *win*, and *although*, or a bilingual English-Afrikaans dictionary, which has information about the same words in two languages. Then consider an encyclopedia like the *Encyclopaedia Britannica*: titles of articles in this work include *fungi*, *brewing*, *funerary rites and practices*, *Tombouctou*, and *Mao Tse-Tung*.

Traditionally, dictionaries are supposed to deal with the question "What does *x* mean?", whereas an encyclopedia deals with the question "What is *x*?" or "Who is *x*?" or "Where is *x*?" An encyclopedia contains encyclopedic information, information about world-famous people, places, and institutions; politics and politicians; characters in literature, the theatre, and cinema; societies and social practices; the universe and natural phenomena. An entry for *weather* in a dictionary will give very different information from that contained in an encyclopedia. It has been neatly said that encyclopedias give information about the *world*, rather than about the *word*. Information in encyclopedias is presented in articles which may be a few lines or a few pages long. They may be arranged alphabetically or thematically. Compare this with the brief, sometimes over-concise information about the *words* of a language that is presented in a dictionary.

The dictionaries we, as lexicographers, spend much of our time discussing present the bones of a language. They are general lexical dictionaries and give what are sometimes referred to as "dictionary definitions". The presence of purely linguistic information usually signals a dictionary rather than an encyclopedia. We talk often about distinct dictionary "types", but the concept of a dictionary may become quite fluid. In many dictionaries, definitions include an element of description over and above what is necessary to isolate the concept. The information may typically describe the function or use of the item concerned. Common examples are definitions for *sugar* or *salt* which describe the use of the substances. These are known as "encyclopedic definitions". Laurence Urdang (2000) reflects on the difference between the two types of definitions in his review of Hartmann and James's *Dictionary of Lexicography*. Incidentally, he

also comments that "attempts at divorcing language from life and treating it as some sort of algebraic formula that was found somewhere engraved on a stone may be regarded by some as logically rigorous, but it is not only inelegant but useless". Dictionaries are often regarded as "containers of knowledge", and dictionary users may be less worried about discriminating between lexical and practical information than lexicographers. As Béjoint (1994: 18) says: "A dictionary is primarily an instrument to be used when one needs a piece of information about a word in particular, about the language in general, or about the world."

Let us look briefly at the dictionaries called "encyclopedic dictionaries" and the sort of information they provide. They are alphabetically arranged and contain linguistic information and encyclopedic or cultural information in the same volume. We might expect them to include more than simply lexical information about "little c" cultural terms such as *family*, *manners*, *drinking*, or *birthdays*, but in fact they rarely include this type of information. Nor do they include entries for locally famous people, places, events, etc. Whatever the language they are catering for, they are more likely to include information about Japan or Paris, Johann Sebastian Bach, or George Washington. This kind of information is part of general world knowledge, world culture, or *civilisation*. I would like to suggest that dictionaries need to contain more locally based items, and some elements at least of "little c" culture.

The tradition of including cultural material in dictionaries

The tradition, especially in the US, of including cultural, or *encyclopedic*, material in dictionaries can now be considered.

In Britain, the inclusion of encyclopedic information in dictionaries was denounced by Dean Richard Trench in 1857, in papers entitled "On some Deficiencies in our English Dictionaries". His grounds for rejecting such information were that it was irrelevant and it cluttered up the dictionary. This was just prior to the start of the *New English Dictionary* project, what was later to become the *Oxford English Dictionary*, and Trench had a profound influence on modern dictionary-making. Encyclopedic information was common in earlier dictionaries, including Samuel Johnson's of 1755, and is still a standard feature of many American dictionaries.

English dictionaries of the 16th and 17th centuries contained information on mythology, the theatre, religious sects, folkloric properties of flowers, and names and descriptions of cities and countries. During the 17th and 18th centuries the authority of the lexicographer increased. Mengham (1993: 112) says: "The lexicographer would determine what should be included in, and what should be excluded from, a body of knowledge that the pragmatic user of his work would learn to regard as the foundation of national language and culture." Johnson and later Noah Webster in the US held positions of power from which they could develop their own moral and political agenda in their dic-

tionaries. Mengham considered that, in Johnson's case, he used his position as a means of "reinforcing the political status quo, religious authority, and cultural authoritarianism ...". A salutary reminder to beware of the dangers of bias in the selection and presentation of encyclopedic information.

In the US there has been a long-standing tradition of lexicographical prescriptivism. In the 18th and early 19th century there were unsuccessful attempts to set up an academy to legislate on linguistic standards. However, the success of Noah Webster's *American Spelling Book* (1783) and *A Compendious Dictionary of the English Language* (1806), and later works of Webster and Joseph Worcester, went some way towards satisfying desires to establish linguistic as well as political independence from Britain. In particular, Webster's *American Dictionary of the English Language* (1828) is a byword for patriotism and nationalism.

Since then, the objectives of American lexicography have been not so much to preserve the language but to indicate good usage and pronunciation. This was especially important following the immigration of people from many countries to the US in search of a better life. Dictionaries were, and still are, tools for education and self-improvement. They were seen as a means of educating foreigners about American ways, American culture, and American English, and regularly included encyclopedic as well as lexical and syntactic information. Dictionaries were also seen as a sign of social and cultural attainment and had an important place, second only to a Bible, in most homes. These dictionaries were compiled for the ordinary person, not the scholar. They were all-purpose reference books, the primary and often the only source of information. They contained scientific, technical, geographical, and biographical information, distributed through the main alphabetical list and in appendices. Webster's *New International Dictionary* (1909), named after Webster, not edited by him, contained sections on place names, quotations, vocabulary tests, proof-readers' marks, and more. A later edition, *Webster's Third New International* (1961), was criticised strongly for departing from this tradition and for becoming descriptive rather than prescriptive. It had no proper names apart from attributive uses, fewer illustrations, and no encyclopedic supplement. The expectations of the public were seriously disappointed.

Today, "the dictionary" is still important in the US. American dictionaries are intended to be user-friendly and satisfy the needs of the average person. They are also highly commercial products and many publishers outside the US have taken note of this. Some modern British dictionaries have followed the American line, notably the *Collins Concise Dictionary*, and the *Oxford English Reference Dictionary*.

Cultural information in modern dictionaries

Assuming the desirability of including encyclopedic or cultural information in dictionaries, we all know that printed dictionaries have space constraints, that

every word has to justify its inclusion. Including cultural content, however interesting and useful to dictionary users, probably means having less lexical content in all but the largest dictionaries. There is certainly little room in bilingual dictionaries, unless they are written on a large scale, though it is here that cultural information may be most appreciated. However, in spite of the constraints, many dictionaries do manage to include a lot of encyclopedic information. Let us look at a few to see what sort of information is given, and then at how it is presented.

Starting with a modern American dictionary, we can consider the second edition of *The Random House Dictionary of the English Language* (1987). This is a large desk dictionary of 2 478 pages and so has considerable scope for encyclopedic information. Within the A-Z section there are names of places in the US and elsewhere. There are also entries for geographical regions, periods of history, even Chinese dynasties. Names of people range from Roman emperors to modern comedians. There are entries for musical compositions, plays, fictional characters, constellations, common plants and birds, etc., etc., and an atlas section — information about the world, not just about the US, though data about the US figure most prominently. There are lists of Chief Justices and Presidents, US States, and US National Parks. Appendices include a directory of colleges and universities in the US and Canada, the Declaration of Independence, and the Constitution of the US. In sum, there is a great deal of general knowledge, specially geared to people in the US, but there is little about daily life, people's attitudes, or patterns of behaviour, except what can be gleaned from the lexis itself.

Turning to British dictionaries, we can look at the fourth edition of the *Collins Concise Dictionary* (1999). The blurb on the back cover says it includes "more people", "more places", "more facts" in its 1 740 pages. It includes people who are household names in all fields, past and present, and facts about countries, cities, and towns all over the world, though it is definitely a British-based dictionary. It describes itself as "an especially handy reference book for the home, college, or office, whether you are studying, writing, or just keeping up with the news". Though there is information on mythology, medals, plants, and animals, most of the entries deal with world-famous people and places. There is no information about daily life or "little c" culture.

The desire to concentrate on general knowledge and achievement culture is not unique to English dictionaries. In France, Larousse is a well-known publisher of both dictionaries and encyclopedias. The *Petit Larousse Illustré* (1998, 1 784 pp) is a monolingual French dictionary and encyclopedia combined. The first alphabetical section contains standard French vocabulary items. The second section contains encyclopedic information: names of famous people, places, buildings, and works of art from all around the world. The book is extensively illustrated with colour photographs, and is excellent as a general reference book of "western civilisation", certainly not an account of French life and culture.

Dictionaries for learners of a language are probably the place where information about the culture of the people whose language they are learning is most needed. Unfortunately, in learner's dictionaries, such as the *Oxford Advanced Learner's Dictionary (OALD)*, there is also the greatest need for information on grammar and synonymy, and for example sentences. Cultural information in such dictionaries is not a priority, but it is increasingly recognised that learners do need cultural information to aid their understanding of written and spoken English. A few years ago, an "encyclopedic" version of the *OALD* was produced. As well as the full dictionary text, this book contained information about countries of the world, famous people and places, famous events and battles, etc., though nothing about the lives of ordinary people. Longman also produced an encyclopedic version of the *Longman Dictionary of Contemporary English (LDOCE)* containing similar information.

In the mid 1990s, when new editions of these books were being planned, Longman retained the combined format and updated the encyclopedic information. At Oxford, we decided to do something different. Research amongst students and teachers of English as a second or foreign language showed that learners wanted to know more about the humbler aspects of British and American culture, effectively the "little c" culture. They could find out about world events and "civilisation" from reference books in their own languages. We developed a list of single items and topic areas, and decided to publish the information as a companion volume to the *OALD*. The *Oxford Guide to British and American Culture*, as it is called, contains information about *all* aspects of British and American culture, and is written in a restricted vocabulary suitable for learners.

In Britain the *Guide* is also bought by native English-speakers, especially those with children, since it is a concise reference book containing information on British and American history, politics, the arts, legends, customs, people, places, and institutions, all in an A-Z format. But the *Guide* also describes how British and American people live, the sports they play, the products they buy, the food they eat, and events that have shaped their lives. Many newspaper and magazine articles allude to people and events that are part of the national consciousness. The *Guide* provides factual information, but also indicates the connotations which some words have for a native speaker. For a full understanding of a passage, readers need to know, for instance, what names such as *Richard Branson*, *Burberry*, and *Bloomingdale's* signify, just as much as they need to know the meaning of *expedition* or *balloon*. But the *Guide* is more than a list of proper names. It contains extended definitions of words like *panhandle* and *pier*, and notes and articles on topics ranging from *folk dancing* to *cricket*, from *class* to *soap operas*, from *manners* to *birthdays*. Articles such as these present information on daily life, common attitudes, and behaviour. This book is published in a separate volume from the associated dictionary, but the information it contains could just as easily form a second section of a lexical dictionary, as in the case of the Larousse dictionary mentioned earlier.

But let us look now closer to home, in South Africa. Let us consider first the *Woordeboek van die Afrikaanse Taal*, specifically volume X (1996), which covers the letter M. Most of the cultural information in this dictionary is to be found in the citations. There are few cultural as opposed to lexical entries, and many of these are Biblical, e.g. *mure van Jerigo* (walls of Jericho) and *Man van Smarte* ("man of grief", i.e. Christ). I also found *Matie* (a student at Stellenbosch), and *Matieland*, and *makweta*, a young black man undergoing the time of initiation in isolation. There are also names of several well-known organisations.

In the case of the *Dictionary of South African English on Historical Principles* (1996), where only South African English, not the entire English language, has to be covered, there is scope for a broader inclusion policy and more detailed treatment of cultural items. Once again, much of the cultural information is to be gleaned from the citations. However, there are entries for many South African institutions, historical and modern, customs, peoples, places, and regions, as well as local plants and animals.

Although not setting out to provide cultural information *per se*, both these dictionaries, because they describe language used mainly in South Africa and draw on its literature for their citations, carry an enormous amount of information about people and their lives. But you do need to grub and fossick around for it.

The *Greater Dictionary of Xhosa* (1989) contains more overtly cultural information throughout the alphabetical section. This is a descriptive and a translating dictionary, covering Xhosa, English, and Afrikaans. It also has a number of addenda containing both grammatical information and descriptions of a wide variety of Xhosa customs and traditions. The foreword to the first volume published, written by John Lamprecht, the then Vice-Chancellor of the University of Fort Hare, sums it up: "The compilers of this Dictionary have succeeded in capturing much of the living breath and feel of the Xhosa world. Here is no mere cold analysis, but a work that breathes the life of the Xhosa peoples themselves."

So far, we have looked at printed works. The world of CD-ROMS and online dictionaries may seem to hold out the possibility of including as much lexical and cultural information as we might desire. Not so. CD-ROMS are expensive to develop. The writing process takes just as long, and since they are a very visual medium, users expect more illustrations. Rights to these are often costly. In many cases the development costs would suggest a purchase price too high for potential users. Online dictionaries are plentiful, and, if you have an internet subscription, free to access. Some are linked to encyclopedias, but producers of these try to recoup costs by restricting access to some of the information to subscribers only. The most famous online products are *Encarta*, a Microsoft production, and *Britannica.com*, the online version of the *Encyclopaedia Britannica*.

The presentation of cultural information

I am going to look now at some of the most common ways of introducing cultural material in dictionaries.

In encyclopedic dictionaries, especially, a whole entry may deal with a cultural item. In the *Random House Dictionary of the English Language* there are entries for *Mozart*, *Fort Sumter*, the bombardment of which started the American Civil War, the *Nile river*, and *Nobel Prize*, amongst many thousands of others. The *Collins Concise Dictionary* is similar. Many general-purpose dictionaries also include entries for culturally-based items, such as *Birthday Honours*, *Phi Beta Kappa*, and *Rastafarian*.

It is possible to include brief cultural points within the definition of a lexical item. An extra phrase or sentence may be sufficient. I mentioned encyclopedic definitions for *salt* and *sugar* earlier. In the *Concise Oxford Dictionary*, *salt* is defined as: "sodium chloride, a white crystalline substance which gives seawater its characteristic taste and is used for seasoning or preserving food". This does not take up much extra space but makes for a more informative definition. In the *Collins Concise Dictionary* the entry for *eucalyptus* indicates that the trees are "widely cultivated for timber and gum, as ornament, and for the medicinal oil in their leaves".

In the *Greater Dictionary of Xhosa* definitions are extended to include cultural information where appropriate. For example, there is cultural information at *intozánè*, where the English version of the text reads:

a forest tree, *Englerodaphne* sp, the bast of which is used for making thread for weaving sleeping-mats.

Again, more extensively, at *úkùsindlékà*, where the English version reads:

trad: prepare a feast for an unexpected member of the family, a relative or visitor; a beast is slaughtered and beer brewed or only beer is prepared; now also: prepare a feast quickly for an unexpected visitor by buying meat and liquor, or serve beer prepared in the traditional manner; friends and neighbours are invited to the occasion.

The material presented at the entry *ixhwêlè* is even more comprehensive:

herbalist who uses roots, leaves and bark of a large variety of medicinal plants in the form of infusions, decoctions, pastes, powders or steam; certain parts of animals (mammals and reptiles) are also used; the medicines are drunk, usually as purgatives, applied externally as washes or plasters, used for gargling, or as enemas, or rubbed into incisions in the skin; besides treating people, herbalists also treat kraals, homes, livestock and crops to protect them from evil influences and to make them

prosper; the knowledge of these medicines may be obtained from a relative, eg a father, or directly from the ancestral spirits who reveal them in dreams; the herbalist operates in the fields of both white and black magic, ie his skill is used to cure ailments and drive away evil influences, alternatively also, when requested, to harm or kill people in a supernatural manner; whereas in the past the herbalist collected medicinal materials himself, many of them are now bought from people who trade in such things.

These definitions, or perhaps they should be called explanations, which are given in Xhosa, English and Afrikaans, are, I believe, one of the great strengths of this dictionary.

Bilingual and multilingual dictionaries need to consider the social and intellectual background of both or all the speech communities they serve. A more detailed explanation of meaning and connotation may be necessary to aid understanding of an item not known to one or more of the speech communities. In the *Tshivenda-English Dictionary* edited by van Warmelo cultural snippets follow the translation. At *khwará*, "pangolin", for instance, we find the following extra paragraph:

Khwará are said to be seen only after rains, therefore believed to fall from the skies with the rain. Their shells therefore serve to make rain, but shedding their blood (if done by the uninitiated) would make the rain stay away. Any man finding a khwará had to carry it to the chief's place alive without delay. Since it can escape no matter how well secured, such a man had no rest until rid of it.

The same dictionary gives fascinating interpretations of Venda proverbs. Again, in Cuenod's *Tsonga-English Dictionary* we find entries like *nthimbu*:

one of betrothed girl's friends (female), who dances on the day the *ndzovolo* cattle for her are brought.

To purists all this may sound a little quaint and out of place in a dictionary, but where else is this kind of cultural information, information associated with individual words, going to be recorded?

In dictionaries such as the *Woordeboek van die Afrikaanse Taal* and the *Dictionary of South African English on Historical Principles*, citations reveal the development of meaning and usage over time but also, as we have seen, provide a wealth of cultural information, past and present. In a similar way, example sentences in learner's dictionaries and elsewhere, when carefully chosen, often provide a cultural context for a word as well as an illustration of usage.

An attractive way of including more detailed cultural information is to add a short note after the lexical entry. A note may fall within the column

structure for the page, but yet allow cultural details to be presented separately from information on grammar and usage. Design features are important, so that the page does not look top heavy. In the *Reader's Digest Afrikaans-Engelse Woordeboek / English-Afrikaans Dictionary*, extra columns on either side of the two main columns contain shaded boxed notes. Most of the notes in this dictionary are to do with usage rather than culture, though some give cultural information. In the *Oxford Guide to British and American Culture* there are many short notes on topics such as *adult education, capital punishment, homelessness, lobbying, nicknames, and nightlife*. Since this is not a lexical dictionary the notes stand independently. The notes are preceded by the headword and begin and end with a ruled line to differentiate them from the rest of the A-Z text.

The *Guide* also includes a set of longer notes or articles on cultural topics, placed as close as possible to their alphabetical position. These are listed in the contents page to aid readers in finding them, and to encourage people to browse. These articles are on major topics, such as *American English, Class, Education, Parliament, and Religion*, and often take up a whole page or even a double-page spread. There are also lighter articles on *Pubs and Pub Names*. Articles are differentiated from the general text by being framed and presented in a different font. The book is intended primarily for learners of English, and key vocabulary for the subject under discussion is highlighted in bold type. Words preceded by an asterisk have their own entry in the *Guide*.

The *Guide* also has a set of colour pages in the middle of the book, on umbrella topics such as *The history of the US* and *The British landscape*. Colour pages are expensive and rarely affordable in dictionaries, but they do make the book attractive to potential purchasers. A black-and-white inset section would be a reasonable substitute and would still allow important cultural topics to be given prominence.

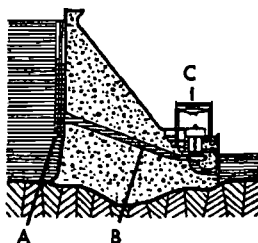
Some information can best be presented as a list, or as a table or chart. The *American Random House Dictionary* contains charts and tables within the A-Z text, some within the column width, some taking up more of the page. Topics include *birthstones, Islands of the World, and United Nations Member Countries*. The *Oxford English Reference Dictionary* puts charts and tables in a 100-page "factfinder" section, but perhaps the most common way of including these in general-purpose dictionaries is to put them in appendices or addenda. Common topics for appendices include lists of signs and symbols, lists of countries and their capital cities and currencies, and advice on punctuation and how to set out a letter or curriculum vitae. The *Greater Dictionary of Xhosa* contains a series of addenda containing information on grammar, e.g. noun classes and object concords, but most addenda in this book are numbered paragraphs describing various aspects of the Xhosa cultural heritage. Subjects include *Tikoloshe* (water sprite), *Mamlambo* (snake), *Ntsikana* (a famous person), *Ukuhlolela*, (searching for a husband), the brewing of beer, and being summoned by the river.

Illustrations are an important means of presenting cultural information, though dictionary editors need to be aware of the high costs involved and that they, rather than the publishers, may be expected to pay for illustrations. Photographs, especially, enhance a book, but reproduction rights are expensive and good quality paper, also expensive, is necessary. Line drawings are a cheaper alternative to photographs, though research shows that these are easily misunderstood by users unfamiliar with certain techniques and conventions, such as cross-sections. Look at the illustration of *dam*¹, for instance, from the *Random House Dictionary*:

dam¹ (dam), n., v., **dammed**, **dam-ming**. —n. **1.** a barrier to obstruct the flow of water, esp. one of earth, masonry, etc., built across a stream or river. **2.** a body of water confined by a dam. **3.** any barrier resembling a dam. —v.t. **4.** to furnish with a dam; obstruct or confine with a dam. **5.** to stop up; block up. [1275–1325; ME < MD, MLG, *dam*; akin to OE *for-demman* to stop up, block]

—Syn. **5.** impede, clog, check, choke.

dam¹
(def. 1)
A, trash rack;
B, penstock;
C, powerhouse



Unless you already know what a dam in the US is like, and how it works, you might find the illustration difficult to interpret. Most illustrations are contained within the text, immediately before or after the entry to which they refer. Alternatively, they may be given as a group in the centre of the book, though this is less satisfactory as they are remote from the entry they illustrate.

Maps are frequently added to dictionaries, often in bilingual dictionaries to show names of countries in the target language. In larger dictionaries there may be a series covering the whole world. In other dictionaries maps may cover only the particular region or country where the dictionary is to be marketed and are more detailed. Some maps are in colour, but a black-and-white drawing is often just as useful.

Endpapers may also be used to present cultural information, if they are not needed for lists of abbreviations or symbols. Topics presented on endpapers may include *flags*, *timelines*, or *lists of presidents*. Timelines can give a graphic summary presentation of local political and social history which would not be appropriate in an A-Z section. Maps may also be presented on endpapers.

Implications for South African lexicographers

We can now draw a few threads together and think of the cultural and lexicographic situation in South Africa.

South Africa is multilingual and multicultural. Although there are some shared aspects of culture, such as political institutions, the justice system, sports, and music, many traditions, practices, events, and heroes are little known outside individual cultural groups. For the Xhosa people, at least, the Xhosa Dictionary Unit is recording a fair amount of cultural information, which is accessible not only to Xhosa speakers but also to English and Afrikaans speakers. Though the dictionary may be targeted primarily at the Xhosa community, its trilingual format means it is more widely accessible. This is something that needs to be seriously considered.

It is impossible to talk about language and culture now without mentioning the African Renaissance. The African Renaissance promotes African culture. It is trying to encourage renewed awareness of Africa and things African, national and local, amongst outsiders, and to engender feelings of honour and pride in their various cultures among Africans. It reconfirms respect for older societies and customs but also encourages ongoing creativity. Laurence Wright (2000, in press) quotes Paul Ricoeur (1965: 283): "Only a living culture, at once faithful to its origins and ready for creativity on the levels of art, literature, philosophy and spirituality, is capable of sustaining the encounter [with] other cultures ..." We need to develop cultures, not just preserve them, and to make them known to a wider audience. Culture is not something that will survive without being nurtured or "cultivated", and the responsibility for this needs to be shared.

Writing cultural dictionaries or reference works in the South African context sounds like an ambitious undertaking. But perhaps it is something we need to consider. I mean specifically books that are useful for schools, students of culture, and for members of other cultural groups, "dictionaries of culture" on the lines of the *Guide to British and American Culture*. If this seems beyond our remit or too ambitious for the moment then perhaps, as lexicographers, we can consider including more cultural information in dictionaries. As we have seen, language and culture are closely related, and all sorts of cultural information can be pinned to a word or name. Who else is in a position to gather together all the threads? Dictionaries containing encyclopedic information, provided they are of a manageable size, often appear more user-friendly, more interesting, and browsable to the general public. They are seen as general reference books and good value for money.

"Dictionary" is a powerful word, and it lends authority to the information contained in any book with "dictionary" in the title. Many lists of subject terms, on topics ranging from psychology to geology, are called "dictionaries". Dictionaries of terminology, learner's dictionaries, and cultural dictionaries, can all be classified as "dictionaries for special purposes". Perhaps there is scope for

future collaboration between dictionary units on a range of cultural dictionaries. Culture can be divisive, insular, and threatening, especially if it appears closed to others. More widely available information may lead to greater interest, greater acceptance.

Transformation of the curriculum in South Africa is still in progress, and we do not yet know what exactly will be included in Curriculum 21. But it will be outcomes-based, and will probably incorporate many features of Curriculum 2005. The latter recognised "Arts and Culture" as a learning area and targeted the following "specific outcomes":

1. demonstrating an understanding of the origins, functions and dynamic nature of culture,
2. demonstrating an ability to access creative art and cultural processes to develop self-esteem and promote healing, and
3. acknowledging, understanding, and promoting historically marginalised arts and cultural forms and practices.

Another highly relevant learning area is, or was, "Language, Literacy and Communication". Specific outcomes here include:

1. showing critical awareness of language use, and
2. responding to aesthetic, affective, cultural and social values in texts.

Culture, especially "little c" culture, is also very much involved in the "Life Orientation" learning area. Students should learn to:

1. respect the rights of people to hold personal beliefs and values, and
2. demonstrate value and respect for human rights as reflected in *ubuntu* and other similar philosophies.

Culture, particularly behavioural culture, is right at the heart of the humanities curriculum, closely linked to language and literacy and also to social studies.

It may be difficult to "teach" culture, but it is possible to help people to become more culturally aware, aware of their own cultural behaviour and aware that other people's cultural practices, though they may be different, are equally valid. Awareness involves having information, information about society and social practices, history, traditions, and values, and I hope we can be amongst those instrumental in supplying it. With this in mind, I would like us to look afresh at the level and type of cultural content of our dictionaries, and to consider whether it meets the needs of prospective users.

References

- Becher, Gabriele. 1995. Culture and Institutions in Translator Training. Teaching Culture-specific Concepts to Prospective Translators: A Case Study. Thelen, M. and B. Lewandowska-Tomaszczyk (Eds.). 1995. *Translation and Meaning, Part 3*: 293-304. Maastricht: Universitaire Pers.
- Béjoint, Henri. 1994. *Tradition and Innovation in Modern English Dictionaries*. Oxford: Clarendon Press.
- Chaffey, William. 1992. Culture-specific Elements in Translation. Lewandowska-Tomaszczyk, B. and M. Thelen (Eds.). *Translation and Meaning, Part 2*: 147-153. Maastricht: Universitaire Pers.
- Crystal, David. 1997. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Goodenough, W.H. 1957. Cultural Anthropology and Linguistics. Garvin, P.L. (Ed.). *Report of the Seventh Round Table Meeting on Linguistics and Language Study*. Washington, DC: Georgetown University Press.
- Gouws, Rufus H. 1996. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 14-31.
- Gumperz, J.J. 1970. Sociolinguistics and Communication in Small Groups. Pride, J.B. and Janet Holmes (Eds.). 1972. *Sociolinguistics: Selected Readings*. Harmondsworth: Penguin.
- Kaschula, Russell H. and Christine Anthonissen. 1995. *Communicating across Cultures in South Africa: Towards a Critical Language Awareness*. Johannesburg: Hodder & Stoughton/Witwatersrand University Press.
- Kotzé, Ernst. 1999. Translating Culture in Bilingual Dictionaries. *Lexikos* 9: 89-107.
- Mengham, Rod. 1993. *The Descent of Language: Writing in Praise of Babel*. London: Bloomsbury.
- Pinnock, Patricia Schonstein. 1994. *Xhosa: A Cultural Grammar for Beginners*. Cape Town: African Sun Press.
- Ricoeur, Paul. 1965. Universal Civilization and National Cultures. *History and Truth*: 271-284. (Trans. Charles A. Kelbley). Evanston: Northwestern University Press.
- Urdang, Laurence. 2000. Review of Hartmann, R.K.K. and G. James: *Dictionary of Lexicography*. *International Journal of Lexicography* 13(1): 35-42.
- Wardhaugh, Ronald. 1986. *An Introduction to Sociolinguistics*. Oxford: Basil Blackwell.
- Wright, Laurence. 2000, in press. Culture and Civilisation in South Africa: Some Questions about the "African Renaissance". *English Academy Review* 16.

Dictionaries and Encyclopedias

- Britannica.com*. 1999-2000. Chicago: Britannica.com Inc.
- Collins Concise Dictionary*. 1999. Glasgow: Collins.
- Crowther, Jonathan and Kathryn Kavanagh (Eds.). 1995. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
- Crowther, Jonathan and Kathryn Kavanagh (Eds.). 1999. *Oxford Guide to British and American Culture for Learners of English*. Oxford: Oxford University Press.
- Cuenod, R. 1967. *Tsonga-English Dictionary*. Johannesburg: Sasavona.

- A Dictionary of South African English on Historical Principles*. 1996. Oxford: Oxford University Press, in association with the Dictionary Unit for South African English.
- Encarta Encyclopædia and World English Dictionary*. 1997-2000. Redmond: Microsoft Corporation.
- Encyclopaedia Britannica*. 1971. Chicago: Encyclopaedia Britannica/William Benton.
- Gove, Philip Babcock** (Ed.). 1961. *Webster's Third New International Dictionary of the English Language*. London: G. Bell & Sons; Springfield: G & C Merriam.
- Johnson, Samuel** 1755. *A Dictionary of the English Language*. London: J. and P. Knapton et al. (facsimile edition, 1979. London: Times Books).
- Longman Dictionary of Contemporary English*. 1995³. Harlow: Longman.
- Murray, James A.H. et al.** *Oxford English Dictionary*. 1933. Oxford: Oxford University Press.
- The Oxford English Reference Dictionary*. 1996. Oxford: Oxford University Press.
- Pahl, H.W. et al.** 1989. *The Greater Dictionary of Xhosa, Volume 3, Q to Z*. Alice: University of Fort Hare.
- Pearsall, Judy** (Ed.). *Concise Oxford Dictionary*. 1999¹⁰. Oxford: Oxford University Press.
- Le Petit Larousse Illustré*. 1998. Paris: Larousse.
- The Random House Dictionary of the English Language*. 1987¹. New York: Random House.
- Reader's Digest: Afrikaans-Engelse Woordeboek / English-Afrikaans Dictionary*. 1987. Cape Town: Reader's Digest Association.
- Van Warmelo, N.J.** 1989. *Venda Dictionary: Tshivenda-English*. Pretoria: J.L. van Schaik.
- Webster, Noah** 1783. *The American Spelling Book*. Boston.
- Webster, Noah** 1806. *A Compendious Dictionary of the English Language*. Boston.
- Webster, Noah** 1828. *An American Dictionary of the English Language*. New York: S. Converse (facsimile edition, 1980. San Francisco: Foundation for American Christian Education).
- Webster's New International Dictionary*. 1909. London: G. Bell & Sons / Springfield: G & C Merriam.
- Woordeboek van die Afrikaanse Taal*. 1950-. Pretoria: Government Printer / Stellenbosch: Bureau of the WAT.

An Integrated Semasiological and Onomasiological Presentation of Semantic Information in General Monolingual Dictionaries as Proposed in H.E. Wiegand's *Semantics and Lexicography*

Phillip Louw, *Bureau of the Woordeboek van die Afrikaanse Taal, Stellenbosch, Republic of South Africa*

Abstract: Herbert Ernst Wiegand is a very important figure in international metalexigraphy. A large part of his research has up to now been unavailable to the majority of the English-speaking world, because it has mainly been published in German. A new book *Semantics and Lexicography* seeks to break through this obscurity by providing English translations of a selection of articles spanning thirty years (from 1976 to 1996), which trace the development of Wiegand's views on semantic information. This book offers a valuable insight into the theoretical and corresponding terminological development that has already had such a remarkable impact on the "practical science" metalexigraphy.

This article focuses on Wiegand's theories on the integration of the semasiological and onomasiological presentation of semantic information in the microstructures of general monolingual dictionaries. The theories are explicated and illustrated by examples from two Afrikaans general monolingual dictionaries and evaluated in order to establish their relevance.

The conclusion is reached that the semasiological presentation of meaning is most appropriately located in the item giving the meaning paraphrase, which should form part of the semantic comment in an integrated microstructure. In the semantic comment, this item must be followed by clearly distinguished items giving onomasiological information, especially about synonymy.

Keywords: DEFINITION, FRAME, HYPONYM, ITEM GIVING THE MEANING PARAPHRASE, LEXICAL PARAPHRASE, LEXICOGRAPHIC PARAPHRASE, METALEXICOGRAPHY, NEAR-SYNONYM, OBJECT-CONSTITUTING KNOWLEDGE, ONOMASIOLOGICAL, SEMANTIC RELATION, SEMASIOLOGICAL, SYNONYM, SYNONYM DEFINITION.

Opsomming: 'n Geïntegreerde semasiologiese en onomasiologiese aanbod van semantiese inligting in algemeen eentalige woordeboeke soos voorgestel in H.E. Wiegand se *Semantics and Lexicography*. Herbert Ernst Wiegand is 'n baie belangrike figuur in die internasionale metaleksikografie. 'n Groot deel van sy navorsing was tot

nou toe ontoganklik vir die meerderheid mense in die Engelssprekende wêreld. 'n Nuwe boek *Semantics and Lexicography* poog om sy werk beter bekend te stel deur Engelse vertalings te verskaf van 'n seleksie van sy artikels wat strek oor dertig jaar (van 1976 tot 1996) en die ontwikkeling van sy sienings oor semantiese inligting uit te lê. Hierdie boek bied derhalwe 'n waardevolle insig in die teoretiese en gepaardgaande terminologiese ontwikkeling wat al so 'n merkwaardige invloed op die "praktiese wetenskap" metaleksikografie gehad het.

Hierdie artikel fokus op Wiegand se teorieë van die integrasie van 'n semasiologiese en onomasiologiese aanbod van semantiese inligting in algemeen eentalige woordeboeke. Die teorieë word verduidelik en geïllustreer deur middel van voorbeelde uit twee Afrikaanse algemeen eentalig verklarende woordeboeke en geëvalueer om hulle relevansie vas te stel.

Daar word tot die gevolgtrekking gekom dat die semasiologiese aanbod van semantiese inligting ten beste inpas in die item wat die betekenisparafrase aanbied, wat deel behoort uit te maak van die semantiese kommentaar in 'n geïntegreerde mikrostruktuur. Hierdie item moet in die semantiese kommentaar gevolg word deur deeglik afgebakende items wat onomasiologiese infor-masie, veral oor sinonimie, oordra.

Sleutelwoorde: DEFINISIE, HIPONIEM, ITEM WAT DIE BETEKENISPARAFRASE AANBIED, LEKSIKALE PARAFRASE, LEKSIKOGRAFIESE PARAFRASE, METALEKSIKOGRAFIE, NABYSINONIEM, OBJEKBEPALENDE KENNIS, ONOMASIOLOGIES, RAAM, SEMANTIESE VERHOUDING, SEMASIOLOGIES, SINONIEM, SINONIEMDEFINISIE.

1. Historical context

The last three decades have seen a tremendous growth in the academic study of the dictionary as a linguistic artefact. The publication of Ladislav Zgusta's *Manual of Lexicography* in 1971 proved to be a watershed in metalexigraphy. His detailed synthesis of insights from lexicology and the systems behind the secondary lexicographic process (i.e. the actual compilation of the dictionary) in particular, provided a foundation for other academics to build on. One such academic was Herbert Ernst Wiegand.

Wiegand has, however, not only expounded on existing theoretical frameworks. In 1984 he published his General theory of lexicography (Wiegand 1984: 15). The focus of his subsequent research has rested heavily on the second subcomponent of his theory of the lexicographic description of language: the textual theory for lexicographic texts. In this textual theory those texts which contribute to the transfer of semantic information in printed dictionaries play an important role, as these are critical to the user perspective.

A large part of his research has up to now been unavailable to the majority of the English-speaking world, because it has mainly been published in German. A new book *Semantics and Lexicography* seeks to break through this obscurity by providing English translations for a selection of articles spanning thirty years (from 1976 to 1996), which trace the development of Wiegand's views on semantic information. This development is plotted from early, pretheoretical groundwork on metalexigraphy (in which discussions take place largely

within the confines of linguistic jargon) through to lucid, metalexigraphical formulations and analyses of items that contain semantic information. The book therefore provides a valuable insight into the theoretical and corresponding terminological development that has had such a remarkable impact on the "practical science" of metalexigraphy.

However profitable a thorough discussion of the historical value of these articles may be, such a discussion does not fall within the scope of this article. Instead, it is the aim of this article to offer a synthesis of some of the most important threads running through these articles and to illustrate the practical application of the theoretical postulates. The medium for this illustration will be Afrikaans general monolingual dictionaries, with the *Verklarende Handwoordeboek van die Afrikaanse Taal* (henceforth HAT) as representative of standard dictionaries and the *Woordeboek van die Afrikaanse Taal* (henceforth WAT) as representative in the comprehensive dictionary class. The focus of the discussion will be on the integration of semasiological and onomasiological presentation of semantic information in the microstructures of general monolingual dictionaries. As not all articles in the book deal with general dictionaries, the focus will be on those that do, but the author also considers relevant postulates from those articles dealing with dictionaries for languages for special purposes.

2. The semasiological and onomasiological presentation of semantic information

Wiegand initially draws a distinction between semasiological and onomasiological presentation in monolingual dictionaries. On a macrostructural level, semasiological presentation corresponds to a strict initial alphabetical listing and on a microstructural level, to an approach emphasizing the denotative meaning of a word. On a macrostructural level onomasiological presentation corresponds to a thematic listing and on a microstructural level, to an approach that emphasizes the semantic relations a word has with other lexical items in a specific language, e.g. synonymy, hyponymy, hyperonymy and antonymy. Wiegand points out that current general-purpose monolingual dictionaries focus too much on the semasiological presentation at the expense of the onomasiological presentation. He proposes a more integrated approach on a microstructural level which will see the semasiological and onomasiological presentation of semantic information in separate article positions within the semantic comment in a dictionary article.

3. The semasiological transfer of meaning

3.1 The term "definition"

The predominant view in metalexigraphy is that the denotative meaning of a lexical item in a dictionary is represented by a lexicographic definition. Defi-

inition types have been identified and qualitative criteria have been set as part of lexicosemantic theory. Yet the term "lexicographic definition" has gone largely unchallenged in metalexical circles as well as in editorial style manuals.

In several of the articles in this book Wiegand questions whether "lexicographic definition" is "an adequate metalexical term" (Wiegand 1994: 241). He attempts to offer terminology with which one can describe the different elements in the lexicographic transfer of meaning. Initially, he replaces "lexical and lexicographic definition" with "lexical and lexicographic paraphrase". This shift seems to be motivated by a lack of faith in the seamless appropriation of the philosophical term "definition" for use in lexicography. By avoiding "definition" several of the terminological potholes, such as the difference between a nominal and a real definition, can be avoided as well. On this point Wiegand (1977: 93) states the following in his 10 theses on lexicography:

Written from an extra-communicative point of view, the lexicographic paraphrases should not be regarded as lexicographic nominal definitions, but as rule formulations for the use of the lemma in conformity with semantic rules. Only in this way can it be adequately explained why and in which way lexical paraphrases, when being read, become texts-in-function, thus enabling the reader to infer the meaning of the lemma.

A lexical paraphrase represents the denotative meaning of a lexical item which is "regarded as the rules of reference and predication for the use of (the lexical item) in habitual texts for (that specific lexical item)" (Wiegand 1983: 148). The lexicographic paraphrase seeks to abbreviate these rules.

These early thoughts are clarified and systematised later in a seminal article "Elements of a Theory Towards a So-called Lexicographic Definition", which was published in German in 1994. In this article the "So-called Lexicographic Definition" is not viewed in isolation, but rather as part of an integrated microstructure within the paradigm of Wiegand's comprehensive textual theory for lexicographic texts, specifically as part of the integrate core of the semantic comment or subcomment (in the case of different senses). Therefore the terminology which is given preference is "an item giving the meaning paraphrase" (Wiegand 1994: 253).

3.2 The traditional approach to the semasiological presentation

This rejection of the term "definition" as not being an "adequate metalexical term" (Wiegand 1994: 241) underlines a more fundamental disagreement with its use in both meta- and practical lexicography. Numerous academic studies have highlighted the variety of "definition types" which can be employed in general monolingual dictionaries. In most of these studies,

preference is given to the so-called definition using *genus* and *differentiae* (see Gouws 1989: 115-116). These "true definitions" (Svensén 1993: 120), which Svensén further describes as being "intensional", expressing "a generic conceptual relationship whereby concepts are arranged in classes according to similarities and differences noted between them" (1993: 122), can be used for most of the syntactic classes (parts of speech) to which a lexical item can belong.

This definition type is firmly rooted in a traditional approach to semantics where componential analysis is the starting point in the construction of a definition. In the introduction to Wiegand's book, Wolski makes it clear that Wiegand does not regard this theory of meaning as entirely suitable for the needs of lexicography. These articles therefore reflect the development of an independent theory of meaning, which is in accordance with Wiegand's textual theory of lexicographic texts. An actional-semantic approach is taken to the expression of the denotative meaning of a lexical item. Denotative meaning is not abstracted from componential analysis, but, as was noted in paragraph 3.1, is "regarded as the rules of reference and predication for the use of (the lexical item) in habitual texts for (that specific l.i.)" (Wiegand 1983: 148).

3.3 The actional-semantic approach to the semasiological presentation

In his later articles Wiegand clarifies why a definition based on componential analysis and presented as a so-called definition using *genus* and *differentiae* can never adequately reflect the denotative meaning of a lexical item (see Wiegand 1994: 250-252). The item giving the meaning paraphrase needs to be based on more pragmatic principles and a different perspective on the constitution of meaning knowledge.

Within the actional-semantic approach Wiegand identifies types of meaning knowledge central to the lexicographer's efforts to represent semantic information in the dictionary. Wiegand (1994: 262) works mainly with the syntactic classes which contain the most words and in this context states:

In the case of lexicalized predicators (adjectives, nouns, verbs) the actional knowledge for the linguistic acting consists of

- (i) the non-encyclopaedic meaning knowledge (the fact that a predicator is habitually used for performing the partial act of predicating ...),
- (ii) the encyclopaedic, object-constituting meaning knowledge (the fact that a predicator is habitually used to refer to something defined by this very predicator or by the habitual use of other predicators — i.e. by language and its use),
- (iii) in some cases, by another kind of meaning knowledge (i.e. that a predicator is used to evaluate the reference object, to express an attitude, etc. ...)

... Meaning descriptions ... have to impart the actional knowledge necessary to infer from these meaning descriptions (i.e. from items giving the meaning paraphrase) this specific knowledge and consequently the rules.

The rules referred to in this statement are the rules of reference and predication in Wiegand's interpretation of denotative meaning, of which the item giving the meaning paraphrase should be an abbreviation.

How does one determine what should form part of such an item giving the meaning paraphrase? Componential analysis is replaced with a pragmatic approach which focuses on the needs of the target user of the dictionary. Whereas the definition, which is a result of componential analysis, is a static entity, the item containing the meaning paraphrase is a dynamic entity which can be different for different types of dictionaries, in each case being adapted to the needs of the target user.

This innovative view on the construction of the item containing the meaning paraphrase can lead to the compilation of far more user-friendly dictionaries. It comes to terms with the fact that a dictionary article is a "simulated question and answer dialogue" (Louw: 1998: 105) between the lexicographer and the dictionary user. Wiegand (1994: 254) reiterates this point by stating that, "producing a dictionary article ... is the formulation of a number of potential answers in a condensed and standardised form to anticipated types of search questions which are integrated in types of use situations". Earlier in the book he states that "only by anticipation can the lexicographer hope to meet the expectations the user has of the codified results of the lexicographer's written acts, the dictionary excerpts" (Wiegand 1976: 14). An item giving the meaning paraphrase should then best be constructed according to a list of "user questions" (Wiegand 1994: 265). These user questions must be in line with the lexicographic needs assessment done as part of the dictionary plan and should be based on detailed empirical research to determine the user perspective.

If an item is constructed within these parameters "one can be reasonably sure that (against a presupposed background of a language community which has reached a preliminary understanding on the use of the lemma-sign in habitual texts) a competent user can infer from a *correct* item giving the meaning paraphrase for a designative lemma-sign ... which he understands correctly, the actional knowledge which as a rule is necessary and sufficient to correctly make a habitual and correct reference with this lemma-sign" (Wiegand 1994: 265).

This shift in emphasis can be illustrated by some examples from Afrikaans monolingual dictionaries. If one employs, for example, componential analysis in order to determine the denotative meaning of the lexical item *olifant* (elephant) one would find an entry similar to the one in the HAT, but different to the one in the WAT.

It could be argued that user's needs have played a decisive role in the construction of the item containing the meaning paraphrase in both these articles.

o·li·fant (-e) 1 Groot, plantetende soogdiervan Afrika, *Loxodonta africana*, en Asië, *Elephas maximus*, met 'n lang, buigsame slurp, waaiervormige ore en waardevolle ivoortande: *So groot soos 'n olifant. Die olifant is die grootste lewende landdier.* 2 (by uitbr.) Groot, lomp ding.

olifant [o(:)'lifant] s.nw., olifante; olifantjie.

1 a Ook, geselstaal, *oolfant*. Enigeen v.d. baie groot, plantetende soogdiere v.d. geslachte *Loxodonta* en *Elephas* (fam. *Elephantidae*, orde *Proboscidea*), wat tot 7 ton kan weeg, ong. 2,30 m hoog word en 'n ouderdom van oor die 100 jaar kan bereik, met 'n neus en bolip wat tot 'n lang, buigbare slurp verleng is, en 'n dik, byna haarlose vel, 'n groot kop, en 'n goed ontwikkelde brein, veral *L. africana* wat in Afrika voorkom, met groot, waaiervormige ore en ivoortande by sowel die bul as die koei, en *E. maximus* wat in suidelike Asië, veral in Indië, voorkom, met kleiner ore en ivoortande slegs b.d. bul: *Die olifant is een van die oorlewende reuse wat vandag nog in Asië en Afrika aangetref word. Die mammoet, wat verwant is aan die Indiese olifant, het duisende jare gelede ... geleef* (A.C. Hoffman: *Diere*, 1942, 149). *Olifante (was) volop. Hulle is van die vroegste dae af meedoënloos gejag om hul tande wat gebruik is vir die vervaardiging van artikels soos biljartballe, klavierklawers en ornamente* (Burg., 24 Mei 1971, 6). *Olifante is kuddediere en as 'n jong olifantjie van sy trop vervreemd raak, soek hy soms die geselskap van ander diersoorte op* (Custos, Okt. 1972, 17).

Vgl. AFRIKAANSE OLIFANT (AFRIKAANS, uitdr.), INDIËSE OLIFANT (INDIËS, uitdr.).

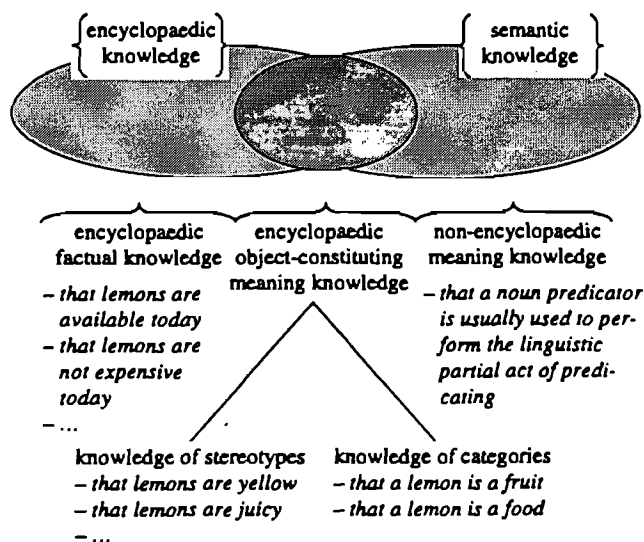
b Enigeen van verskeie uitgestorwe diere v.d. fam. *Elephantidae* of v.d. orde *Proboscidea* wat a.d. olifant (OLIFANT 1 a) herinner: *Uit opgrawings ... blyk dit dat hy (die mammoet) besonder groot was ... Ander uitgestorwe olifante ... is die mastodon, die stegedon, die platkopolifant en 'n soort met vier lang tande in die bokaak* (P.J. Schoeman: *Afr. vriende* 1, 1975, 35).

2 Iem. of iets wat baie groot, lomp of swaar is: *Ek (was) in die kleedkamer saam met hierdie olifant ... Baasspelerig en lui — alles op gesag van die ... boer wat sy ring aan daardie vet vinger gesteek het* (J.M. Gilfillan: *Frambose*, 1993, 13).

In the desk dictionary a somewhat cryptic, yet distinguishing paraphrase is given, while the paraphrase in the comprehensive dictionary is more detailed and specific. Both items provide more information than a definition constructed by means of a strict componential analysis and presented in the form of a definition using genus and differentiae. A probable question from metalexigraphical circles would be: is the additional information provided encyclopaedic, rather than semantic? In order to answer this question within the terms of reference of this article, one would have to return to Wiegand's theory of the kinds of meaning knowledge.

3.4 Semantic and encyclopaedic information

In Wiegand's semantic theory the boundaries between semantic and encyclopaedic knowledge are not as discrete as in traditional semantic theories. The relationship between the two types of knowledge are visually represented as follows in Wiegand (1994: 268):



The item containing the meaning paraphrase must not only reflect the "non-encyclopaedic meaning knowledge" but also the "encyclopaedic object-constituting knowledge". In the two articles above this is certainly the case.

In the article from the HAT, very little information is added to what would be gained from a strict componential analysis. A supplementary component is included in the item giving the meaning paraphrase, viz. the fact that the ivory tusks are valuable. Taxonomy is also given. This is a procedure which Wiegand opposes, preferring taxonomic details in a separate article position. One can agree that the target user group of the HAT probably does not need these details, as these users only require a very broad perspective on the object that is described. In general though, this item from the HAT succeeds in its goal by conveying the necessary object-constituting knowledge.

The needs and expectations of the WAT's target users differ greatly from those of the HAT's target user group. As a comprehensive dictionary, it has to provide enough information for the user to gain a detailed and specific perspective on the object described. Whereas the HAT would for example suffice with "very big" as a size description, the WAT gives a more detailed account of the size and weight of an elephant. The HAT only gives a cursory physical description of an elephant, enough (one could argue) to distinguish it from other "very big mammals", but the WAT paints a detailed picture not only of an

elephant's distinguishing features, but also of the differences between the two main species of elephant.

Both these dictionaries have transferred object-constituting knowledge by means of the item giving the meaning paraphrase. Yet judgements of adequacy in this regard depend solely on the context of the dictionary search which was conducted.

Neither dictionary stops here in their attempts to adequately describe the animal. The HAT proceeds to use own examples to enhance the meaning transfer. In the WAT's case, additional encyclopaedic information is presented by means of citations, which help to further improve the user's world knowledge.

3.5 Frames

In order to systematise this user-oriented approach to the transfer of denotative meaning in monolingual dictionaries, the concept of "frames" has been introduced by Wiegand (1989: 573) and expounded upon by Konerding (1993). This concept refers to a set of "thematic question groups" (Wiegand 1994: 272) which are set up for a specific type of lexical item. In the case of *natriumlamp*, such a set could constitute more or less the same as the set for tools proposed by Wiegand (1994: 272-273) and the one for musical instruments quoted by Smit (2000: 178-179).

Five broad groups of questions can be identified:

1. How can the (outer) form of the object be described? (Wiegand 1994: 272)
2. What are the "predicates characterising the hierarchically superordinate whole of which the object is an ingredient/part"? (Smit 2000: 179)
3. In which way (how) does the object (type) come into existence? (activity/production) (Wiegand 1994: 273)
4. What are the "predicates characterising activities in which the object functions/plays a role"? (Smit 2000: 179)
5. Which other objects is the object under consideration similar to and in which respects are they different? (Wiegand 1994: 273)

According to Smit (2000: 178-179), who discusses the groups of questions set up by Konerding (1993), question groups 1 and 2 fall under the subheading "constitutive relationships and features of the object", question group 3 under "phases of existence and distribution" and question groups 4 and 5, together with a separate single question on "other names for the object", under "the meaning of the object for people".

In the WAT's article for *natriumlamp* the object-constituting knowledge (as part of the denotative meaning) is imparted effectively by providing answers to some of the above questions.

natriumlamp [na:'tri:ɔemlamp / na:'tri:ɔem-lamp] s.nw. Gasontladingslamp wat gew. bestaan uit 'n u-vormige buis gevul met natriumdamp en 'n klein hoeveelheid neongas, waarvan die kleurkwaliteit swak, maar die ligopbrengs hoog is, en wat hfs. vir straatverligting gebruik word: *Daar word tans op groot skaal gebruik gemaak van ontladingsbuisse, bv. die natriumlamp, die kwikdamp-lamp* (Afr. Kernens., 1965, 864). *In ... kunstmatige dagliglampe word elektrisiteit ... deur ... onaktiewe gasse en dampe (gestuur) ... Ons het bv. ... natriumlampe wat geel lig uitstraal* (Le Roux – McLachlan: Alg. Nat. J.S., 1950, 268).

Its outer form and the "activities in which the object functions/plays a role" (Smit 2000: 179) are disclosed, but not the way the object is produced, as this is not considered to be truly relevant to the target user of the dictionary. Some idea is given as to "the hierarchically superordinate whole of which the object is an ingredient/part" (Smit 2000: 179) by means of the traditional first element of a "genus and differentiae definition". In this case, however, the procedure has some value, because the superordinate chosen is not hierarchically far removed from the lemma (as "lamp" would be for example). "Gasontladingslamp" is a direct superordinate and as such is a valuable point of departure to the users of the dictionary. By scientifically constructing such an item giving the meaning paraphrase, the WAT succeeds in conveying the necessary semantic information. Further encyclopaedic information about the object is available in the examples for those users who need or wish to know more than is required to identify the object.

3.6 Qualitative criteria

In Wiegand's opinion, the use of frames for construction of items containing the meaning paraphrase also seems to obviate the need to identify traditional qualitative criteria. This includes criteria meant to encourage "good defining practice" (Landau 1984: 132) such as "priority of essence", "substitutability", "brevity", "simplicity", etc. (see Landau 1984: 132-138 for a more comprehensive discussion of these criteria). Wiegand (1994: 235-237) is especially harsh on substitutability, questioning its "relevance" and usefulness as a criterion. He (Wiegand 1994: 271-272) further re-emphasises an earlier call for "a re-orientation away from the traditional concepts of definition and, therefore, away from such fruitless alleged problems which also appear in their wake, e.g. circularity, incompleteness, vagueness, and redundancy". On this point Wiegand (1994: 272) adds the following in summary:

These albeit difficult problems with respect to definition theory are of only marginal importance for lexicography as a scientific practice because every good lexicographer knows how to handle, for instance,

circularity. Neither do these tasks relate to any tasks in dictionary research that are still of importance today.

These problems should, if the correct procedure is followed while defining, not arise. This once again shows the value of a well-designed dictionary plan in which scientifically sound structural guidelines (drawn up with the target user's needs in mind) form an integral part.

4. The onomasiological transfer of meaning

Not all the semantic information inferable from the frames needs to be presented in the item containing the meaning paraphrase. Several of the questions pertain to the semantic relations existing between the lemma and other lexical items in the lexicon. This would include questions from question groups 2 and 5, as well as questions pertaining to "other names for the lemma". In these cases Wiegand prefers segregation. The semantic relations should be given by items in other article positions which are separated from the position of the item giving the meaning paraphrase by clear, unambiguous structural markers. Different article positions should preferably be allocated to the various items, reflecting the different semantic relations.

As was mentioned in paragraph 2, the semantic relations under discussion include hyponymy and hyperonymy, synonymy and relations of semantic opposition. Of these, synonymy is primarily emphasized in Wiegand's book, with three of the nine articles dealing almost exclusively with either synonymy in general monolingual dictionaries or synonymy in dictionaries for special purposes.

4.1 Synonymy in general monolingual dictionaries

The term "synonymy" does not escape redefinition. More so than "definition" perhaps, the term "synonymy" has been a bone of contention in metalexicographical circles. The common misconception of synonymy as referring to "two words meaning the same thing" has been severely criticised. Firstly, it is usually pointed out that synonymy is a relation between lexical items and not merely between words. Secondly, the inherent vagueness of the phrase "meaning the same thing" is criticised. It does not address the complexity of synonymy as a semantic relation. In fact, as Louw (1998: 176) points out, "there are few if any absolute synonyms in a language" (on this point see Louw and Nida 1988: 15). Contextual and other differences which exist between two lexical items with similar reference, have necessitated the distinction between absolute and partial synonymy. Rather than viewing these terms as polar opposites, it is useful to view synonymy as a scale with these two types of synonymy at various ends. One could argue that even though there are few

absolute synonyms in a given language, there are usually many that are close to them on the scale of synonymy and can therefore be classified as near- or near-absolute synonyms. In this class, the contextual differences are relatively slight and the interchangeability of synonyms would not cause substantial embarrassment to speakers, as for example with a standard language and vulgar pair. It is mainly the class of near-synonyms that concerns Wiegand.

In this context Wiegand (1983: 146) redefines lexical synonymy as occurring between two lexical items when "the rules of usage ... (are) similar to such an extent that the rules of reference and predication are the same in habitual texts". The lexicographer must represent this "tolerance relation" of substitutability effectively when presenting lexicographic synonyms.

4.2 On the so-called "synonym definition"

Even in his earlier articles, Wiegand criticizes the treatment of synonyms in general monolingual dictionaries. The presentation criticized is mainly the so-called "synonym definition". This type of definition has generally been accepted in metalexicographical circles, though with some reservations. These reservations have mainly targeted the inadequacy of lexicographic article-external cross-referencing between the "synonym definition" and corresponding target entry, as well as that between the article where the "full definition" appears with the synonym mentioned and the article with the synonym definition as target entry.

With regard to standard or smaller general monolingual dictionaries, the prevailing view is that "synonym definitions" are necessary. Landau (1984: 270) states that "synonym definitions are not necessarily a mark of lazy or inept editing; they may be the best solution to the problem of too little space". Sven-sén (1993: 119) concurs, but also explicates the major reservation as follows:

Using synonyms and near-synonyms as definitions saves space, and the method is entirely valid when the need for semantic precision is not too great, but one has to be on guard against synonyms that have multiple meanings. If a synonym is polysemous, it must not stand as a complete definition, but must be disambiguated ...

Unfortunately this is often not done. Consider, for example, the treatment of the near-synonyms *kyker* and *oog* in HAT.

Both these lemmas are polysemous and in one sense of each, they are used to refer to the same object. At *kyker* 2 a so-called synonym definition is given, i.e. "*oog*". It is not specified for which sense of *oog* this reference is valid. Only an experienced speaker of Afrikaans would know that the first sense of *oog* is the relevant one and if this speaker already knew this, there would be no need to consult the dictionary. The problem is compounded by two other factors. Firstly "*kyker*" is not even mentioned under *oog* 1 as a possible synonym. Sec-

only; a further "synonym definition" ("oogappel") is inserted into the same article position under *kyker* 2 and separated from "oog" by a comma. The conclusion to which a dictionary user would probably come is that these two are synonyms of *kyker* and of each other (since commas are usually used to separate so-called synonyms in the semantic comment). Yet "oogappel" (in its literal use) refers to the pupil of the eye whereas "oog" refers to the whole eye. They cannot be listed together without further semantic and contextual guidance. Furthermore, *oogappel* is itself a polysemous lemma and no attempt is made at *kyker* to disambiguate the reference. This confusing use of so-called synonyms and so-called synonym definitions is an obstacle to the effective transfer of semantic information in a single-volume general monolingual dictionary.

ky'ker (-s) 1 Iemand wat kyk, bekyk; toeskouer. *Baie kykers, maar min kopers.* 2 Oog, oogappel: *Blou kykers.* ∇ *Dit is ooglopend Abel en Ella se dogter: daar is Ella se gitswan kykers en haar digte, swan hare* (Etienne van Heerden). 3 Optiese instrument: (Ver)kyker, (toeneel)kyker. 4 Iemand wat na die televisie kyk: *Gereelde kykers.*

oog (oë; ogie) 1 Elk van die organe waarmee gesien word, waarmee mens of dier kyk; gesigsorgaan: 'n *Nooi met mooi bruin oë. Sommige insekte het saamgestelde oë. Die kleur, appel van die oog. Die haat, liefde in iemand se oë sien.* ∇ *Jou oë is nat van die trane van gister* (C.L. Leipoldt). *Blou oë waarin die hemel woon* (A.G. Visser). *So ver as die oog reik, so ver mens kan sien. Met jou oë iets soek, probeer om dit raak te sien. Met die bloe oog lees, sonder bril, vergrootglas.* 2 Iets wat soos 'n oog lyk of aan 'n oog laat dink: *Die oog van 'n naald. Hakies en ogies gebruik aan 'n kledingstuk. Die oog van 'n fontein, stroom, rivier, plek waar dit ontstaan, bv. ∇(U)it dieselfde oog (burrel tog nie) vars en bitter water (op nie)* (Jak. 3:11). *Oog van 'n aartappel, loot, waar dit uitloop.* 3 Uiterlik: *Hierdie tuin maak 'n (mooi) oog.* 4 (nie alg.; geselst.) Kroeg: *Heeldag in die oog deurbring.* 5 Middel-punt van 'n sikloon e.d.

The problem is even greater in comprehensive general monolingual dictionaries, where cross-references are not only article-external. Comprehensive dictionaries often comprise several volumes and cross-references are often also volume-external. The user is then expected to combine the correct pieces of information by making use of remote article-external and volume-external non-lemmatic addressing procedures, which are very difficult manoeuvres.

Furthermore, a comprehensive dictionary often takes a long time to complete, with any number of years passing between the publication of its first and last volumes. Cross-references are therefore often made to lemmas which will only appear in much later volumes. These "synonym definitions" are therefore not easily "disambiguated" by means of a reference to a specific sense of a lemma, because the article for that synonym has not been constructed yet. Comprehensive dictionaries are also only revised after all volumes have been

published, which means that the user will have to contend with imprecise "synonym definitions" for a long time.

As with "definition", Wiegand rejects both the term "synonym definition" and its application. He (Wiegand 1976: 25) seeks to redefine this item (which was traditionally referred to as a "gloss" at that time) in the very first article, referring to "dictionary excerpts in the form of lemma word" instead. According to Wiegand, this type of dictionary excerpt is not suitable for use in dictionaries, as a synonym cannot fulfil the same function as a lexicographic paraphrase given as the item containing the meaning paraphrase, i.e. an abbreviated rule formulation.

In the later articles this view is systematically supported by distinguishing, on the grounds of separate functions, separate article positions for items giving the meaning paraphrase on the one hand and for word synonyms on the other. With the textual theory for lexicographic texts as broad framework, word synonyms are to be presented as part of the integrated microstructure as follows:

Word synonyms ... belong into the SK (semantic commentary) yet not in the semasiological, but in the onomasiological part. Here, they are best listed under an explicit cross-reference by synonyms (e.g., "Sy" or "≈"), which can be understood as a semantic commentary symbol, functioning as a cross-reference ...; the synonyms should be printed in such a way that the font differs from the one used for the ME (meaning explanation) ... In each SK, the ME is therefore obligatory and word synonyms are listed if there exist any (Wiegand 1983: 150).

Wiegand (1983: 150) adds that this separation of semasiological and onomasiological information is important, because "word synonyms are not 'hidden' in the ME, instead, semantic knowledge of the lexical structure is made explicit". This is illustrated in the treatment of *onbewerk* and *onverwerk* from the WAT.

The WAT does employ so-called "synonym definitions", but with accurate specification of the sense to which the synonym applies, in this case "Onverwerk (ONVERWERK 1)" at *onbewerk* 2. The synonym definition is met by a "synonym mention" in the other article. This mention is close to Wiegand's proposal for the treatment of all synonyms in general monolingual dictionaries. It is positioned close to the item giving the meaning paraphrase and separated from it by means of a typographical structural marker, here a semi-colon. A second structural marker, "sin.", is added for more clarity and this is followed by the synonym, in this case *onbewerk*, which is presented in italics, whereas the item giving the meaning paraphrase is given in roman font. It is also separated by means of typographical structural markers from its superceding information categories. To complete the picture, some field labels in the microstructural treatment of the synonym, such as "(minder gebruikelik)" in this case, are also included in the synonym mention, in order to indicate the restrictions on use

onbework [ɔnbəvrək' / ɔnbəvrək' / 5:bəvrək'] b.n.w., onbeworkte.

1 Wat nie bewerk (BEWERK 3) is nie; soms sin. met *braak*: *Die plaas (was) nog onbework en een groot boskasie* (Landbw., 20 Okt. 1970, 25). *Oumansoutbos (hou) minder van bewerkte as van onbeworkte grond* (Landbw., 4 Apr. 1972, 58).

▪ Ook byw.: *Die mielies (is) ... onbework geplant* (Boerd. S.A., Nov. 1971, 26).

2 (minder gebruiklik) Onverwerk (ONVERWERK 1): *Onbeworkte bokhaar, velle, wol. Die handel in onbeworkte pelse het ... uitgebrei* (C.S. Grobbelaar in Afr. Kinderens. VIP, 1964, 287). *Onbeworkte mis word vir veel minder (as vloeibare mis) verkoop* (Landbw., 10 Nov. 1970, 63).

onverwerk [5:fərverk' / ɔnfərverk'] b.n.w., onverwerkte.

1 Wat geen verwerking ondergaan het nie, of wat dit slegs in 'n geringe mate ondergaan het; sin. *onbework* (minder gebruiklik); teenoor *verwerk*: *Onverwerkte grondstowwe, leer, materiaal, metaal, olyfolie, produkte, tabak, wol.* = *Daar kan ... gedink word aan die heffing van 'n klein uitvoerbelasting op die uitvoer van onverwerkte metallurgiese ertse* (Tegnikon, Mrt. 1966, 8). *Daar (was) verkope van onverwerkte Russiese stene buite die Sentrale Verkooporganisasie* (Beeld, 9 Mrt. 1994, Sakeb., 1).

▪ Ook byw.: *Goud onverwerk uitvoer.* = *Die riool (sal) onverwerk weggerig word totdat die ontwikkeling van die oord 'n ander metode regverdig* (Burg., 6 Sept. 1986, 3).

2 Waarvan die implikasies nog nie deurdink of die negatiewe impak nog nie geabsorbeer of geassimileer of draagliker gemaak is deur 'n verwerkingsproses nie; teenoor *verwerk*: *Die verkragter (het) probleme met homself... byvoorbeeld onverwerkte frustrasie en aggressie, seksuele probleme en 'n behoefte aan mag* (Beeld, 22 Mrt. 1994, Vroueb., 17). *Drie maande ... (na die miskraam) was ek weer swanger, maar jare daarna sou ek ... nog baie terneergedruk voel. Dit was onverwerkte verdriet* (Huisg., 14 Aug. 1986, 132).

▪ Ook byw.: *Hierdie stuk van sy lewe lê onverwerk in sy binneste ... Dertig jaar van wrok, verbittering ... om hom lê die wye see van die eensaamheid* (C.M. v.d. Heever: Simson⁸, 1940, 163).

3 Wat nog nie geprosesseer of aangepas is in 'n nuttige of bruikbare vorm nie; teenoor *verwerk*: *Statistiese inligting in onverwerkte vorm.* = *Die onverwerkte gegewens (moet)... eers verwerk word. Deur die toepassing van sekere beskrywende statistiese tegnieke kan ... die gegewens in 'n sinryker vorm gerangskik word* (N.J. Möller, vert.: Stat. Begr., 1968, 22).

▪ Ook byw.: *Die aansoekvorms van tienduisende Oos-Kapenaars wat ... om staatspensioene aansoek gedoen het, lê steeds onverwerk by staatsdienskantore* (Burg., 7 Sept. 1996, 1).

which exist for that specific synonym. A criticism of the WAT is that sense specification is not presented at this item, but only at the so-called synonym definition. The ideal would be a repetition of the full item giving the meaning paraphrase, with synonyms presented in separate article positions in the format displayed by the WAT, but with an accurate sense specification as employed in the WAT's "synonym definitions". This would also be the most effective way of curbing the problems created by synonym references between volumes.

Consequently, one can agree that, in keeping with the user perspective, Wiegand (1983: 150) is correct in remarking that "the 'semasiological principle of the item giving the meaning by means of word synonyms' ... should ... be substituted by an 'onomasiological principle of supplementing the lexicographic meaning explanation by word synonyms'".

This would benefit users of comprehensive dictionaries, who would probably also benefit most from a repetition of a full meaning paraphrase and two-way synonym references in well-demarcated article positions. This can be an effective answer, if the lexicographer, as part of the planning of the semantic component of an article (preferably by using frames), anticipates as many of the possible near-synonyms as possible within the dictionary's macrostructural scope. This is a demanding brief, but seems to be in keeping with Wiegand's (1976: 24) view of the lexicographer as "an empirical scientist". The corpus, combined with other lexicographic sources and the lexicographer's own linguistic intuition, should provide an adequate picture of the lemma's synonymous relations with other lexical items.

Whether the repetition of "meaning paraphrases" is a viable option for smaller dictionaries, is a different matter. In commercially published standard, desk or smaller paper dictionaries, the maximum level of textual condensation is pursued. In this regard "synonym definitions" have always been an important space-saving (and therefore money-saving) tool. It is hard to see Wiegand's vision of a separate slot for synonymy in this predatory milieu. Perhaps the best one can hope for is that a suitable "synonym definition" with a precise cross-reference will be employed and met by an equally precise mention of that lemma as a synonym in the article of the more frequent lexical item, which should also contain a full meaning paraphrase. This mention, in the form of an item, should be separate from the item giving the meaning paraphrase and should be introduced by a unique structural marker. An adequate distinction of the contextual differences between the synonyms should then be drawn by means of labels. The main concern is that the decision be based on the needs of the target users as revealed by a detailed empirical study.

4.3 Other semantic relations

One of the most challenging tasks of the lexicographer is to adequately underpin the semantic relations between a specific lexical item and others in the macrostructural scope of the dictionary. Corpus material is often inadequate and decisions of this nature usually rely on "the language intuition, language competence and therefore world knowledge as well as world conception of the lexicographers, taking into account the empirically established use of language" (Wiegand 1976: 36) and on secondary sources such as thesauri and antonym dictionaries. In order to make the lexicographer's task easier, it is therefore important that there be a scientific system of presenting semantic relations, which can be applied consistently throughout the dictionary.

In the articles selected for publication, Wiegand focuses on presenting such a system for the lexicographic treatment of synonymy, but he also provides useful guidance for the treatment of other semantic relations. In a one-volume semasiological dictionary (such as the HAT) Wiegand (1982: 133) finds the indication of synonymy sufficient, but for "a multi-volume integrated dictionary" (such as the WAT) he proposes the same treatment for hyponymy, antonymy, etc. as he did for synonymy, i.e. an item in an article position separate from the item giving the meaning paraphrase and well-demarcated by a unique structural marker.

Attempts have been made in the WAT to reflect relations of semantic opposition in this way. A separate article position is employed and this item is separated from the preceding definition or synonym by means of a semicolon. No further distinction is, however, made as to whether the item denotes antonymy or shows that the user is dealing with a complementary pairing. Both types of item are introduced by the structural marker "teenoor".

Unfortunately the WAT does not have a consistent system for dealing with hyponymy and hyperonymy which is in line with the one proposed by Wiegand. As pointed out in 3.5 under the discussion of *natriumlamp*, superordinates or hyperonyms are usually given as the points of departure in items giving the meaning paraphrase, but there are no separate article positions containing hyperonyms as independent items.

With regards to hyponymy there is a tendency not to present hyponyms or cohyponyms, because it is not, for example, seen to be in the interests of the target users to present "lemoen", "pomelo", etc. as cohyponyms at *nartjie*. Only in what the lexicographer believes to be exceptional or difficult cases where the two objects are, for example, closely related and therefore easily confusable, are cohyponyms given. This would then be presented near the end of an article and introduced by "Vgl." ("Cf."). There are two problems here. Firstly, this placement causes a separation of this type of item from the other items giving onomasiological information. In fact, it separates the hyponym or cohyponym from all other items in the semantic comment, thereby making the user's search path harder and less predictable. Secondly, the article position introduced by "Vgl." does not only contain cohyponyms. Partial synonyms and other confusable words (even though a separate article position for these does exist) find their way into the "Vgl."-item. The resulting loss of uniqueness is a further stumbling block on the user's inner search path.

In this regard, Wiegand has a valid point of view. The onomasiological content of these items is an integral part of the minimal frame for each lexical item, and the item itself should be a crucial part of an effective, integrated microstructure, specifically as part of the semantic comment presented in close proximity to the item giving the meaning paraphrase.

5. Conclusion

In this article I have attempted to clarify, apply and also bring together two major threads running through this selection of articles. The semasiological presentation of meaning has been seen to be located most appropriately in the item giving the meaning paraphrase, which should form part of the semantic comment in an integrated microstructure. In the semantic comment this item must be followed by clearly distinguished items giving onomasiological information, especially on synonymy.

This article has, however, merely begun to show the potential of Wiegand's theories (such as the frame theory) for the improvement of dictionaries' user-friendliness. Lexicographers working on dictionaries for specific languages need to explore this potential for the specific types of dictionaries they wish to compile. We are entering an exciting new lexicographic dispensation in South Africa and the lexicographers of both the new and the existing National Lexicographic Units would do well to take cognisance of Wiegand's unique and powerful theories when planning the secondary lexicographic processes of their respective dictionaries.

Bibliography

A. Dictionaries

- Louw, J.P. and E. Nida. 1988. *Greek-English Lexicon of the New Testament Based on Semantic Domains. Vol. 1.* New York: United Bible Societies.
- Odendal, F.F. 1994³. *Verklarende Handwoordeboek van die Afrikaanse Taal (HAT).* Doornfontein: Perskor.
- Van Schalkwyk, D.J. (Ed.-in-chief). 2000. *Woordeboek van die Afrikaanse Taal. (Vol. XI).* Stellenbosch: Bureau of the WAT.

B. Other sources

- Gouws, R.H. 1989. *Leksikografie.* Cape Town: Academica.
- Hartmann, R.R.K. (Ed.) 1984. *LEXeter '83 Proceedings.* Tübingen: Max Niemeyer.
- Hausmann, F.J., et al. (Eds.) 1989-1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie.* Berlin: Walter de Gruyter.
- Konerding, H.-P. 1993. *Frames und lexikalisches Bedeutungswissen. Untersuchungen zur linguistischen Grundlegung einer Frametheorie und zu ihrer Anwendung in der Lexikographie.* Tübingen: Max Niemeyer.
- Landau, S.I. 1984. *Dictionaries: The Art and Craft of Lexicography.* Cambridge: Cambridge University Press.

- Louw, P.A. 1998. *Kriteria vir 'n standaard vertalende woordeboek*. Unpublished M.A. thesis. Stellenbosch: University of Stellenbosch.
- Smit, M. 2000. Using Frames in Special-Field Lexicography: An Ethnomusicological Case Study. *Lexikos* 10: 173-188.
- Svensén, Bo. 1993. *Practical Lexicography*. Oxford: Oxford University Press.
- Wiegand, Herbert E. 1976. Synonymy and its Role in Monolingual Lexicography. Wiegand, Herbert E. 1999: 11-54.
- Wiegand, Herbert E. 1977. Thinking about Dictionaries: Current Problems. Wiegand, Herbert E. 1999: 55-94.
- Wiegand, Herbert E. 1982. On the Meaning Explanation of Sentence Adverbs in Monolingual Dictionaries. Wiegand, Herbert E. 1999: 113-138.
- Wiegand, Herbert E. 1983. Synonyms Appearing in Major Alphabetical Dictionaries of Contemporary German. Wiegand, Herbert E. 1999: 139-152.
- Wiegand, Herbert E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, Reinhard R.K. (Ed.) 1984: 13-30.
- Wiegand, Herbert E. 1989. Die lexikographische Definition im allgemeinen einsprachigen Wörterbuch. F.J. Hausmann et al. (Eds.) 1989-1991: 530-588.
- Wiegand, Herbert E. 1994. Elements of a Theory Toward a So-called Lexicographic Definition. Wiegand, Herbert E. 1999: 203-282.
- Wiegand, Herbert E. 1999. *Semantics and Lexicography. Selected Studies (1976-1996)*. Edited by Antje Immken and Werner Wolski. Tübingen: Max Niemeyer.
- Zgusta, L. 1971. *Manual of Lexicography*. The Hague: Mouton.

The Use of Examples in Polyfunctional Dictionaries

D.J. Prinsloo, *Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa and*

R.H. Gouws, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa*

Abstract: Dictionary compilation for the African languages in South Africa stands on the eve of the African Renaissance and prospective compilers of dictionaries for African languages are in need of a sound perspective and guidelines in respect of the compilation and choice of examples. The aim of this article is to analyse and evaluate some current approaches towards the handling of examples of usage as a data category in modern dictionaries and to suggest ways in which this information category can be improved by compiling, selecting and shaping examples to render optimal transfer of information and to enhance information retrieval. The emphasis will be on example phrases and sentences given in the dictionary article.

Keywords: LEXICOGRAPHY, DICTIONARY, METALEXICOGRAPHY, AFRICAN LANGUAGES, EXAMPLES, AUTHENTIC EXAMPLES, CONSTRUCTED EXAMPLES

Opsomming: Die gebruik van voorbeelde in polifunksionele woordeboeke.

Die samestelling van woordeboeke vir die Afrikatale in Suid-Afrika staan op die vooraand van die Afrika Renaissance en voornemende samestellers van woordeboeke vir die Afrikatale benodig 'n gesonde perspektief en riglyne ten opsigte van die samestelling en die keuse van voorbeelde. Die doel van hierdie artikel is om sommige van die jongste benaderings ten opsigte van die hantering van gebruiksvoorbeelde as 'n datakategorie in moderne woordeboeke te analiseer en te evalueer en om metodes aan die hand te doen waarvolgens hierdie inligtingskategorie verbeter kan word deur die samestelling, keuse en afronding van voorbeelde om optimale inligtingsoordrag te bewerkstellig en om inligtingsontsluiting te verbeter. Die klem sal val op voorbeeldfrases en -sinne wat in die woordeboekartikel aangebied word.

Slutelwoorde: LEKSIKOGRAFIE, WOORDEBOEK, METALEKSIKOGRAFIE, AFRIKATALE, VOORBEELDE, OUTENTIEKE VOORBEELDE, GEMAAKTE VOORBEELDE

Introduction

One of the developments in modern-day lexicography, as indicated in numerous publications, e.g. Hartmann (1989), Gouws (1989), Wiegand (1998), Van der Merwe-Fouché (1999), is the emphasis on an approach according to which the compilation of dictionaries is guided by the user perspective. A user-orientated dictionary should lead to enhanced information retrieval procedures.

Improving the quality and appropriateness of examples is one of the ways of enhancing the process of information retrieval. The inclusion of examples should benefit all target users of a dictionary but it is of special importance to the encoding user who needs maximum guidance within the physical limitations of a dictionary article. Illustrative examples play a vital role in dictionaries, and the dictionary conceptualisation plan of any new lexicographic project should make provision for a systematic presentation of this data type in the data distribution structure. This also applies to the planning of dictionary projects to be compiled by the National Lexicography Units (NLUs) of all the official languages of South Africa. These dictionaries should primarily be directed at the communicative needs of the members of the relevant speech communities and should endeavour to equip these target users with the necessary communicative skills to function successfully in a multilingual and multicultural environment. Consequently, the functional inclusion of examples, illustrating actual everyday language usage, is of prime importance. It adds to the quality and user-friendliness of a dictionary.

Examples can have different functions, including an explanatory function, in a dictionary article (cf. Rademeyer 1992), which may never be underestimated but has to be maximized and utilised to the utmost. This is of particular interest to the dictionaries compiled for the African languages. It has been stated repeatedly in publications such as Mbogho (1985) and Gouws (1990) that to date, dictionaries for African languages are the products of limited efforts. When planning the dictionaries to be compiled by the NLUs the lexicographers have to be aware of the problems of the past and they have to identify and avoid these pitfalls. In the case of examples, attention should focus on e.g. information regarding the purpose and characteristics of good examples, the question of authentic versus constructed examples, the decoding and encoding functions of examples and the needs of the target user in especially learner's dictionaries. One of the problems confronting the lexicographer is the fact that, also with regard to the presentation and treatment of examples, theoreticians offer different views and suggestions. They are biased towards one or other ideological method, often reflecting a one-sided view of the compilation or choice of examples. In this article an attempt will be made to create a more balanced perspective on the selection and presentation of examples and their distribution in the article structure. Guidelines for the use of examples will be formulated, especially for prospective compilers of dictionaries for African languages.

The dictionary plan, which has to be implemented when the dictionaries of the NLUs are compiled, should display a sound theoretical basis (cf. Gouws 2000b). The lexicographic process may include no haphazard procedures or an arbitrary and random choice and presentation of data categories. This also applies to the selection and presentation of illustrative material. The role of the dictionary compiler as "the deliberate controller and manipulator of illustrative

material" (Cowie 1989: 63) has to be reassessed to ensure the consistent application of a well-designed lexicographic system.

Examples and dictionary typology

No single dictionary can be everything to everyone. It is therefore of paramount importance that the planning of every dictionary has to be preceded by an in-depth analysis of the needs and reference skills of the intended target users in order to determine the typological criteria to be adhered to (cf. Hartmann 1989). Dictionaries represent a reliable form of communicative empowerment. The dictionaries compiled for a specific speech community should be a response to the specific communicative needs of that speech community. A speech community existing in a monolingual environment has different needs compared to a speech community which is part of a multilingual society.

Within multilingual South Africa, dictionaries compiled for the African languages should aim to meet the needs of the members of the relevant speech communities. Although it has to be seen as the ultimate ideal to compile a comprehensive monolingual dictionary for each one of the South African languages, such a dictionary may not be regarded as a short-term target of the NLUs. Given the communicative needs of the members of the different South African speech communities, the compilation of more restricted dictionaries, e.g. learner-orientated bilingual or descriptive dictionaries or school dictionaries should be regarded as one of the primary objectives of the NLUs. In the past dictionaries of these typological categories have often neglected illustrative examples as a functional data category.

The comprehensive lexicographic process in South Africa (cf. Gouws 2000) has to include the formulation of a dictionary plan to be applied within each one of the NLUs. The application of the dictionary plan should result in the publication of a first as well as subsequent dictionaries. More specifically, the dictionary plan should not only aim at the publication of the first dictionary but also at the development of a range of dictionaries for each of the NLUs. Such an approach will have definite implications for the typological choice and all the structural components of the target dictionaries. This includes the macro- and microstructures as well as the access and mediostructure and, very important, the data distribution structure. The development of the South African lexicographic process has to be regarded as an evolutionary process. This has direct implications for the structure of the dictionaries to be compiled during the first phase of the process because these dictionaries should constitute the basis for the dictionaries of the next phase.

The presentation and treatment of examples should be seen within the context of the more comprehensive lexicographic process where both the dictionary typology and the choice of the type of microstructure play a decisive role. A decision by a specific National Lexicography Unit to start with a mono-

lingual or a bilingual dictionary of a restricted nature, e.g. a school dictionary or a desk dictionary, does not imply the omission of illustrative examples. Instead, it implies a much stricter set of selection criteria for the inclusion of entries representing this data category. Contrary to the system adhered to in many older dictionaries, it should be emphasised that the treatment of entries in dictionaries of a restricted nature may not leave the intended target user with the feeling that the dictionary has not assisted him/her in achieving the purpose of his/her dictionary consultation effort. In the past, school dictionaries have too often been the result of an editorial cut and paste process. A more comprehensive dictionary had been used as a point of departure and the school dictionary was often compiled by deleting macrostructural and microstructural entries in an arbitrary way. Illustrative examples represent one of the data categories which often did not survive to feature in a school dictionary. Such an approach totally eschews the needs of the target users of the dictionary.

This paper presupposes that illustrative examples represent a functional data category in any general translation or descriptive dictionary and that dictionary typology, within this broad category of general descriptive and translation dictionaries, should not detract from the value of examples. In such dictionaries, the comment on semantics should contain a search zone allocated to this data type. Any new dictionary project plan should make provision for an evolutionary lexicographic process which allows lexicographers to use the first dictionary as a basis for the following ones. This implies that each data category presented in the first dictionary, should be dealt with in such a way that it could be expanded on in a second, more comprehensive dictionary. The responsibility is on the lexicographer to ensure that each data category in a restricted dictionary is presented in a scientific and theoretically sound and adequate way. The typological nature of a dictionary may never be an excuse for the presentation and treatment of any given data category in a manner not motivated by sound theoretical principles.

It should thus be emphasized that dictionaries should display a structure aimed at the successful transfer of those data categories needed by the target users. This is of particular relevance to the dictionaries compiled for the African languages in the newly established NLUs. The specific nature of the microstructure and the article structure is of extreme importance for the inclusion, positioning and treatment of examples. Having decided on a specific microstructure and an article structure, the lexicographer should apply and adhere to these decisions in a consistent way.

Moving from an unintegrated to an integrated and eventually to a semi-integrated microstructure

Dictionaries can display different types of microstructures and the compilation of any dictionary should be in accordance with the criteria of the relevant type. Wiegand (1996b) makes provision for, among others, three major types of

microstructures: integrated, unintegrated and semi-integrated microstructures. For both monolingual descriptive and translation dictionaries, a semi-integrated microstructure could be seen as an ideal. However, this is a type of microstructure best suited for a more comprehensive type of dictionary and not for desk and school dictionaries. Furthermore, it presupposes quite sophisticated dictionary-using skills. In the remainder of this paper the emphasis will be on the other two types of microstructures. The discussion of a semi-integrated microstructure could come to the fore again in a model for a more comprehensive dictionary.

The use of an unintegrated microstructure implies that the translation equivalents and the cotext entries do not appear in the same text block. Positioning the translation equivalents and the cotext entries in different text blocks does not necessarily imply a lack of a coordination between these entries or the absence of an addressing procedure between cotext entry and translation equivalent. In a dictionary with short articles a well-devised unintegrated microstructure could serve the user well. However, the presentation of two separate search zones should not lead to an arbitrary ordering of entries on an intra-search zone level (cf. Gouws 2000a). This is unfortunately the case in many of the existing South African bilingual dictionaries due to an unpredictable and inconsistent selection and presentation of examples and other cotext entries.

The use of an unintegrated microstructure in many South African dictionaries often tends towards uncoordinated text blocks. The situation is often aggravated by the lack of a microarchitecture and consequently the absence of a clear indication of the boundaries of the different search zones. An unintegrated microstructure compels the lexicographer to use structural markers to ensure the identification of the different search zones as well as the necessary inter-search zone coordination of translation equivalents and cotext entries. The absence of such structural markers can lead to a special form of internal textual condensation in an article which requires more sophisticated dictionary using skills than the lexicographer should expect from the target user of the dictionary. Although the examples, given as cotext entries, can be coordinated with the relevant member of the translation equivalent paradigm if a relation of addressing equivalence prevails (cf. Gouws 2000b), this nonlemmatic addressing procedure illustrates the occurrence of nonspecific distant addressing. The nonspecific and nondirect link between the cotext entry and its address makes it more difficult to achieve successful dictionary consultation. Compare the article of the lemma sign *stryd* in *Groot Woordeboek/Major Dictionary*:

stryd, fight, strife, struggle, contest, conflict, combat, action, war; *die* ~ *AANBIND met*, join issue with; *die* ~ *BESLIS*, win the battle; *DIE STRYD* (*pol., hist.*), the Struggle; *die* ~ *om (die) BESTAAN*, the struggle for existence, *die* ~ *GEWONNE gee*, admit defeat; *IN* ~ *met*, in

conflict with, contrary to; ... *die goeie* ~ *STRY*, fight the good fight; ...
ten ~ *e TREK teen*, go to war against; ...

Contrary to this example, the functional application of an unintegrated microstructure presents the illustrative examples in a text block separated from either the text block containing the translation equivalent paradigm or the different subcomments on semantics given for the various polysemous senses of the lemma sign. The success of such a microstructural approach is the clear and unambiguous coordination between these text blocks so that each illustrative example entry has a definite and clearly indicated address in the preceding text block.

An integrated microstructure contains both the translation equivalents and the cotext entries in the same text block with a cotext entry following each translation equivalent to assist the encoding user in his/her attempt to master the correct usage. The microarchitecture of an article displaying an integrated microstructure can confuse users if the lemma sign represents a polysemous lexical item and translation equivalents have to be given for each one of the polysemous senses of the lexical item represented by the lemma sign. Users then may have to struggle through a lengthy article to reach the needed translation equivalent. One of the major advantages of an integrated microstructure is the coordination between translation equivalent and cotext entry, as a result of specific and direct, i.e. nondistant, nonlemmatic addressing. Compare the article of the lemma sign *maak* in *Tweetalige Aanleerderswoordeboek/Bilingual Learner's Dictionary*:

maak 1. make [a] *My mother can make clothes. My ma kan klere maak ...* 2. go *Ducks go "quack". Eende maak "kwaak". ...* 6. put *She tried to put the baby to sleep. Sy het die baba aan die slaap probeer maak.*

If applied well and consistently, both integrated and unintegrated microstructures assist the user to access the needed data categories presented in the comment on semantics of a bilingual dictionary. However, both these types of microstructures can also impede the information retrieval process if applied inconsistently or in an article not suited for the specific type of microstructure. Wiegand (1996b) has developed an additional microstructure, which includes the advantages and excludes the disadvantages of both the other types. This type of microstructure is called a semi-integrated microstructure. Articles with a semi-integrated microstructure display a characteristic microarchitecture with differentiated text blocks in the comment on semantics, to allow for both an integrated and an unintegrated lexicographic treatment. For reasons mentioned earlier in this section, the semi-integrated microstructure will not be discussed in more detail in this paper.

The illustrative examples in a bilingual dictionary can be regarded as the data category of which the presentation is most influenced by the specific

microstructural type. The plan of any new dictionary project should therefore also pay thorough attention to the specific type of microstructure most suitable to respond to the real needs and reference skills of real target users.

The purpose and properties of good examples

Just as any dictionary needs to have a genuine purpose (cf. Wiegand 1998), each and every data category included in a dictionary article should adhere to the realisation of the purpose of the dictionary. This can only be done if the lexicographer has a definite and clear view of the function of each data category. The inclusion, presentation, nature and extent of each data category should be governed by a set of criteria which places that data category within the broader frame of the purpose of the dictionary. Some data categories may have a monofunctional occurrence whereas others may be multifunctional. Although illustrative examples constitute a data category with a whole range of functions, one of the basic aims of the lexicographer by including and presenting examples should be to guide the user on a variety of characteristic features of the lexical item represented by the lemma sign which functions as guiding element of the specific article. Examples play an important role in guiding the user to *know* the word. Laufer (1992: 71) formulates this as follows:

Knowing a word would ideally imply familiarity with all its properties ... When a person "knows" a word, he/she knows the following: the word's pronunciation, its spelling, its morphological components, if any, the words that are morphologically related to it, the word's syntactic behaviour in a sentence, the full range of the word's meaning, the appropriate situations for using the word, its collocational restrictions, its distribution and the relation between the word and other words within a lexical set ... The foreign language learner knows a much smaller number of words ... In many cases word knowledge is only partial, i.e. the learner may have mastered some of the word's properties but not the others.

Research over a broad spectrum (cf. Cowie 1989, Gouws 1989, 2000, Rademeyer 1992, Rundell 1998, and Atkins, Rundell and Gouws 1998) reveals some of the main characteristics of examples. According to this research, examples should help to

- disambiguate senses,
- distinguish one meaning from another,
- clarify an abstract definition,
- supplement the information in a definition,
- show or indicate the selectional range,

- place the word in context,
- place the word in context,
- specify the semantic range,
- indicate the collocational behaviour, including typical collocations,
- illustrate the grammatical patterns,
- specify the word order,
- give pragmatic uses,
- note stylistic features,
- indicate appropriate registers,
- reflect the word history,
- be accurate, especially those quoting measurements, technical data, etc., and
- stimulate the user to capture the features or characteristics of the word in question and use the examples as a model to create examples of his/her own.

Characteristics of bad examples

There are numerous pitfalls in the way of constructing good examples, such as register mismatch, examples that are confusing, distracting, contain irrelevant detail, etc. Atkins, Rundell and Weiner (1997: slide 9c25) list the following typical problems:

- natural, typical ... but completely pointless (e.g. Sicilian: *a quarrel between two Sicilians*)
- includes distracting or irrelevant detail
- seriously atypical (e.g. by proxy: *You can create an international incident by proxy*)
- highly context-dependent (e.g. gravitate: *He gravitated, naturally, to Newmarket*)
- register mismatch (e.g. latter: *We have to decorate the kitchen and the hall — I'd rather do the latter (room) first*)
- confusing (e.g. black vb: *They blacked all coal from mines that had continued working during the strike*)
- or just missing when you need it ...

The incorrect handling or unnecessary introduction of extralinguistic factors such as race, sex, politics, culture, etc. also render bad examples, e.g. reference to female subjects, companies or the morally unaccepted. Compare Atkins, Rundell and Weiner (1997: slide 9c26) in this regard:

nag His wife nags (at him) all day. | She kept nagging (her husband) for a new car. | She nagged her son into leaving the house ... [LDOCE1]

crap [Company name]'s is crap — let's go to Burger King! [Longman Language Activator — early draft]

just the job Thanks for that screw — it was just the job! [LDOCE1]

It is also not advisable to refer to living persons or current events. Such examples enhance the relevance of the dictionary and are interesting to the user but easily become outdated, consequently the user might judge the whole dictionary as dated.

Corpus examples versus constructed examples: Introduction

The past few decades saw the rise of huge computerized corpora which brought a new dimension to dictionary compilation. The traditional method according to which examples were "made up" became increasingly challenged by the so-called "authentic" examples (unmodified utterances). The debate intensified since 1987 when Cobuild started using authentic examples in dictionaries. It unfortunately became an ideological struggle resulting in loss of perspective where supporters of a strategy only emphasized the virtues of their strategy and highlighting the shortcomings of the alternative method rather than taking the best of both alternatives. A more balanced perspective will be attempted in the following paragraphs.

In favour of corpus (authentic) examples and biased against constructed examples

Arguments in favour of authentic examples reflecting negatively on constructed examples are summarized in terms of Cowie (1978: 129), Fox (1987), Sinclair et al (1987, as quoted by Cowie 1989: 58), Laufer (1992), and Potter (1998: 357-362).

Authentic examples

- are almost always superior to constructed examples,
- are grammatically correct,
- are situationally appropriate,
- support the explanations,
- give accurate collocations,
- illustrate usage,
- provide a reliable guide for speaking and writing in the English of today,
- have actually occurred in the language,
- represent the language as it is actually spoken and written,

- show all the features of normal discourse, such as the use of pronouns rather than nouns, of linking words such as *and* and *therefore*, etc.,
- provide valid and accurate information on which learners can rely,
- include one or more useful collocations and, where appropriate, a range of grammatical patterns, and
- give the learner the guarantee that a piece of language does occur.

Constructed examples

- are often isolated, self-contained sentences because lexicographers tend to produce sentences with too much information in them,
- are sometimes odd and not very likely to occur in a communicative act,
- are artificial and may not reveal the most typical usage of a word,
- are really part of the explanations,
- have no independent authority or reason for their existence, and in many cases are used to clarify the explanations,
- give no reliable guide to composition in English and would be very misleading if applied to that task,
- do not say "This is how the word is used" but rather "This will help you to understand the sense", and
- could be hazardous without consulting a corpus.

Finally it is stated that

- usage cannot be invented, it can only be recorded,
- much of the information given in constructed or corpus-aided examples is reliable neither about the contexts in which a word or phrase is typically used, nor about the words that typically occur with it, and
- in some cases it can actually be misleading.

Potter (1998: 359-362) gives a detailed discussion on how constructed examples such as the one for the entry *bloom*, "I was just admiring the blooms in your garden", could be hazardous without consulting a corpus.

It would therefore have been ridiculous to have studied real language in order to find out the facts of the language, and then to have abandoned this and concocted fake examples for the dictionary. (Fox 1987: 138)

In favour of constructed examples and against corpus examples

Hornby (as quoted by Cowie 1978: 58), Laufer (1992), Minaeva (1992) and others argue that learners are better served by examples that are to a greater or lesser extent invented by lexicographers rather than by examples selected from a corpus. According to them constructed examples

- are as useful or more useful to learners than those taken directly from a corpus with little or no modification,
- demonstrate the linguistic points the lexicographer wishes to convey, without any distraction or added difficulty which may be introduced by examples taken directly from real texts,
- are more helpful in the comprehension of new words than authentic examples,
- can be made to include detail, whether syntactic or lexical, which throws light on the meaning or use of the entry word,
- can be judiciously shaped in the interest of the learner,
- can include the significant syntactic detail, and
- are more useful because they seem to be less dependent on the learner's general lexical knowledge than authentic examples.

Authentic examples

- do not, even when used extensively, guarantee that the various needs which dictionary examples have to meet, will automatically be fulfilled,
- place on the compiler the additional burden of selecting (or even editing) examples so that meaning, syntax and style are effectively illustrated, and
- are overloaded with irrelevant extralinguistic information.

Perspective on corpus examples versus authentic examples

From the arguments in favour of or against authentic examples and constructed examples presented above, it is clear that both sides claim superiority. It is also unfortunate that most lexicographers defend only one alternative by focusing only on its virtues whilst ignoring its deficiencies, rather than trying to capitalize on the virtues of both methods. Such an ideological struggle does not serve the interests of lexicography. The two extremes are manifested or summarized in the following quotations:

There is rarely any need to consider thinking up examples. There is sufficient evidence of how the word has been used: we can see what its collocates are; we can check whether the subject or the object of a verb is more typically a person or a thing ... (Fox 1987: 147)

versus:

Lexicographer's examples are more helpful in comprehension of new words than the authentic ones. In production of the new word, lexicographer's examples are also more helpful, ... the usefulness of constructed

examples seems to be less dependent on the learner's general lexical knowledge than the usefulness of the authentic examples. (Laufer 1992: 75)

One can also assume that in many cases lexicographers are not "free" to choose between authentic and constructed examples but have to follow "company policy". Furthermore, both sides claim that their viewpoints are supported by user feedback and experiments conducted on groups of users.

As a step towards a clearer perspective, general claims such as the superiority of one method over the other, authority, natural versus unnatural, more versus less thought-provoking, typical versus not typical, situationally appropriate versus inappropriate, user preference of one over the other, etc. should be avoided. It is risky to claim that the positive objectives listed here can only be obtained by using authentic examples. It will be more sensible to adopt a policy of taking the best from both options and avoiding the weak points of both at the same time.

To deal with the issue of natural versus unnatural: It is risky to generalize in stating that corpus examples are natural and constructed one's are not.

Lexicographers who are educated native speakers of the language are bound to have correct intuitions about their mother tongue, about the grammaticality of the word, its typical use and its typical environment. These intuitions are not necessarily less correct than the intuitions of those language users who are represented in the corpus and are therefore not less reliable. (Laufer 1992: 72)

Editors are constantly constructing examples to meet the learner's productive needs. Such examples ... often achieve their aim precisely because of some "unauthentic", pedagogically contrived feature of wording or typography ... Whatever the merits of authentic specimens of usage, it still has to be convincingly shown that they can take the place of invented examples for several productive purposes. (Cowie 1989: 58)

Thus it has to be concluded that a debate on this level has no gain. If some constructed examples are less "natural" than authentic ones, the problem does not lie in the inability of the lexicographer as a mother-tongue speaker to produce sentences as well-formed or natural as those uttered or written by other speakers of the language as taken from the corpus, but unnatural because too much information is stuffed into a single sentence, rendering it unlikely to occur in a natural conversation.

Fox (1987: 141-144) formulates this as follows:

Isolated examples carry too much meaning content for one sentence. ... *"The teacher used to cane me when I behaved badly"* ... Sentences occurring in

a natural context are not neat little isolated wholes. They carry a lot of loose ends — they follow on from what has been said and they lead into what will be said ... Language is not a series of isolated sentences ...

And Cowie (1989: 59) says the following:

Compilers who invent their examples may, of course, fall into the opposite trap of sacrificing linguistic naturalness to explanatory fullness.

On the other hand, Minaeva (1992: 78-79) points out that corpus examples, although "very natural", can carry an abundance of irrelevant sociolinguistic information such as names of people and places:

To use these dictionaries to the best advantage, one should possess a considerable amount of background knowledge because illustrative word-combinations and sentences abound in sociolinguistic information. The user should be an expert in science, literature and music ("Einstein was a mathematical *genius*." OALD; "I *myself* feel that Muriel Spark is very underrated." COBUILD; ...), know geography and history ("They've got a small farm in Devon." ... COBUILD ...); ... They are overloaded with irrelevant extralinguistic information ... irrelevant sociolinguistic information ...

Compare also the comments of Rundell (1998: 335) in this regard:

The risk here, illustrated rather too often in COBUILD1 but only very occasionally in COBUILD2, is that wholly authentic examples can sometimes show mystifyingly irretrievable contexts (for example in COBUILD1's example at *gravitate*: *He gravitated, naturally, to Newmarket*); atypical uses ... or too much irrelevant and — to the learner — distracting material ... So there is still a place for the more "pedagogical" example, typically now produced by modifying an actually occurring sentence, which allows the lexicographer to focus on specific linguistic points without baffling the user.

A second pitfall related to naturalness is to formulate constructed examples in one tense, mood or actuality, such as in the present tense positive, whilst the specific word more frequently occurs in, say, the past tense negative. The same holds true for constructing active examples, whilst such words are more frequently used passively.

There are certain clear limitations to the lexicographer's intuition. Without the use of a corpus, the lexicographer has no chance of finding the hundreds of senses and subsenses of a word like *run* on intuition. Furthermore, Fox (1987: 146) states that an experiment on intuition in respect of collocates revealed that

on intuition the lexicographer is inclined to think in terms of "semantic sets" rather than of "words which are actually likely to occur in the near vicinity of each other".

When we sit and intuit how words are used, we are likely to get it wrong. We also know that as soon as we start playing around with examples, making them more "accessible" or more "regular", we are liable to take the life out of them, or worse, mislead the user of the dictionary. (Fox 1987: 148)

The first step towards a clearer perspective on the issue of authentic versus constructed examples will be to view them as a *continuum* with end points *authentic (taken from the corpus)* as the one extreme and *constructed examples* as the other extreme with certain discreet *categories* in between, such as *edited corpus examples, corpus-based examples, and constructed examples taking corpora into consideration*. Compare the following oversimplified table illustrating this continuum:

Extreme	Intermediate categories			Extreme
Authentic (corpus examples) taken directly from a corpus without editorial modification	Slightly edited/modified corpus examples	Heavily edited/modified corpus examples	Partially invented, based on a corpus	Constructed examples

Humble (1998: 593) rightfully concludes:

The bone of contention came to be considered almost as an *ideological problem* and not as a practical one for which the solution *may lie somewhere in the middle*: each kind of example was considered intrinsically better or worse (emphasis added). (Humble 1998: 593).

Although lexicographers in principle opt for one of the extremes, there are certain factors which necessitate the selection of examples from the intermediate stages. Humble (1998: 593) considers factors such as (a) encoding versus decoding, (b) the target user's skills e.g. "beginners", "intermediate" or "advanced" learners, and (c) frequency of use of words.

Despite a certain amount of research into the issue ... the jury is still out on the relative merits of corpus-based and lexicographer-produced examples. But it is really no longer relevant to characterize the argument as concerning a simple choice between the authentic and the invented. All reputable monolingual learners' dictionaries now base every aspect

of their text on corpus data, so the differences now lie in the degree to which corpus material is "processed" on its way into the examples. (Rundell 1998: 334)

Encoding versus decoding and the learner's profile

According to some traditional points of view (cf. Kromann, Riiber and Rosbach 1984a, 1984b), the ideal situation would necessitate the compilation of at least four and possibly eight bilingual dictionaries for any given language pair. Certain applications of the active-passive principle imply that one has to distinguish four different functions for each member of any given language pair and that a separate dictionary should be compiled to comply with each one of these functions. This implies the compilation of dictionaries for text reception and text production for mother-tongue and non-mother-tongue speakers of both languages involved (cf. Hausmann 1977, 1986). The specific function of a dictionary should then determine the nature and extent of the illustrative examples to be used in that dictionary. However, even for the most sophisticated languages the idea of eight or even four dictionaries per language pair has to be regarded as totally impractical. From a user-perspective, such an idea is also not attainable. The current approach in dictionary research, which favours a user-driven lexicography, is also in opposition to such an impractical point of view.

There is quite a difference between dictionaries with a passive and those with an active function. Wiegand (1996a: XV) emphasises the fact that the formulation of a theory of bilingual lexicography may not lead to a situation where theory is isolated from lexicographic practice. Consequently Wiegand (1996b: 2) indicates quite clearly that the active-passive principle should not be regarded as a principle determining the typological nature of a dictionary. Within a theory of bilingual lexicographic texts and with a distinct focus on the needs of the target users of any given dictionary, it should rather be used to differentiate between various functions prevailing in one bilingual dictionary and even in one article of a given dictionary. Wiegand argues convincingly against the need to compile four (or even eight) bilingual dictionaries for a given language pair. According to him, only one general translation dictionary is sufficient for a given language pair if the data distribution structure and presentation in such a dictionary make provision for the different needs of different users by adopting a polyfunctional approach.

The dictionary conceptualisation plan of any bilingual dictionary should clearly indicate whether it is a monofunctional or a polyfunctional product and the lexicographer should know well in advance what the profile of the intended target user is and whether the search zone in which the illustrative examples are accommodated should focus on an encoding or a decoding function or on both these functions. The importance of examples for both the

encoding and the decoding function may never be underestimated. Fox (1987: 138-139) says:

It is important that we realize that learners' needs for encoding are at least as urgent as their needs for decoding. More so, perhaps, in these days when such strong emphasis is placed on communication ... We must ensure that the information we give students will genuinely help them to produce language of their own which is as near as possible to that produced by native speakers. In the past, dictionaries have been seen more as aids to decoding than aids of encoding. There is no reason why they should not succeed in doing both ...

Dictionaries have traditionally concentrated more on the grammatical, partly perhaps because they have been seen as tools for interpreting language rather than for creating it.

The use of examples for an encoding and/or decoding function is not only relevant to learner's dictionaries or bilingual dictionaries. Each and every lexicographer who decides to include illustrative examples in a dictionary should be aware that the entries given in the presentation of this data category, have to be functional text segments of the specific dictionary in which they appear. In a decoding function (cf. Cowie 1989: 57), examples can help to clarify individual meanings and can help the user to distinguish between related meanings. On an encoding level they help the user to select the correct grammatical pattern(s) for a given word or sense, to form acceptable collocations and to produce language according to stylistic norms typical of native speakers of the given language.

Although extremely important, the mono- versus polyfunctional nature of a dictionary may not be regarded as the only determining factor with regard to the choice of examples. Besides the function of the dictionary, the user's level of proficiency as well as the lemma in question should not only determine the choice of examples, but it should also influence the lexicographer's choice between constructed versus authentic examples. The dictionary conceptualisation plan should therefore provide for a data distribution structure which allows a differentiated approach with regard to the use of illustrative examples.

The treatment of words referring to very specific concrete objects, e.g. *radio* and *television*, do not benefit greatly from exemplification. Fox (1987: 137) states that examples in articles headed by lemma signs representing this kind of words teach the user nothing more about the actual object, but instead give information on how the word is likely to be used, e.g. *turn on the television* and *watch television*.

The lexicographer has to realise that the functional use of examples is a choice which will lead to entries accessible to the user. They should not be too difficult to understand, although, according to Humble (1998: 594), learners

profit most from an input which is slightly beyond their capacities. The examples must however be comprehensible and reproducible. The appropriate context, the typical grammatical structure as well as the pragmatic context in which a word occurs should form part of the information transfer achieved by the correct choice of examples. Whether the user is a learner or a mother-tongue speaker, whether the dictionary has a monofunctional or a polyfunctional approach, the choice of each example has to be done in such a way that it responds to the needs of at least one section of the target users of the dictionary.

If the dictionary is aimed at learners, Humble (1998: 596) stresses the special qualities of examples:

They should be given examples resembling the sentences they aim to produce. Words can combine to make sentences according to the *open-choice principle* or to the *idiom principle*. Frequent words have a tendency to form idioms often, but can be used on their own as well, in an *open-choice* setting, even if this modality is not always the vernacular one and examples of these congruent uses are hard to find in a corpus. Constructed examples, however, if congruent, use the words in their core meaning, and are especially useful for beginning and intermediate learners.

The data distribution structure of a dictionary (cf. Bergenholtz, Tarp and Wiegand 1999) should also determine the extent of the use of examples and the consequent balancing of data categories in the dictionary article. The extended use of examples at the cost of the inclusion or treatment of other data categories needs to be discussed at an early stage of the formulation of the dictionary conceptualisation plan. Although their functional value may never be underestimated, illustrative examples should always be regarded as one of the data categories which assist the user towards a better understanding and use of the treated language.

Decisions regarding the use of examples may not be taken in an arbitrary way. The dictionary conceptualisation plan should give a meticulous exposition of the function, choice, nature and extent of examples. Lexicographers should then adhere to these criteria.

References

- Atkins, B.T. Sue, Michael Rundell and Edmund Weiner. 1997. *Salex'97. A Training Course in the Compilation of Monolingual Dictionaries*. Unpublished course material of a tutorial held at the Dictionary Unit for South African English, Rhodes University, Grahamstown, 15-26 September 1997.

- Atkins, B.T. Sue, Michael Rundell and Rufus H. Gouws. 1998. *Afrilex-Salex'98. A Training Course in the Compilation of Bilingual Dictionaries*. Unpublished course material of a tutorial held at the University of Pretoria, Pretoria, 7-18 September 1998.
- Bergenholtz, H., S. Tarp and H.E. Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research: 1762-1832*. Berlin: De Gruyter.
- Cowie, A.P. 1978. The Place of Illustrative Material and Collocations in the Design of a Learner's Dictionary. Strevens, P. (Ed.). 1978. *In Honour of A.S. Hornby: 127-139*. Oxford: Oxford University Press.
- Cowie, A.P. 1989. The Language of Examples in English Learner's Dictionaries. R.R.K. Hartmann (Ed.). *Lexicographers and their Works*. Exeter: University of Exeter Press.
- Du Plessis, M. 1993. *Tweetalige Aanleederswoordeboek/Bilingual Learner's Dictionary*. Cape Town: Tafelberg.
- Eksteen, L.C. et al. (Eds.). 1997¹⁴. *Groot Woordeboek/Major Dictionary*. Cape Town: Pharos.
- Fontenelle, Thierry et al. (Eds.). 1998. *Euralex '98 Proceedings*. Louvain: University of Louvain.
- Fox, Gwyneth. 1987. The Case for Examples. Sinclair, John M. (Ed.). 1987: 137-149.
- Gouws, R.H. 1989. *Leksikografie*. Cape Town: Academica.
- Gouws, R.H. 1990. Information Categories in Dictionaries with Special Reference to Southern Africa. Hartmann, R.R.K. (Ed.). *Lexicography in Africa: 52-65*. Exeter: University of Exeter Press.
- Gouws, R.H. 2000. Toward the Formulation of a Metalexigraphic Founded Model For National Lexicography Units in South Africa. Wiegand, H.E. (Ed.). 2000. *Wörterbücher in der Diskussion IV: 109-133*. Tübingen: Max Niemeyer.
- Gouws, R.H. 2000a. Doeltaalgerigtheid teenoor lemmagerigtheid in vertalende woordeboeke. *Tydskrif vir Geesteswetenskappe*. 40(1): 39-47.
- Gouws, R.H. 2000b. Strategies in Equivalent Discrimination. Zettersten, A. and V.H. Pedersen. (Eds.). 2000. *Symposium on Lexicography IX*. Tübingen: Max Niemeyer.
- 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. 1977. *Einführung in die Benutzung des neufranzösischen Wörterbücher*. Tübingen: Max Niemeyer.
- Hausmann, F.J. 1986. Grundprobleme des zweisprachigen Wörterbuchs. *Symposium on Lexicography III: 137-154*. Tübingen: Max Niemeyer.
- Hausmann, F.J. et al. (Eds.). 1989. *Wörterbücher. Ein internationales Handbuch zur Lexikographie / Dictionaries. An International Encyclopedia of Lexicography. / Dictionnaires. Encyclopédie internationale de lexicographie*. Berlin: Walter de Gruyter.
- Humble, Philippe. 1998. The Use of Authentic, Constructed and "Controlled" Examples in Foreign Language Dictionaries. Fontenelle, Thierry et al. (Eds.). 1998: 593-599.
- Kromann, H.-P., T. Riiber, P. Rosbach. 1984a. "Active" and "Passive" Bilingual Dictionaries. The Šerba Concept Reconsidered. Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83: 207-215*. Tübingen: Max Niemeyer.

- Kromann, H.-P., T. Riiber, P. Rosbach.** 1984b. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. Wiegand, H.E. (Ed.). *Studien zur neuhochdeutschen Lexikographie. Germanistische Linguistik* 3-6/81: 159-238. Hildesheim: Georg Olms.
- Laufer, Batia.** 1992. Corpus-based versus Lexicographer Examples in Comprehension and Production of New Words. Tommola, Harnu et al. (Eds.). 1992: 71-76.
- Mbogho, Kalumbo.** 1985. Observations on Bilingual Lexicography involving Bantu and Indo-European Languages. *Babel* 31(3): 152-162.
- Minaeva, Ludmila.** 1992. Dictionary Examples: Friends or Foes. Tommola, Harnu et al. (Eds.). 1992: 77-80.
- Otto, A.N.** 1989. *Kriteria vir 'n Afrikaanse aanleerderwoordeboek*. Unpublished D.Litt. thesis. Stellenbosch: University of Stellenbosch.
- Potter, Liz.** 1998. Setting a Good Example. What Kind of Examples Best Serve the Users of Learners' Dictionaries? Fontenelle, Thierry et al. (Eds.). 1998: 357-362.
- Rademeyer, L.** 1992. *Die funksies van verbale voorbeeldmateriaal in eentalige woordeboeke*. Unpublished M.A. thesis. Stellenbosch: University of Stellenbosch.
- Rundell, M.** 1998. Recent Trends in English Pedagogical Lexicography. *International Journal of Lexicography* 11(4): 315-342.
- Sinclair, John M.** (Ed.). 1987. *Looking Up. An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London: Collins ELT.
- Tommola, Harnu et al.** (Eds.). 1992. *Euralex '92 Proceedings*. Tampere: University of Tampere.
- Van der Merwe-Fouché, M.** 1999. *Gebruikersperspektief in die Afrikaanse leksikografie*. Unpublished D.Litt. thesis. Stellenbosch: University of Stellenbosch.
- Wiegand, H.E.** (Ed.) 1996. *Wörterbücher in der Diskussion II*. Tübingen: Max Niemeyer.
- Wiegand, H.E.** 1996a. Zur Einführung. Wiegand, H.E. (Ed.). 1996: VII-XV.
- Wiegand, H.E.** 1996b. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. Wiegand, H.E. (Ed.). 1996: 1-82.
- Wiegand, H.E.** 1998. *Wörterbuchforschung*. Berlin: De Gruyter.

Standards retrodigitaler Wörterbücher*

Michael Schlaefer, *Deutsches Wörterbuch,*
Akademie der Wissenschaften in Göttingen, Deutschland

Zusammenfassung: In den vergangenen Jahren sind zunehmend die Möglichkeiten der rückwirkenden digitalen Textaufbereitung auch für Wörterbücher genutzt worden. Für die vorliegenden deutschsprachigen Wörterbücher zu älteren Sprachstufen und für wortgeschichtliche Wörterbücher stellen sich dabei besondere Probleme. Diese Probleme betreffen nur zu einem Teil technische oder informatische Zusammenhänge. Sie ergeben sich vor allem aus dem lexikographisch-technischen und dem inhaltlichen Zustand der Wörterbücher. Wie die vorliegenden digitalen Wörterbuchbeispiele erkennen lassen, erweisen sich die Möglichkeiten, durch die Retrodigitalisierung für diese Wörterbücher neue, vor allem systematische Nutzungsmöglichkeiten zu erschließen, als teilweise sehr begrenzt. Das Verhältnis von Aufwand und Ergebnis muß im Einzelfall kritisch geprüft werden. Im Zweifelsfall muß die Erarbeitung neuer Wörterbücher den Vorrang vor einer aufwendigen kostenintensiven digitalen Aufbereitung älterer, inhaltlich überholter Werke haben.

Stichwörter: LEXIKOGRAPHIE, DIGITALES WÖRTERBUCH, HISTORISCHES WÖRTERBUCH, MITTELHOCHDEUTSCHES WÖRTERBUCH, GRIMMSCHES WÖRTERBUCH, GOETHE-WÖRTERBUCH, DEUTSCHES RECHTSWÖRTERBUCH, MIKROSTRUKTUREN, LEXIKOGRAPHISCHE METASPRACHE, WÖRTERBUCHKRITIK, ZUKÜNFTIGE WÖRTERBÜCHER

Abstract: Standards for retrodigitalised dictionaries. In recent years, text retrodigitalisation has been applied increasingly to dictionaries. Retrodigitalisation faces special problems in respect of both dictionaries of older language stages and historical dictionaries of German currently in use. There are technical problems as well as issues of computerisation that need to be addressed, but the main obstacle is the poor quality of lexicographical content and technique. Current retrodigitalised dictionaries show the limitations of retrodigitalisation, especially for systematic data retrieval. Cost-benefit analysis should be applied to each individual case. In case of doubt, the creation of new dictionaries should be preferred to the large-scale and expensive retrodigitalisation of older lexicographically antiquated works.

Keywords: LEXICOGRAPHY, DIGITAL DICTIONARY, HISTORICAL DICTIONARY, MITTELHOCHDEUTSCHES WÖRTERBUCH, GRIMMSCHES WÖRTERBUCH, GOETHE-WÖRTER-

* Der Artikel beruht auf einem gleichlautenden Kurzreferat anlässlich der 3. Arbeitstagung deutschsprachiger Akademiewörterbücher vom 22.3. – 25.3.2000 in Zürich. Für ihre Unterstützung bei der Durchsicht und Fertigstellung des Manuskripts danke ich Dr. H. Albrand, U. Härtel, D. Holste, F. Michaelis und F. M. Wohlers.

BUCH, DEUTSCHES RECHTSWÖRTERBUCH, MICROSTRUCTURES, LEXICOGRAPHICAL METALANGUAGE, DICTIONARY CRITICISM, FUTURE DICTIONARIES

Die nachträgliche Umsetzung von Printprodukten in das digitale Medium besitzt seit einiger Zeit eine gewisse Konjunktur, auch wenn der Anteil an retrodigitalen Wörterbüchern im Vergleich zu Literatursammlungen noch gering ist. Die Frage nach den für solche Retrodigitalisierungen beschreibbaren Möglichkeiten und Standards betrifft eine Reihe von technischen und lexikographischen Fragen, auf die im weiteren vor allem unter dem Blickwinkel der Benutzung digitaler historischer deutscher Wörterbücher kurz eingegangen werden soll.

Die Frage der technischen Standards der Retrodigitalisierung läßt sich knapp unter dem Blickwinkel betrachten, ob ein Text als Bildfolge von Buchseiten oder als Folge von Zeichen erfaßt wird. Die digitale Bildseite, das Image, liefert eine authentische Ansicht einer Buchseite des gescannten Exemplars einschließlich ggf. vorhandener farbiger Initialen, handschriftlicher Einträge, Stockflecken usw. Die digitalen Bildseiten lassen sich über Indizierungen nach Seiten, Kolumnentitel, Stichwörtern u. dgl. erschließen und bei der Benutzung aufrufen. Moderne Hochgeschwindigkeitsscanner erlauben bei der technischen Umsetzung des Printproduktes in die digitale Ebene ein sehr ökonomisches Vorgehen.

Die Image-Digitalisierungen bieten den Vorteil der authentischen Seitenwiedergabe, sie gestatten aber nicht, nach Wörtern bzw. Zeichenfolgen zu suchen. Solche Suchzugriffe machen es erforderlich, Zeichen für Zeichen des Printtextes zu erfassen und das Original in einer digitalen Volltextversion zu edieren. Dabei können Texterkennungsprogramme sehr wirksame Unterstützung bieten, wenn die Vorlage bestimmte Bedingungen erfüllt. Sauber gedruckte Antiquadrucke lassen sich so mit fast hundertprozentiger Fehlerfreiheit über solche OCR-Programme erfassen. Für Frakturdrucke oder Drucke mit häufig wechselnden Schriftarten und Schriftgrößen, wie dies in Wörterbüchern vielfach üblich ist, ist die automatische Zeichenerfassung noch sehr fehleranfällig und es führt z.Z. kaum ein Weg daran vorbei, die Texte abzuschreiben.

Das Abschreiben erweist sich dann als weitgehend problemlos, wenn die Texte auf dem normalen Zeichensatz beruhen und als Fließtexte gestaltet wurden. Unbeschadet dieser idealtypischen Situation ist grundsätzlich ein editorisches Verfahren nötig, wenn die typographische Gestaltung des Printtextes in die digitale Version übertragen werden soll. Die digitale Edition der Texte setzt vorrangig ein Kodierungssystem voraus, das eine zeichengenaue und layout-äquivalente Wiedergabe der Textoberfläche erlaubt. Ein entsprechend kodierter Text kann später unter Voraussetzung der erforderlichen Software für Folgeausgaben exakt reproduziert oder effizient überarbeitet werden. Sollen systematische inhaltliche Suchzugriffe auf Teile des digitalen Volltextes erschlossen werden, müssen die betreffenden inhaltlichen Ebenen oder Einheiten

maschinenlesbar kodiert werden. Da Wörterbücher zahlreiche mikro- und makrostrukturelle Inhaltsebenen aufweisen, kann diese inhaltliche Auszeichnung u.U. eine sehr anspruchsvolle Aufgabe darstellen.

Die Textauszeichnung (text encoding oder text markup) kann technisch mit einer Auszeichnungssprache wie SGML (Standard Generalized Markup Language) erfolgen. Dadurch entsteht ein langfristig und plattformneutral nutzbares Textdatenformat. Das Grundelement einer markup-Sprache bildet die DTD (Document Type Definition), die die gültige Grammatik für die Strukturierung eines Dokumentes enthält und alle Markierungen der Textelemente regelt. Die zur Verfügung stehenden wissenschaftlichen oder industriellen Standards können hier als durchaus entwickelt bezeichnet werden.

Retrodigitale Texte lassen sich netzgestützt oder als CD-ROM-Versionen für Benutzer erschließen. Als Repräsentanten der beiden skizzierten technischen Standards seien die Image-Versionen des Adelung'schen Wörterbuchs, des Deutschen Wörterbuchs von H. Paul, des Mittelhochdeutschen Wörterbuchs von M. Lexer und des Deutschen Universalwörterbuchs der Duden-Redaktion genannt. Die Neigung der Verlage, retrodigitale Wörterbuchtexte als CD-Versionen und nicht als Netzversionen anzubieten, ist im wesentlichen auf die Sorge vor unkontrollierter Weiterverwertung der Datenbestände zurückzuführen.

Zur Behandlung der technischen Zusammenhänge lexikographischer Retrodigitalisierungsansätze sei exemplarisch auf die Arbeit von Th. Gloning und Ch. Schlaps (1999) hingewiesen. Hier soll weiterführend der Frage möglicher Benutzungsbedingungen nachgegangen werden. Generalisierend kann zunächst dazu festgestellt werden, daß alle retrodigitalen Wörterbuchversionen den Bedürfnissen der dezentralen Literaturversorgung entsprechen. Diese Nutzungsbedingung ist unabhängig von der Art der Digitalisierung als Image oder Volltextversion. Mitunter kann die Imageversion sogar Vorteile bieten, die bei zeichengetreuer Wiedergabe nur schwer erreichbar sind. Die Imageversion des Adelung'schen Wörterbuchs (Adelung digital) z.B. kann an beliebigen digital ausgestatteten Arbeitsplätzen benutzt werden, ohne daß Bibliotheksgänge anfallen, Kopierverbote zu beachten sind usw. Diese Version bietet zudem das seitengetreue Druckbild und damit eine Eigenschaft, die beim Umgang mit älteren Wörterbüchern besonders wünschenswert erscheinen kann.

Ein weiterer Aspekt der Literaturversorgung durch retrodigitale Wörterbücher ergibt sich aus der Hypertextverkettung verschiedener Werke zu einem Verbund. So kann z. B. das Duden Universalwörterbuch (Deutsches Universalwörterbuch digital) in der "PC-Bibliothek" des Brockhausverlages mit einer Reihe anderer Wörterbücher wie dem Rechtschreib-, Bild- oder Fremdwörterbuch im digitalen Verbund benutzt werden. Für die digitalen Versionen des Lexerschen Mittelhochdeutschen Wörterbuchs und des Mittelhochdeutschen Wörterbuchs von Benecke-Müller-Zarncke gilt entsprechendes (Mittelhochdeutsche Wörterbücher im Verbund). Damit wird es den Benutzern ermöglicht,

verschiedene lexikographische Werke kombiniert am Bildschirm zu benutzen. Neben solchen bibliothekstechnischen Zusammenhängen ist die verlegerische bzw. bearbeitungsbezogene Nutzung der retrodigitalen Produkte anzusprechen. Die einmal retrodigitalisierte Volltextversion kann in der Weiterverwendung effizienter überarbeitet und für andere Wörterbuchvorhaben ausgewertet werden, als dies bei Printprodukten möglich ist. Eine neue Auflage kann auf der bestehenden Textbasis rascher erarbeitet werden. Auch wird es möglich, neue Wörterbücher aus vorhandenen zu gewinnen, indem Texte erweitert, selektiert oder aus verschiedenen Quellen zusammengeführt werden. Als dritte Ebene der Nutzung retrodigitaler Wörterbücher ist ihre systematische Auswertbarkeit in Form von Zeichenfolgensuchen zu nennen. Derartige Suchen sind in Printprodukten sehr umständlich und zeitraubend. Wer einmal alphabetische Printwörterbücher unter selektiven systematischen Gesichtspunkten z.B. bei der Ermittlung von Wortfeldern, Wortfamilien oder anderen Wortschatzausschnitten mit nichtalphabetischer Ausrichtung benutzt hat, weiß den grundsätzlichen Vorteil selbst einfacher digitaler Suchzugriffe zu schätzen. Hier liegt prinzipiell eine große Chance der digitalen Volltextversionen bei der Erleichterung und Effektivierung der Wörterbuchbenutzung.

Die skizzierten technischen Nutzungsverbesserungen haben zu einer verbreitet sehr optimistischen Einschätzung der künftigen retrodigitalen Wörterbuchwelt geführt. Dieser Optimismus verbindet sich mit Utopien u.a. über die Einbindung der Wörterbuchbenutzung in ein künftiges digitales Infotainment oder die Vernetzung des weitgehend retrodigitalisierten Printbuchbestandes (Storror u.a. 1996).

Den zweifellos mit der Benutzung retrodigitaler Wörterbücher verbundenen Vorteilen sind jedoch aus der Kenntnis der Qualität der vorliegenden Wörterbücher einige grundsätzliche Nachdenklichkeiten gegenüberzustellen. Diese Nachdenklichkeiten sollen in ihrem Kern exemplarisch an einem Beispiel für eine offensichtlich gescheiterte systematische Nutzung einer retrodigitalen Wörterbuch-Volltextversion skizziert werden.

In einem Benutzungsexperiment wurde versucht, die wortgeschichtliche Markierung *modern*, die im Artikel ABBRECHEN des digitalen Paulschen Wörterbuchs (Paul digital) erscheint, als Ausgangspunkt für die Zusammenstellung aller Lemmata zu benutzen, in deren Beschreibung dieses Merkmal auftritt. Die Volltextsuche mit der betreffenden Zeichenfolge erbrachte 28 Treffer ohne flektierte Formen. Darunter befand sich das Stichwort MÓDERN und eine bunte Mischung objekt- und metasprachlicher Vorkommen einschließlich einer Form des dänischen Adjektivs in einer Quellenangabe im Artikel ENTWERFEN. Unstrittig liegen unter den Treffern auch Nachweise für die gesuchte wortgeschichtliche Markierung u.a. in den Artikeln STRAND, POTENZ, VERWARNEN usw. vor. Die Aufgabe des Benutzers besteht allerdings hier nicht nur in der Selektion des Suchbegriffs von Formalhomographen. In Kenntnis der Freiheiten der Beschreibungssprache der Wörterbücher wurde vorsorglich auch nach inhaltlich ähnlichen wortgeschichtlichen Markierungen für *modern*

unter der Zeichenfolge *heute* gesucht. Unter den 2000 Treffern für die Zeichenfolge fanden sich zahlreiche wortgeschichtliche Klassifikationen, die synonymisch zu *modern* zu verstehen waren. Ähnlich, wenngleich vom Umfang her deutlich geringer war das Ergebnis mit dem Suchwort *jünger*. Der Versuch, einen mit diesem wortgeschichtlichen Merkmal bestimmten Lexikausschnitt systematisch auf der Grundlage der digitalen Volltextsuche zu ermitteln, ist, wie das Experiment zeigt, kaum gezielt und kontrolliert durchzuführen, weil die digitale Textstrukturierung unscharf ist und die zugrundeliegenden lexikographischen Informationen nur in der Zeichenebene durchsuchbar sind. Man kann sich ausmalen, wie ein Suchergebnis unter vergleichbaren Bedingungen aussähe, wenn die Suchform im Wortstamm flexivisch und orthographisch stark variiert erschiene. Für ein Verb wie *reiten* müßten historische Formen wie *reytten*, *raiten*, *reuthen* usw. alle separat ermittelt werden.

Man kann das skizzierte Ergebnis mit dem Hinweis auf die geradezu radikale Schlichtheit der digitalen Aufbereitungsform dieses Wörterbuchs relativieren. Die Ursache für die Resultate liegt auf der Hand. Da in der behandelten digitalen Version die Zeichenfolgen nur als ausdrucksseitige Einheiten, nicht als sprachliche Zeichen mit Inhalts- und Ausdrucksseite existieren, wird die Zeicheneinheit nur über die Ausdrucksseite konstituiert. Damit ein rechnergestützter Zugriff mit demselben Zeichenverständnis erfolgen könnte, wie es für eine herkömmliche Lesenutzung vorausgesetzt wird, müßte praktisch jede Zeichenfolge der retrodigitalen Version mindestens in ihrem Lemmawert kodiert werden.

Der editorische Standard des digitalen Paulschen Wörterbuchs beschränkt sich auf eine vereinfachte Erschließung der Textoberfläche des Originals. Inhaltlich und datenstrukturell sind nur Stichwörter und Artikel als Einheiten gekennzeichnet bzw. die Stichwortverweise mit Linkfunktionen unterlegt worden. Im interessierenden Bereich ist daher für die Benutzung nichts weiter zu erwarten als die Verknüpfung von Textwortformen und Stichwortformen. Für den wenig anspruchsvollen Benutzer mag dies deutlich mehr sein als nichts, für denjenigen, der z.B. den Robert Électronique kennt, ist dieses Mehr noch sehr nahe am Nichts.

Im Sinn der international erreichten digitalen Editionsstandards wäre zu erwarten, daß z.B. Mikro- und Makrostruktur der Artikel formal und inhaltlich angemessen ausgezeichnet und erschlossen würden. Dabei sind gestufte Lösungen denkbar. Mindestens aber wäre die Nutzung typographischer Textoberflächenmerkmale wie fett, mager, recte und kursiv als Hilfsindikatoren für Inhaltsebenen wie Objekt- und Metasprache zu fordern. Entsprechende Zuordnungen von Textformaten und inhaltlich bestimmten Textebenen sind programmunterstützt und damit automatisch möglich, außer natürlich bei Durchbrechungen des typographischen Systems. Weitergehende inhaltliche oder strukturelle Markierungen betreffen u. a. die Unterscheidung von Formalhomographen wie *modérn* und *móderm* ebenso wie die Unterscheidung von

Stichwörtern, Verweiswörtern, Belegwörtern, Verwendungsbeispielen, metasprachlichen Termini oder flexivisch-orthographischen Varianten der Wörter.

Im Vergleich zum Paulschen Wörterbuch zeigen sich bei jüngeren Retrodigitalisierungsansätzen im deutschsprachigen Raum Ergebnisse, die als Fortschritte zu interpretieren sind. Diese Einschätzung betrifft u.a. das retrodigitale Lexersche Wörterbuch, das im Verbund mit dem Mittelhochdeutschen Wörterbuch von Benecke-Müller-Zarncke und dem Findebuch netzgestützt zur Verfügung steht. Diese Netzversion bietet neben der Volltextsuche eine datenbankorientierte Abfragemöglichkeit auf den Ebenen Stichwort, grammatische Angabe, Siglen, Belegzitate, Etymologie und Bedeutungsangaben. Die z.Z. netzgestützt benutzbaren Datenbankebenen der Mittelhochdeutschen Wörterbücher weisen noch einige technische Einschränkungen auf, die hier jedoch vernachlässigt werden sollen. Der originale Wörterbuchtext kann zudem artikelweise im Zusammenhang gelesen und benutzt werden. Dabei sind alle Literatursiglen mit einem Quellenverzeichnis verlinkt. Über Links sind auch alle Stichwörter innerhalb des Verbundes vernetzt. Damit bietet diese Retrodigitalisierungsversion eine gegenüber der digitalen Version des Paulschen Wörterbuchs erheblich verbesserte technisch-inhaltliche Erschließung. Sie stellt auch eine modellhafte Lösung für Wörterbuchnetze dar.

Die Benutzung der Datenbankabfragen der Mittelhochdeutschen Wörterbücher im Verbund läßt sich an einigen Beispielen demonstrieren. Die Suche nach der Bedeutungsangabe "furchtsam" führt auf 15 Artikel, in denen ein gleichlautendes beschreibungssprachliches Element vorkommt. Diese Artikel können per Link unmittelbar von der Seite mit den Suchergebnissen angesteuert werden. Damit ergibt sich einerseits eine sehr komfortable Verbindung von Datenbank- und Artikelkomponenten, andererseits wird der ursprüngliche Charakter der Artikel als zusammenhängende Texte in wünschenswerter Form in die Benutzung eingebunden. Die Datenbankabfragen zeigen allerdings auch problematische Resultate. Weshalb das metasprachliche Element "Mark" zwar für den Artikel MARC stn. ausgewiesen wird, nicht jedoch für den Artikel HAGEMARCE "geschlossene, eingehgte mark", ist nicht ersichtlich. Möglicherweise kann man diese Erscheinung auf die von den Bearbeitern erwähnte noch nicht abgeschlossene Implementierung aller erforderlichen Softwarevoraussetzungen zurückführen.

Alle Abfragen beruhen offensichtlich auf einer 1:1-Zeichenfolgenabbildung. Der Benutzer kann pro Suchgang nur eine Textwortform angeben. Die Suche nach *Feldweg* in der Ebene der Bedeutungsangaben führt auf den Artikel VUOTERWEC, nicht aber auf den Artikel HAGELKRIUZE, in dem die beschreibungssprachliche Textwortform *feldwegen* auftritt. Dazu muß die Wortform *feldwegen* separat gesucht werden. Damit bietet diese retrodigitale Version zwar die Voraussetzungen für eine strukturierte Suche in verschiedenen Artikelerebenen, sie macht es aber erforderlich, daß der Benutzer seine Suche gut konzipiert und alle Varianten des interessierenden Suchwortes selbst erzeugt, da eine Lemmatisierung nicht erfolgt ist. Da mittelhochdeutsche Wörterbücher,

wie man annehmen darf, vorzugsweise von einem engen Expertenkreis benutzt werden, kann man diese Benutzungsvoraussetzung ggf. tolerieren. Grundsätzlich aber stellt die fehlende Lemmatisierung der Textwortformen eine ganz erhebliche Benutzungsbarriere dar, weil die Konstruktion aller möglichen Flexions- und Schreibvarianten einen großen Arbeitsaufwand und im historischen Bereich sehr gute Sprachkenntnisse voraussetzt.

Für das Goethe-Wörterbuch ist ein retrodigitaler Prototyp (Gloning u.a. 1999) vorgestellt worden, der außer den Artikelgrenzen und den Stichwörtern ähnlich gestufte Zugriffe erlauben wird, wie sie für die Mittelhochdeutschen Wörterbücher skizziert worden sind. Als systematische Abfrageebenen sind Stichwörter, Bedeutungserläuterungen, Belege, Literaturnachweise und Synonymangaben vorgesehen. Ferner ist in diesem Prototypen die Vernetzung der Wörterbuchbelegstellen mit dem jeweiligen literarischen Werktext vorgesehen. Diese Vernetzung erscheint besonders attraktiv, weil sie den Beleg aus seiner lexikographischen Individuation wieder in den textuellen Zusammenhang bringt, ohne den er in vieler Hinsicht problematisch bleibt. Freilich läßt sich ahnen, daß eine über das Modell hinausgehende Umsetzung dieses nützlichen Plans vorläufig an verlagsrechtlichen Barrieren scheitern oder zu gebührenpflichtigen Lösungen führen wird. Ganz ähnlich sind Pläne zu beurteilen, die sich auf eine Vernetzung von Wörterbüchern und aktuellen Enzyklopädien beziehen.

Die Bearbeiter der retrodigitalen Goethe-Wörterbuchversion machen auch auf die Möglichkeit aufmerksam, verschiedene Ansichten der vorhandenen Artikel erzeugen zu können. So soll es möglich gemacht werden, sich z.B. nur die Bedeutungsstruktur eines Artikels unter Ausblendung aller anderen Textelemente zeigen zu lassen. Ferner wird angestrebt, die Belege ergänzend zur gedruckten Version nach verschiedenen Selektionsgesichtspunkten z.B. für den "minimalen Leser" oder einschließlich zusätzlichen Materials abzubilden.

Für das Deutsche Rechtswörterbuch wird ebenfalls z.Z. eine Retrodigitalisierung vorbereitet (Deutsches Rechtswörterbuch digital). Ein Prototyp ist netzgestützt für die Buchstaben I, N, O zugänglich. Das Startmenü bietet eine Buchstabenliste, über die man auf eine Stichwortliste und von dort auf die fensterweise aufgegliederte digitale Artikelfassung kommt. Für das Stichwort *Igel* wird zunächst die Wortklasse Maskulinum angegeben. Die im gedruckten Artikel unter den Gliederungsmarken I und II enthaltenen Positionen werden in zwei weiteren Fenstern als IGEL I und IGEL II überschrieben und mit den jeweiligen metasprachlichen Angaben versehen. Zu IGEL I findet sich die im Originalartikel enthaltene Erklärung "als zum Fang freigegebenes Tier". Ein zugeordnetes Unterfenster bietet als "Navigationshilfe" Links auf Belege und Lemmaliste. Klickt man den Link "Belege" an, öffnet sich ein weiteres Fenster mit der Beleg- und Fundstelle zu IGEL I, das untergliedert ist in je ein Feld für den Belegtext, die Datierung, eine bibliographische Sigle und eine Fundstellennummer. Die bibliographische Sigle ist als Link angelegt und führt auf die vollständige bibliographische Angabe.

Während der Zugriff auf die verschiedenen Informationsschichten der Artikel in der digitalen Version des Rechtswörterbuchs prinzipiell den bisher beschriebenen Ansätzen zu vergleichen ist, wird durch die stark nach Fenstern segmentierte Information der Textzusammenhang der Artikel beeinträchtigt. Artikelübergreifende Recherchezugriffe sind z.Z. in dieser Version noch nicht möglich. Daher kann auch keine Aussage darüber gemacht werden, ob die für das Mittelhochdeutsche Wörterbuch kritisierte fehlende Lemmatisierung hier erfolgte. Positiv hervorzuheben ist in dieser Hinsicht jedoch die vom Ausgangsmenü anzusteuernde Schreibvariantenzuordnung, die die Schreibung *ygel* im Beleg des besprochenen Artikelabschnitts dem Lemma *Igel* zuordnet oder vom *imbs* auf *Imbiß* führt. Eine weitere Nutzungsmöglichkeit der digitalen Version des Deutschen Rechtswörterbuchs wird unter dem Teilmenü "Wortbelegung" geboten. Dort kann das nicht im Wörterbuch abgebildete Belegmaterial ab dem Alphabetabschnitt *Mo-* eingesehen werden.

Ein knapper retrodigitaler Ausschnitt aus M. Heynes Deutschem Wörterbuch (Heyne digital) ist abschließend zu nennen. Das Startmenü bietet unter der Rubrik "Index" in Submenüs verschieden strukturierte Zugriffe auf den Stichwortbestand. Zum einen kann die gesuchte Stichwortform nicht nur als ganzes Wort, sondern auch nach bestimmten Positionen in Komposita und auf Wortarten eingegrenzt gesucht werden. Zum anderen kann für die Ausgabe festgelegt werden, ob die Stichwortform mit Zusatzinformationen wie der Etymologie, den belegten Textwortformen und Verweiszielen angezeigt werden soll. Auf diese Weise können selektive Ansichten der Artikel gezeigt werden. Die Rubrik "Suche" erschließt den Weg zu einer Volltextsuche auf der Ebene nichtlemmatisierter Zeichenfolgen. Diese Suche kann auch selektiv nach bestimmten Textebenen wie Objekt- und Metasprache oder den Textwortformen zu den Stichwörtern gesteuert werden.

Gegenüber dem digitalen Paulschen Wörterbuch können solche Zugriffe technisch als durchaus entwickelt bezeichnet werden. Man wird Ch. Schlaps und Th. Gloning (1999) in der Beurteilung zustimmen, daß es trotz des erreichten Entwicklungsstandes zu den metalexikographischen Zukunftsaufgaben gehören wird, benutzerorientiert nach weiteren sinnvollen Erschließungsmöglichkeiten zu suchen, damit die hohen Investitionen in das erforderliche Markup für die Wörterbücher auch lohnen. Um aber einen solchen Kanon entwickeln zu können, muß nicht nur die Frage nach den gewünschten Nutzungsmöglichkeiten geklärt werden. Es muß auch geprüft werden, welche Standards in den Wörterbüchern vorliegen und von welchen Voraussetzungen Markup-Versuche ausgehen können. Dazu sollen im Anschluß einige Sachverhalte in historischen Wörterbüchern betrachtet werden. Darunter werden lexikographische Standardwerke gefaßt, die wie z. B. das Grimmsche Wörterbuch, das Deutsche Wörterbuch von H. Paul oder das Deutsche Wörterbuch von M. Heyne die deutsche Sprache in ihrer historischen Gesamtdimension beschreiben.

Die lexikographische Texterstellung in diesen historischen Wörterbüchern zielt vielfach auf eine im Zusammenhang lesend entstehende Informationsentnahme durch die Benutzer, nicht darauf, standardisierte Informationen im systematischen Zusammenhang zu erschließen. Dabei wirken die von Vorstellungen des 19. Jahrhunderts geprägten lexikographischen Formen bzw. Benutzungsprojektionen bestimmend. So werden vor allem Textgestaltungsmöglichkeiten ausgeschöpft, die den Einzelartikel bzw. das Einzelwort anschaulich werden lassen. Die lexikographische Strukturkonsistenz ist ebenso wie die metasprachliche Konsequenz und die Ausführung von Vernetzungen meist nur ansatzweise entwickelt. Der Anteil frei umschriebener, elliptisch formulierter bzw. impliziter Information erweist sich als sehr hoch. Diskursive Tendenzen überlagern oder durchkreuzen die unterschiedlich entwickelten Gliederungsansätze.

Für die Textstrukturierung, die zur späteren systematischen digitalen Auswertung erforderlich ist, entstehen aufgrund solcher Textformen erhebliche Barrieren. Ein effizientes Markup, das über die automatische Kennzeichnung von typographischen Anfangs- und Endsignalen hinausgeht, ist kaum möglich. Eine solche automatische Auszeichnung der Texte wäre auch dann nicht mehr sinnvoll, wenn die typographischen Formate zwar als Indikatoren für inhaltliche Sachverhalte zu bestimmen sind, jedoch in einer diskursiven Verflechtung erscheinen, die bei einem selektierenden digitalen Zugriff zu weitgehend sinnlosen Textschnitzeln führte. Die häufig unscharfe mikrostrukturelle Gliederung der Artikel in den Printversionen kann nur dann für systematische digitale Suchen erschlossen werden, wenn zuvor eine Restrukturierung durchgeführt wird. Dabei muß der ursprüngliche Wörterbuchtext aufgrund kompetenzgestützter Interpretationen manuell ausgezeichnet werden. Dieser Arbeitsgang ist auch dann zwingend erforderlich, wenn typographische Formate im Text in der inhaltlichen Indizierungsfunktion ambivalent bleiben oder inkonsequent verwendet wurden. Wie die Untersuchungen von K. Casemir zum Grimmschen Wörterbuch zeigen, ist das typographische System vom Ansatz her mehrdeutig und unterliegt zudem beträchtlichen Schwankungen. Der metasprachliche Beschreibungstext ist z.B. nicht nur regulär in kursiver Normalschrift ausgeführt, sondern erscheint u.a. auch recte gesperrt und kursiv gesperrt. Lemmavarianten werden in Versalien sowie — wie im übrigen auch die Belegtexte in Normalschrift — recte und kursiv ausgeführt usw. Eine automatische Umsetzung der typographischen Grenzschnitte in Kodierungen des Markups müßte hier zwangsläufig zu einer beträchtlichen Ambivalenz der Zuordnung von typographischem Format und Inhaltsebene führen.

Die Defizite, die hier im lexikographisch-technischen Bereich erkennbar werden, lassen sich zweifellos mit einigem Aufwand durch Nachbearbeitung beheben. Dabei stellt sich naturgemäß die Frage nach der Relation von Aufwand und Ergebnis. Diese Frage stellt sich noch entschieden schärfer unter dem Blickwinkel der lexikographisch-inhaltlichen Qualität.

In der bisherigen Diskussion um die Volltextretrodigitalisierung tritt die Beurteilung der inhaltlichen Wörterbuchqualität gegenüber dem Problem der Wörterbuchform und den datentechnischen Problemen weitgehend zurück. Offensichtlich geht man allgemein davon aus, daß die verfügbaren Werke konzeptionell und inhaltlich trotz des z.T. beträchtlichen Alters und bekannter Kritikpunkte so unstrittig sind, daß auch große finanzielle Investitionen gerechtfertigt erscheinen. Diesem Understatement entsprechen nicht alle Beobachtungen im hier berücksichtigten Spektrum.

Lexers Mittelhochdeutsches Wörterbuch oder das Grimmsche Wörterbuch in seiner abgeschlossenen Ausgabe werden seit langem als veraltet beschrieben (u.a. Schlaefler 1999: 94-95). Neue Werke oder Teilbearbeitungen sollen die alten Werke ersetzen. Dazu werden mit öffentlichen Mitteln erhebliche Aufwendungen gemacht. So wird für das Grimmsche Wörterbuch seit 1960 eine Neubearbeitung der Teile A-F durchgeführt. Für die Sprachstufe des Mittelhochdeutschen wurde unter Hervorhebung der Mängel der vorhandenen Wörterbücher (Mittelhochdeutsches Wörterbuch in der Diskussion 1988) durch die Göttinger Akademie der Wissenschaften (Jahrbuch 1998: 281) ein neues Langfristprojekt eingeleitet. Die kontinuierlichen Nachträge zum Lexerschen Wörterbuch einschließlich der Publikation des Findebuchs lassen ebenso wie Vergleiche der alten und der neuen Grimm-Ausgabe erkennen, daß in den alten Werken tatsächlich Lücken bestehen und aktuelle Forschungsinteressen nicht mehr adäquat befriedigt werden können. Dabei geht es offensichtlich nicht nur um Stichwortlücken, sondern auch um die Zusammensetzung der Quellenbasis, den Artikelaufbau und die Art der Wortbeschreibung.

Solche Mängel betreffen die Möglichkeiten, mit den retrodigitalen Wörterbuchversionen neue, vor allem systematische Recherchemöglichkeiten zu erschließen, in ganz erheblicher Weise, wie die folgenden Befunde veranschaulichen. Da Lexers Beschreibungssprache nicht nur sehr spontan und übersetzungsorientiert, sondern auch sehr auf Knappheit angelegt ist, sind weggelassene Bedeutungsangaben (Artikel WIDERSPRÜCHIG) oder referierende Verweise wie *dasselbe* (Artikel WIDERVEHTE) *ib.* (Artikel WIDERSPILN) ganz geläufig. Dazu kommen lateinische Interpretamente (Artikel WIDERSPRECHUNGE) als Bedeutungsangaben. In der beschriebenen Weise entsteht eine sehr uneinheitliche Informationsoberfläche, die zwar bei der Artikellectüre, nicht aber bei einem selektiven systematischen Zugriff kompensiert werden kann. Die oben skizzierte datenbankorientierte Suchabfrage für das Lexersche Wörterbuch findet damit nicht nur an der fehlenden Lemmatisierung der Textwörter ihre Grenze. Sie bleibt zwangsläufig defizitär, weil das verarbeitete Printprodukt schon Defizite und Inkonsistenzen aufweist. Die beste DTD kompensiert nicht das Fehlen der Bedeutungsangaben. Ganz analog zeigen sich beschreibungssprachliche Inkonsistenzen in anderen Wörterbüchern (Dummer u.a. 1998: 219). Hier scheint es geraten, die für retrodigitale Wörterbücher prophezeiten "neuen Nutzungsmöglichkeiten" (Mittelhochdeutsche Wörterbücher im Verbund) vor allem im Hinblick auf systematische Zugriffe

mit großer Zurückhaltung zu beurteilen. Die neue, systematische Nutzung ersetzt im skizzierten Fall für wissenschaftliche Zielsetzungen keineswegs die traditionelle Lesenutzung der Artikel.

Die Aufwendungen für die Retrodigitalisierung erscheinen vor dem skizzierten Hintergrund nicht ohne weiteres schlüssig. Man wird aber das Argument zulassen müssen, daß es gegenwärtig keine besseren mittelhochdeutschen Wörterbücher gibt und daß das begonnene neue mittelhochdeutsche Wörterbuch voraussichtlich erst nach dreißig und mehr Jahren abgeschlossen zur Verfügung stehen wird. Für die heutige Forschergeneration und vermutlich auch noch für die folgende besteht durchaus ein Bedarf an digitaler Wörterbuchinformation, der mit dem vorgelegten Wörterbuchverbund provisorisch zu befriedigen ist. Hingegen kann die Begründung, die Forschungstradition müsse zugänglich gehalten werden, nur bedingt als Rechtfertigung für diese Art der aufwendigen Volltexterschließung betrachtet werden. Zum einen ist die Literaturversorgung mit den Printexemplaren der mittelhochdeutschen Wörterbücher durchweg zufriedenstellend, zum anderen könnten mit einer erheblich kostengünstigeren Image-Digitalisierung, die netzgestützt bereitgestellt wird, bestehende Literaturversorgungslücken geschlossen oder eine Dezentralisierung der Literaturversorgung rasch erreicht worden sein. Daß der retrodigitale Text der älteren mittelhochdeutschen Wörterbücher für eine künftige Überarbeitung oder Neuausgabe des Originals verwendet würde, scheint eine kaum realistische Perspektive.

Für den historischen neuhochdeutschen Bereich ergeben sich hinsichtlich der inhaltlichen Voraussetzungen für die Wörterbuchretrodigitalisierung ähnliche Sachverhalte wie bei den mittelhochdeutschen Wörterbüchern. Hier ist besonders die Rolle des Grimmschen Wörterbuchs zu beachten. Das Grimmsche Wörterbuch nimmt von der zweiten Hälfte des 19. Jahrhunderts bis in die Gegenwart eine dominante Rolle im schriftsprachlichen neuhochdeutschen Segment der Wörterbuchlandschaft ein. Diese Dominanz ist zwar durch die Entstehung des Frühneuhochdeutschen Wörterbuchs und eines entwickelten Spektrums von gegenwartssprachlichen Wörterbüchern jeweils an den Periodengrenzen bzw. durch das Deutsche Rechtswörterbuch und das Deutsche Fremdwörterbuch auf speziellen lexikalischen Ebenen in gewisser Weise eingeschränkt, jedoch nicht aufgehoben worden. Für den Kernbereich des historischen Neuhochdeutschen und den größten Teil des Alphabets steht damit ausschließlich die abgeschlossene Ausgabe des Grimmschen Wörterbuchs zur Verfügung.

Ganz fraglos stellt diese Ausgabe die umfangreichste Dokumentation zur historischen deutschen Lexik dar. An kritischen inhaltlichen und konzeptionellen Einschätzungen des Grimmschen Wörterbuchs fehlt es aber ebenso fraglos seit seiner Gründung nicht (vgl. Schlaefer 1999). Innerhalb dieser Einschätzungen kommt im gegebenen Zusammenhang den Stimmen ein besonderes Gewicht zu, die monieren, daß die konzeptionellen und materiellen Grundlagen der Artikel oft kaum einzuschätzen sind und daß die Eigenwilligkeiten

der Bearbeiter zu einer hochgradigen qualitativen Uneinheitlichkeit der Artikel geführt haben. Viele Artikel sind in ihrem Informationswert ohne die Aufarbeitung wörterbuchgeschichtlicher Hintergrundkenntnisse nur schwer oder gar nicht zu beurteilen. Die Erstellung einer wissenschaftlichen Ansprüchen genügenden retrodigitalen Volltextversion dieses Wörterbuchs steht damit vor deutlich größeren Problemen, als sie etwa für das Lexersche Wörterbuch beschreibbar sind. Zur Veranschaulichung derartiger Qualitätsprobleme sei exemplarisch auf einige Beobachtungen zum Artikel FINGER verwiesen.

Der von J. Grimm in der abgeschlossenen Ausgabe bearbeitete Artikel FINGER steht prototypisch für ein gelehrtes, aber unorganisiertes lexikographisches Arbeiten, das sich letztlich einer sinnvollen systematischen Erschließung entzieht. Unter dem Gliederungsabschnitt 1) (Deutsches Wörterbuch 3: 1650) werden Belege gebucht, die das Stichwort gar nicht enthalten, die weiteren 22 Gliederungsabschnitte mischen nach wenig nachvollziehbaren Gesichtspunkten Phraseologisches, Welt- und Völkerkundliches. Unter Gliederungsabschnitt 2) heißt es einleitend *der ausgestreckte finger, zumal der rechten hand deutet, (...) gewinkt wird mit dem zeigefinger oder mit der ganzen hand, gedroht wird mit dem zeigefinger, der vierte finger heiszt der namenlose oder ungenannte (impudicus), finn. nimetöin sormi, und wenn auch den nordamerikanischen Mönilarris der dritte finger 'der ohne namen' genannt wird, ist es doch unser vierter, weil sie den daumen besonders rechnen und vom ersten finger zu zählen beginnen.* (Deutsches Wörterbuch 3: 1651). Eine sprachgeschichtliche Information zum Stichwort *Finger* ist in diese Passage nur mühsam hineinzulesen. Und auch wenn es möglich wäre, in einer digitalen Version hier das Stichwort *Zeigefinger* zu finden, böte der Text doch auch dazu wenig sprachlich Relevantes. Die Hinweise auf die kulturhistorischen Implikationen der Fingergestik ist im Sprachwörterbuch ebenso überflüssig wie der Hinweis unter 3), der sich darauf bezieht, welche Fingerformen an Frauen als schön empfunden werden. Unter den Gliederungsabschnitten 5) und 6) des Artikels FINGER fehlt der Beschreibungstext ganz, dafür wird das Belegmaterial unter 13) um ein altirisches Zitat ergänzt. Die Quellenlage ist fast ausschließlich literarisch. Soweit es den neuhochdeutschen Zeitraum betrifft, sind schriftsprachliche schweizerische und österreichische Quellen weitgehend vernachlässigt worden.

Solche und ähnliche Befunde lassen sich auch in den anderen Teilen des Wörterbuchs machen. Dazu kommt in den späteren Teilen das Problem des völlig unübersichtlichen Artikelumfangs. Artikel wie GRUND oder STEHEN bedürften dringend der Neufassung und dies nicht nur, weil veraltete Textausgaben zitiert werden oder sekundäre Wörterbuchzitate 30% der Belege ausmachen, sondern weil sie umfangsbedingt nicht mehr die Merkmale der Textsorte "Wörterbuchartikel" aufweisen. Hier würde auch die Möglichkeit, verschiedene Ansichten der Artikel wie z.B. die, nur Bedeutungsgliederung im Zusammenhang ausfiltern zu können, zu nichts führen, weil Abfolgen von "Bedeutungsangaben" wie unter STEHEN II, A, 7, b, v, bb) allenfalls noch in der diskursiven Textverflechtung einen Sinn machen. Daß solche Elemente

ernsthaft in einer artikelübergreifenden datenbankorientierten Abfrage nach beschreibungssprachlichen Einheiten eine verwertbare Information ergäben, ist nicht ersichtlich. Unter der bezeichneten Stelle finden sich folgende durch Zählung einem Gliederungsabschnitt zugewiesene beschreibungssprachliche Angaben: bb) *besonders häufig*, cc) [ohne Beschreibungstext], dd) *dafür auch*, (...) *gern in freiere gebrauchswesen übergehend*, o) [ohne Beschreibungstext], c) *mannigfache adverbiale und präpositionale ausdrücke werden mit seite gebildet*, α) [ohne Beschreibungstext], β) [ohne Beschreibungstext], γ) *mit näherer bestimmung im gen. (od. possessivpron.); so schon ahd.* (Deutsches Wörterbuch 10, 2, 9: 1495-1496). Die vielen eingestreuten Einschübe mit *ferner*, *auch*, *gewöhnlich*, die vielfach eigene Gliederungsabschnitte konstituieren, wurden nicht berücksichtigt. Die retrodigitale Volltextversion wird hier nicht mehr als den zwingend im Zusammenhang zu lesenden Text erschließen können. Falls die Zeichenfolgen nicht lemmatisiert werden, ist auch auf die Belege für *stehen* eben nur dann ein Zugriff möglich, wenn der Benutzer weiß, daß er *ston*, *stund*, *stand*, *st* *M*nd, *stant*, *sten* usw. suchen muß.

Andere Inhaltsprobleme der abgeschlossenen Ausgabe des Grimmschen Wörterbuchs können nur angedeutet werden. Auch in der retrodigitalen Version des Grimmschen Wörterbuchs bleibt die BLINDSCHLEICHE eine *blinde giftige schlange*, und unter dem Stichwort JUDE wird es weiterhin in dumpfem Antisemitismus lauten *von ihren (der Juden) schlimmen eigenschaften werden namentlich ihre unreinlichkeit, sowie ihre gewinnsucht und ihr wuchersinn in mannigfachen wendungen betont*. Für die 13 Stichwörter der Spalte 2391 in Band 4,2 von JUNGFRAUENMÄGDLEIN bis JUNGFRÄULEIN werden wie in vielen anderen Artikeln entweder gar keine oder nur kryptische Bedeutungsangaben stehen. SINKEISEN wird als Stichwort erscheinen, obwohl es nur niederdeutsch über einen Wörterbucheintrag nachgewiesen ist. Dafür werden auch in der retrodigitalen Version des DWB fast alle Fremdwörter fehlen und darüber hinaus alles, was nach dem jeweiligen, oft mehrere Generationen zurückliegenden Bearbeitungsschluß zu buchen gewesen wäre. Der Mängelkatalog könnte um zahlreiche Beobachtungen verlängert werden.

Keine noch so sorgfältige retrodigitale Aufbereitung wird diese Probleme beheben können. Die Frage nach Aufwand und Ergebnis einer Volltextdigitalisierung, die schon für die mittelhochdeutschen Wörterbücher gestellt worden ist, kann daher für das Grimmsche Wörterbuch kaum noch positiv beantwortet werden. Die Abfragen, die an digitale Volltextversionen auf den Ebenen der Stichwörter, der Objekt- oder Metasprache, der Etymologien, Belege und Quellennachweise gerichtet werden könnten, bleiben in weit höherem Maß, als es aus den angedeuteten Defekten des Lexerschen Wörterbuchs ersichtlich wurde, Stückwerk. Sie führen auf wenig Nützliches wie im Artikel *Finger* oder auf Abzulehnendes wie im Artikel *Jude*. Angesichts solcher Gegebenheiten fällt es schwer, sich vorzustellen, worin der mit einer Volltextversion erreichbare retrodigitale "Mehrwert" für die Wörterbuchbenutzer bestehen sollte. Die Tatsache, daß es für den fraglichen Objektbereich kein besseres Wörterbuch gibt

und daß die Neubearbeitung des Grimmschen Wörterbuchs nur einen sehr beschränkten Buchstabenbereich revidiert anbietet, sollte weniger retrodigitale Bemühungen als lexikographische Neuansätze anspornen.

Retrodigitale Wörterbuchversionen, so ist zusammenfassend festzustellen, bieten in der Literaturversorgung und bei bestimmten systematischen Nutzungszugriffen zweifellos eine Verbesserung der bestehenden Möglichkeiten. Lexikographisch und inhaltlich sind die retrodigitalen Wörterbücher jedoch stets bestenfalls so gut wie ihre Printvorlagen. Die im einzelnen erreichbare retrodigitale Qualität hängt daher nicht nur von der Stringenz der DTD ab, sondern vor allem von der Qualität des verwendeten Printproduktes. Diese Qualität ist für einen Teil der Wörterbücher nachdrücklich in Frage zu stellen. Die retrodigitalen Produkte heben diese Mängel nicht auf, sondern tradieren sie. Systematische Abfragemöglichkeiten dürfen nicht darüber hinwegtäuschen, daß sie nur datentechnisch systematisch sind, inhaltlich jedoch eine vielfach lücken- und fehlerhafte Basis erschließen. Das Operieren mit dem Systematikbegriff bei retrodigitalen Wörterbüchern aus dem behandelten Spektrum sollte daher sehr zurückhaltend erfolgen, wenn nicht schwerwiegende Mißverständnisse in Kauf genommen werden sollen. Die inhaltlichen Mängel der historischen Lexikographie sind kaum durch Retrodigitalisierungen, verbreitet nicht einmal durch Überarbeitung, sondern nur mit zeitgemäßen lexikographischen Konzepten zu beheben. Diese Einsicht ist, wie die Argumentation für die Neubearbeitung des Grimmschen Wörterbuchs vor vierzig Jahren zeigt, nicht ganz neu und völlig unabhängig von der Entwicklung retrodigitaler Techniken.

Literaturhinweise

- Adelung *digital*. <http://www.ub.uni-bielefeld.de/diglib/adelung/grammati/>.
- Adelung, J. Ch. 1808. *Grammatisch-kritisches Wörterbuch der hochdeutschen Mundart, mit beständiger Vergleichung der übrigen Mundarten, besonders aber der Oberdeutschen, I-IV*. Wien: Soltau.
- Benecke, G. F., W. Müller und F. Zamcke. 1854-1861. *Mittelhochdeutsches Wörterbuch, I-III*. Leipzig: S. Hirzel.
- Bergmann, R. (Hg.). 1998. *Probleme der Textauswahl für einen elektronischen Thesaurus*. Beiträge zum ersten Göttinger Arbeitsgespräch zur historischen deutschen Wortforschung, 1. und 2. November 1996. Hg. im Auftrag der Akademie der Wissenschaften in Göttingen v. R. Bergmann. Stuttgart: S. Hirzel.
- Casemir, K. 2000. Zur Digitalisierung des Deutschen Wörterbuchs von Jacob Grimm und Wilhelm Grimm. *Sprachwissenschaft* 25: 77-102.
- Deutsches Fremdwörterbuch*. 1913 ff. *Deutsches Fremdwörterbuch, I ff.* Straßburg (Berlin u.a.): Trübner (de Gruyter).
- Deutsches Rechtswörterbuch*. 1914 ff. *Wörterbuch der älteren deutschen Rechtssprache, I ff.* Weimar: Böhlau.
- Deutsches Rechtswörterbuch digital*. <http://www.rzuser.uni-heidelberg.de/~cd2/drw/frameset.htm>.
- Deutsches Wörterbuch digital*. <http://gaer27.uni-trier.de/GrimmWB/grimmwb.htm>.

- Deutsches Wörterbuch*. 1854-1971. Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm, I-XVI. Leipzig: S. Hirzel.
- Deutsches Wörterbuch. Neubearbeitung*. 1960 ff. Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm. Neubearbeitung. Hg. v. der Berlin-Brandenburgischen Akademie der Wissenschaften und der Akademie der Wissenschaften zu Göttingen, I ff. Leipzig (/Stuttgart): S. Hirzel.
- Deutsches Universalwörterbuch digital*. Duden. Deutsches Universalwörterbuch. o. J. A-Z, 3. A. Version 1.1, PC-Bibliothek, Mannheim/Wien/Zürich: Dudenverlag.
- Dummer, S., F. Michaelis, M. Schlaefer. 1998. Zur Digitalisierung historischer Wörterbücher. *Lexikos* 8: 1-29.
- Findebuch*. 1992. Findebuch zum mittelhochdeutschen Wortschatz. Mit einem rückläufigen Index. [Bearb. v.] K. Gärtner u.a. Stuttgart: S. Hirzel.
- Gloning, Th. und Ch. Schlaps. 1999. Prototypen für ein elektronisches Goethe-Wörterbuch. *Sprache und Datenverarbeitung. International Journal for Language Data Processing* 2: 20-33.
- Goethe-Wörterbuch*. 1978 ff. Goethe-Wörterbuch. Hg. v. der Berlin-Brandenburgischen Akademie der Wissenschaften, der Akademie der Wissenschaften zu Göttingen und der Heidelberger Akademie der Wissenschaften, I ff. Stuttgart/Berlin: Kohlhammer.
- Heyne digital*. Heyne, M. 1890-1895. *Deutsches Wörterbuch, I-III*. Leipzig: Hirzel. <http://Grimm.ADW-Goettingen.gwdg.de/heyne/hdw.html>.
- Heyne, M. 1890-1895. *Deutsches Wörterbuch, I-III*. Leipzig: Hirzel.
- Jahrbuch* 1998. Jahrbuch der Akademie der Wissenschaften zu Göttingen 1998 (1999). Göttingen: Vandenhoeck u. Ruprecht.
- Lauer, B. (Hg.). 1999. *Die Brüder Grimm und die Geisteswissenschaften heute*. Ein wissenschaftliches Symposium der Brüder Grimm-Gesellschaft e. V. in der Paulinerkirche zu Göttingen am 21. und 22. November 1997. Schriften der Brüder Grimm-Gesellschaft, N.F. 30, Kassel: Brüder Grimm-Gesellschaft.
- Lexer digital*. Lexer, M. 1872-1878. *Mittelhochdeutsches Handwörterbuch, I-III*. Leipzig: S. Hirzel <http://gaer27.uni-trier.de/MhdWB/>.
- Lexer, M. 1872-1878. *Mittelhochdeutsches Handwörterbuch, I-III*. Leipzig: S. Hirzel.
- Mittelhochdeutsches Wörterbuch in der Diskussion*. (1988). Symposium zur mittelhochdeutschen Lexikographie, Hamburg, Oktober 1985. Tübingen: Niemeyer.
- Mittelhochdeutsche Wörterbücher im Verbund*. o. J. Mittelhochdeutsche Wörterbücher im Verbund. Benecke/Müller/Zarncke — Matthias Lexer — Nachträge zum Lexer — Findebuch. Die wichtigsten lexikographischen Hilfsmittel für das Studium älterer deutscher Texte auf CD-ROM und im Internet. Entwickelt im Fach Germanistik der Universität Trier in Zusammenarbeit mit dem Kompetenzzentrum "Neue Publikationsformen für geisteswissenschaftliche Grundlagenwerke", in Verbindung mit der Akademie der Wissenschaften und der Literatur zu Mainz. (<http://gaer27.uni-trier.de/MWV-online/MWV-online.html>).
- Paul, H. 1992. *Deutsches Wörterbuch*. Neubearb. A. v. H. Henne u. G. Objartel unter Mitarbeit v. H. Kämper-Jensen. Tübingen: Niemeyer.
- Paul digital*. Paul, H. 1992. *Deutsches Wörterbuch*. Neubearb. A. v. H. Henne u. G. Objartel unter Mitarbeit v. H. Kämper-Jensen. o.O., o.J.
- Ris, R. 1998. *Der schweizerische Anteil in den deutschen Großwörterbüchern*. Bergmann, R. (Hg.). 1998: 113-124.

- Le Robert Électronique*. 1994. *Le Robert Électronique Dos-Macintosh-Windows* (CD-Rom) Paris: Robert.
- Schlaefer, M.** 1999. *Das Grimmsche Wörterbuch in der deutschen Wörterbuchlandschaft*. Lauer, B. (Hg.) 1999: 93-124.
- Storrer, A. und K. Freese.** 1996. Wörterbücher im Internet. *Deutsche Sprache* 2: 97-153.

Using Frames in Special-Field Lexicography: An Ethnomusicological Case Study*

Maria Smit, *Department of Music, University of Stellenbosch, Stellenbosch, Republic of South Africa*

Abstract: The distinction between encyclopedic and semantic knowledge is well-known. The influential German metalexigrapher, Wiegand, also identifies a third kind of knowledge, which he calls special-field object-constituting knowledge. It is precisely this kind of knowledge users need in order to know to which object a particular lexical item refers, otherwise they cannot use that lexical item referentially in communication.

This article focuses on Konerding and Wiegand's (1994) discussion of frames and on their suggestions with regard to frame-based dictionary articles. Certain types of information are of great importance if the object-constituting knowledge is to be conveyed in dictionary articles. The article also investigates the possibilities of using frames in special-field lexicography. A sample article from an ethnomusicological database is presented in which the frame for artefacts, as worked out by Konerding, is applied and discussed. By means of frames, lexicographers can systematically ensure that the special-field object-constituting knowledge is indeed present in the articles they write.

Keywords: SPECIAL-FIELD LEXICOGRAPHY, FRAME THEORY, FRAMES, LEARNER'S DICTIONARIES, SPECIAL-FIELD ENCYCLOPEDIA DICTIONARIES, SPECIAL-FIELD OBJECT-CONSTITUTING MEANING KNOWLEDGE, MATRIX FRAMES, MINIMAL FRAMES, ITEM CLASSES, PREDICATE CLASSES

Opsomming: Die gebruik van rame in vakensiklopediese leksikografie: 'n Etnomusikologiese gevallestudie. Die onderskeid tussen ensiklopediese en semantiese kennis is welbekend. Die invloedryke Duitse metaleksikograaf, Wiegand, identifiseer ook 'n derde tipe kennis, wat hy vakensiklopediese objek-konstitutiewe kennis noem. Dit is juis hierdie soort kennis wat gebruikers nodig het om te weet na watter objek 'n bepaalde leksikale item verwys, anders kan hulle nie die leksikale item in 'n verwysende funksie in kommunikasie gebruik nie.

Hierdie artikel fokus op Konerding en Wiegand (1994) se bespreking van rame en hulle voorstelle in verband met raam-gebaseerde woordeboekartikels. Sekere tipes inligting is van uiterste belang om objek-konstitutiewe kennis in woordeboekartikels weer te gee. Die artikel ondersoek ook die moontlikhede om rame in vakleksikografie te gebruik. 'n Voorbeeldartikel uit 'n etnomusikologiese databasis word aangebied, waarin die raam vir artefakte, soos uitgewerk deur Koner-

* This paper was presented at the Fifth International Conference of the African Association for Lexicography, held at the University of Stellenbosch, Stellenbosch, 3-5 July 2000.

ding; toegepas en bespreek word. Deur middel van rame kan leksikograwe op 'n sistematiese wyse seker maak dat die objek-konstitutiewe kennis inderdaad in die artikels wat hulle skryf, aanwesig is.

Sleutelwoorde: VAKLEKSIKOGRAFIE, RAAMTEORIE, RAME, AANLEERDERWOORDE-BOEKE, VAKENSIKLOPEDIËSE WOORDEBOEKE, VAKENSIKLOPEDIËSE OBJEK-KONSTITUTIEWE BETEKENISKENNIS, MATRIKSRAME, MINIMALE RAME, ITEMKLASSE, PREDIKAAATKLASSE

1. Introduction

Sometimes when speakers look up special-field terms, be it in general monolingual dictionaries, special-field dictionaries or in dictionaries for special purposes, they need to know to which object a particular expression refers, otherwise they are unable to use the expression or lexical item referentially during communication. This article discusses Konerding and Wiegand's suggestions on frame-based dictionary articles to meet this specific user need, and a sample article from an electronic ethnomusicological database being compiled at the University of Stellenbosch¹ is shown.

The electronic database can be classified as a hybrid, because it contains features of a special-field encyclopedic dictionary² and a learner's dictionary. Pilegaard (1994: 211) uses the expression "encyclopedic bilingual LSP dictionary"³, because in his view (Pilegaard 1994: 212), users (especially students) need explanations of the meaning of lexical items as well as illustrations of their use.⁴ Users need "subject information and language information", as he (Pilegaard 1994: 215) puts it. If there is no subject information, "it seriously restricts the potential use of the dictionary". Pilegaard (1994: 217) proceeds: "This problem is particularly acute where the user is short of time or does not have competent primary sources of information at hand; a situation which is the rule rather than the exception both for scholars, students and professionals working with LSP communication."

2. Wiegand's typology of special-field dictionaries

In his discussion of special-field lexicography, it seems wrong for Wiegand (1988: 742) that one would try to distinguish between language dictionaries and encyclopedias simply on the basis of the distinction between semantic items and encyclopedic items that the articles might contain. He (Wiegand 1988: 743) claims that researchers have problems with such a distinction, because they only approach it from a linguistic or philological angle. Wiegand suggests it would be more fruitful if one would approach the classification problem from the point of view of social action theory or communication theory. He therefore pleads for a classification of reference works according to their genuine

purposes. The genuine purposes of reference works are that they can be used to obtain information on the object (for example, the language) from the lexicographical data in them.⁵

Wiegand (1994: 105) defines the genuine purpose of a special-field dictionary by stating that it is an object used to obtain information that may be considered as special-field knowledge. He (Wiegand 1988: 745) notes that one could classify all reference works as belonging to two classes, either to class K1 or class K2. Class K1 may be the class of linguistic objects, and class K2 the class of nonlinguistic objects or "things" ("Sachen"). If reference works are classified as classes K'1, 2, 3, etc., class K'1 is the class of all reference works with a genuine language-lexicographical purpose, whereas class K'2 is the class of all reference works with a genuine subject-lexicographical purpose. A genuine language-lexicographical purpose exists when a potential user may obtain information about a linguistic object from the lexicographical data of class K'1.⁶ A genuine subject-lexicographical purpose exists when a potential user may obtain information about an object or "thing" ("Sache") from the lexicographical data of class K'2.⁷

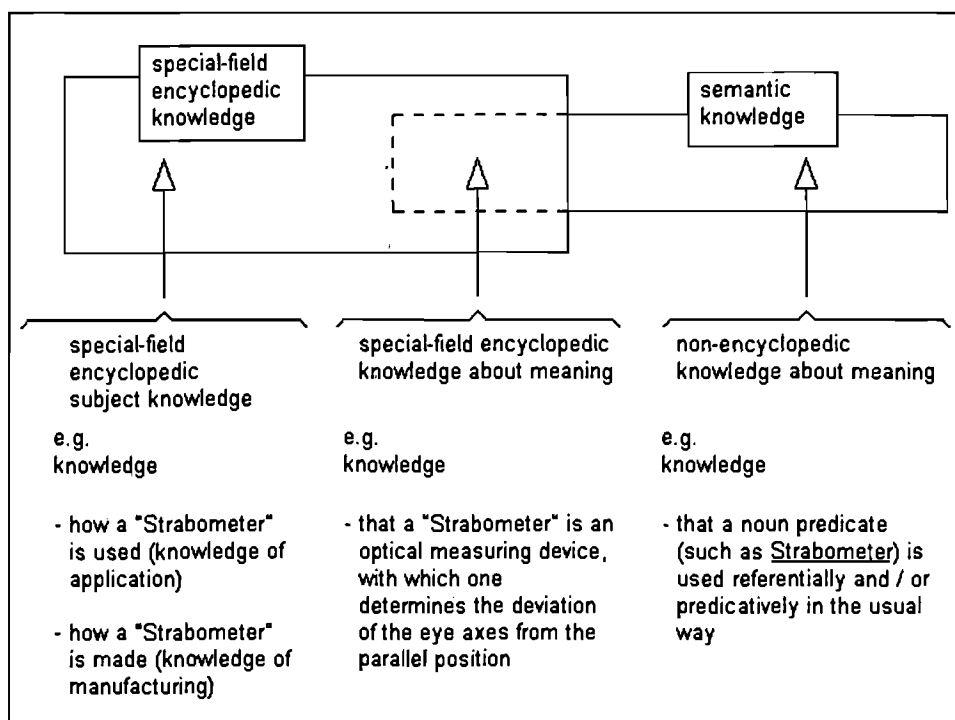
In Wiegand's (1988: 746-747) view there is, however, another possibility which lies between these two distinct classes of reference works. Instead of distinguishing only between semantic data on the one hand and encyclopedic data on the other hand, Wiegand (1994: 125) also distinguishes a third type of data. He hopes that this would contribute to a better understanding of the article structures of special-field dictionaries.

This third possibility can be found in the so-called "Allbuch" in which language lexicography and subject lexicography combine.⁸ Wiegand (1988: 747) argues that this type of reference work forms a class of its own, distinct from language lexicography and subject lexicography. It is now possible to present a typology with three types of reference works, namely: (i) special-field language dictionaries, (ii) special-field subject dictionaries, and (iii) "Allbücher" or encyclopedic dictionaries.⁹ Although these three types will not be discussed in detail, some important distinctions by Wiegand are however highlighted, because of their relevance for this article.

In order to make a finer distinction between linguistic and encyclopedic knowledge, Wiegand (1994: 114-115) explains how a medical student, Oskar, might consult a medical dictionary to find out what the expression *Strabometer* means. Oskar does not know the referential and predication rules for the expression *Strabometer*. This means that Oskar does not know how to use the expression, and therefore does not know the meaning of *Strabometer*. Oskar has a knowledge gap, which can be regarded as a special-field lexical knowledge gap ("fachlexikalische Wissenslücke"). This type of gap does not necessarily imply a special-field knowledge gap of the object ("fachliche Sachkenntnis-lücke") as well, but one may assume that this is the case. Wiegand (1994: 114) explains that according to the latest psycholinguistic perceptions, one may assume that adult native speakers always link the internal representation of

subject knowledge to the internal representation of an expression when an internal representation of the subject knowledge indeed exists. In simple terms this means that if one does not know to which object a particular expression refers, one cannot use the expression referentially or predicatively in a usual way.

Wiegand (1994) therefore not only distinguishes between semantic knowledge or nonencyclopediaic meaning knowledge ("nichtenzklopädisches Bedeutungswissen") and subject knowledge or special-field encyclopediaic subject knowledge ("fachenzklopädisches Sachwissen"), but also identifies a third type of knowledge. This he calls special-field encyclopediaic object-constituting meaning knowledge ("fachenzklopädisches gegenstandskonstitutives Bedeutungswissen"). It is object-constituting because it constitutes the referential object to which the expression refers and enables speakers to refer to the object in written or spoken texts. The following figure illustrates this:



According to Wiegand (1994: 116), this last type of knowledge can be described as knowledge about meaning which enables speakers to know how linguistically to deal with a particular referential lexical expression in a special field, that refers to some object. If one acquires this knowledge from a text as in Oskar's case, this knowledge has no specific visual component. This is why

Wiegand (1994: 116) adds that illustrations in special-field dictionaries are useful.

Wiegand (1994: 118) suggests that all items with meaning paraphrases in dictionary articles of special-field language dictionaries are simultaneously semantic and special-field encyclopedic data. Therefore, they are called semantic special-field encyclopedic data ("semantische-fachenzyklopädische Daten"). From this type of data a user can obtain the special-field object-constituting meaning knowledge belonging to the lemma sign.

In view of his characterisation of the object-constituting meaning knowledge, Wiegand (1994: 125) claims that one can now allocate sets of classes of items with the same genuine purpose to any of these three classes of data. Consequently, one is in the position to plan microstructure programmes for specific types of special-field dictionaries, in which one can treat the different types of dictionary articles in accordance with the dictionary type. Such microstructure programmes can be constructed by means of so-called frames, which Wiegand has researched in detail.¹⁰

3. The frame theory applied to lexicography

According to Wegner (1989: 893), frame theory developed in the domain of artificial intelligence. Its subject domain is the reconstructive explanation of cognitive processes by means of computer simulation. Programmes are developed which can produce utterances, answer questions with regard to the contents of texts, summarise texts or determine the theme of texts. Researchers in text linguistics and lexical semantics who developed the frame theory are, amongst others, C.J. Fillmore and M. Minsky. Konerding (1993: 21) refers to Minsky¹¹ who describes a frame as "a *data structure* for representing a stereotyped situation":

It is a *collection of questions* to be asked about a hypothetical situation. It can be viewed as an *organized matrix* of slots for given states of affairs, assigned with specific terminal values on the basis of the particular states of affairs to which the frame is applied.

Konerding's (1993) systematic and scientific study with regard to the use of frame theory in lexicography accounts for the representation of stereotypical everyday knowledge that dictionary users need when consulting dictionaries. Apart from Hellwig's¹² work on text theory, Konerding also applied Minsky's suggestions in order to establish the predicates which can be employed to describe nouns used in their usual senses. In addition, the study by Balmer und Brennenstuhl¹³ proved to be invaluable in the determination of predicate schemes.

Even though Konerding's examples, which are all nouns, come from general monolingual dictionaries, Konerding and Wiegand (1994: 138) claim that

the results apply to special-field dictionaries as well. They (Konerding and Wiegand 1994: 111-112) also state that, because of the scientific method used in determining the frames, learner's dictionaries can also use these frames to convey semantic knowledge of nouns in a systematic way.

Konerding worked out a typology and also tested this typology by means of computer research. The following types of nouns were identified for standard contemporary German and consequently used in the research: (i) concrete object, subclassified as natural object and artefact, (ii) organism, (iii) person/actor, (iv) event, (v) action/interaction/communication, (vi) institution/social group, (vii) (part of the) environment (of people), (viii) part/piece, (ix) whole/collection/entirety, and (x) condition/feature (of).

For each of these types of nouns, so-called matrix frames have been drawn up by Konerding and tested on several existing dictionary articles. These frames were constructed by means of several questions,¹⁴ which were grouped together according to themes. Konerding then also tested these frames on nominal phrases and different types of texts in which nouns were present. Konerding and Wiegand (1994: 106) claim that the results of Konerding's method concur with the latest results of cognitive psychology.

Out of the matrix frames, Konerding developed so-called minimal frames, which can be used in lexicography. This article focuses only on one of these frames, namely, the frame for artefacts, and applies it to a term from African music in par. 4.

According to Konerding and Wiegand (1994: 118-119), it is precisely the object-constituting meaning knowledge ("gegenstandskonstitutives Bedeutungswissen") that users need in order to be able to refer to objects in communicative situations, which can be determined and presented by means of frames. Konerding (1993: 405-410) lists several predicate classes which can be used to describe the object-constituting knowledge of nouns in the class artefacts. Not all of these, however, contribute to convey the object-constituting meaning knowledge to users, according to Konerding and Wiegand (1994: 121-122). Some of them contribute to convey the so-called special-field encyclopedic knowledge (Wiegand 1994: 117).

In par. 4 below, Konerding's predicate classes are adapted to suit the needs of the ethnomusicological database under discussion. Some of the classes have, for example, been grouped together for the sake of a logical flow of information. In addition, this list of predicate classes by Konerding was compared to De Vale's (1990) framework, which deals with all the aspects which she considers important when one wants to classify musical instruments within the field of organology. There is a remarkable resemblance. The predicate classes which, according to Konerding and Wiegand (1994: 121-122), contribute to the object-constituting meaning knowledge, are marked with an asterisk (*). The numbers 1-5 indicate the groups of questions Konerding (1993) used before establishing the predicate classes.

a. Constitutive relationships and features of the object:

1. predicates characterising the form, colour and measurements of the object (*);
 - predicates characterising additional discernible features of the object (*);
2. predicates characterising the hierarchically superordinate whole of which the object is an ingredient/part;
 - predicates characterising the essential parts of the object (*);
 - predicates characterising those objects which are used in the production phase of the object or parts of the object;
 - predicates characterising the arrangement, functional roles and special features of parts of the object;
 - predicates characterising the features of taking the object apart;
 - predicates characterising the applicability of the features of the object; and
 - predicates characterising other exceptional features (*).
- b. Phases of existence and distribution:
3. predicates characterising the conditions under which the object originates/the phases of production of the object;
 - predicates characterising the conditions under which the object originates/the phases of production and the later phases of existence of the object;
 - predicates characterising the distribution of the object; and
 - predicates characterising the availability of the object.
- c. The meaning of the object for people:
 - other names for the object;
4. predicates characterising activities in which the object functions/plays a role (*);
 - predicates characterising roles of the object in such activities (*);
 - predicates characterising the meaning of the object for the people using it;
 - predicates characterising other products of people which are used in the object;
5. predicates characterising similar objects, the differences between such objects and general categories within which the objects falls (*);
 - predicates characterising theories in which the object plays a role; and
 - predicates characterising information which has been published or made available about the object.

Wiegand (1992: 261-265) discusses three options for the presentation of predicate classes as part of a dictionary article.

3.1. The first option

Within the framework of Wiegand's lexicographical theory (Konerding and Wiegand 1994: 125), each of these predicate classes corresponds to a certain

item class ("Angabeklasse").¹⁵ In the examples Konerding and Wiegand (1994: 126) studied there are fewer item classes than predicate classes, because some predicate classes are grouped together. They (Konerding and Wiegand 1994: 139) suggest that ideally, every predicate class should receive its own lexicographical item class, especially in the case of, for example, learner's dictionaries. From a didactical point of view, the measure to which redundancy will occur in such cases, will actually be desirable.

For the noun class artefacts Konerding and Wiegand (1994: 127) recommend, also for the sake of computerising the dictionary, at least the item classes representing the following types of knowledge: (i) class of items concerning the knowledge of categories, (ii) class of items concerning the knowledge of functions, (iii) class of items concerning the knowledge of form and components, and (iv) class of items concerning the knowledge of other features.

3.2. The second option

If one wants to compile a learner's dictionary, it is recommended (Wiegand 1992: 263, and Konerding and Wiegand 1994: 137) that one formulate the parts of dictionary articles in which the object-constituting meaning knowledge appears, without too much text condensation.¹⁶ On the other hand, to obtain standardised condensed dictionary articles, one should formulate rules for lexicographers which prevent them from formulating too freely. These rules can also be based on frames.¹⁷

3.3. The third option

Regarding Wiegand's (1992: 264) interpretation, recent research in cognitive psychology and semiotics indicates that it might be useful to incorporate illustrations into the texts to facilitate the understanding of object-constituting knowledge by users. There are two ways of illustration: by means of language (illustrative examples) and by means of pictures or graphs.

3.4. Concluding remarks on the options

Wiegand (1992: 264) claims that one can formulate dictionary articles which contain items with a meaning paraphrase on the basis of frames. He points out that minimal frames do not distinguish clearly between factual encyclopedic meaning knowledge and object-constituting knowledge in all cases. He adds that this is not always necessary in dictionary articles which aim at conveying semantic knowledge.¹⁸ One should, however, always make sure which predicate classes should be taken into account when conveying object-constituting meaning knowledge. Wiegand (1992: 265) contends one could also systematically form competence examples¹⁹ which describe the use of the lemma sign as

well as further knowledge about the reference object on the basis of these predicate classes. In addition, one could produce texts in which the object-constituting meaning knowledge is systematically supplemented by factual encyclopedic meaning knowledge.

Frame-based dictionary articles have the advantage over traditional dictionary articles in that they are not only formulated on the basis of the individual lexicographer's language or subject competence, on his/her idiosyncratic ways of formulating and writing texts, and on relatively arbitrary descriptive traditions. Rather, the descriptive methods which are used can be regarded as objective and scientifically sound (Konerding and Wiegand 1994: 146).

4. Dictionary articles in the South African ethnomusicological database

In much the same manner as Kammerer's (1999: 232-236, 252-253) exposition of a frame-based dictionary article, a sample article from the ethnomusicological database, that of *uhadi*, is presented and discussed.

It is important to note that Bergenholtz, Tarp and Wiegand (1999: 1780) found encyclopedic data among various types of articles in existing special-field dictionaries. Sometimes a single article contains encyclopedic information, but single articles may also contain cross-references to so-called synoptic articles ("Synopseartikel"). The latter type of article may contain special-field encyclopedic subject knowledge and convey data which does not only refer to the lemma sign of the particular synoptic article, but which is also applicable to the lemma signs of the article(s) cross-referring to it. In addition, encyclopedic data can also be found in appendices such as an introduction to the special field, an overview of the special field, or even grammatical appendices, or illustrations of important items in the special field such as historical tables (Bergenholtz, Tarp and Wiegand 1999: 1787). The advantages of these appendices are that special-field issues can be dealt with in a coherent way and repetition of data in single articles can be avoided. A reference system to guide the user to these appendices is of course necessary.

This type of reference system is also built in in the database under discussion. It is self-evident that the predicate classes which account for the special-field object-constituting meaning knowledge of artefacts, as identified by Konerding and Wiegand (1994: 122), will have to be present in the article itself and not in the synoptic article or the appendices.

In the following table, the suggested information types of the article *uhadi* are listed. The predicate classes of Konerding's frame are presented in the left column, and the actual data of the database in the right column. The predicate classes necessary to convey object-constituting meaning, according to Konerding and Wiegand (1994: 121-122), are marked by an asterisk (*). Apart from the information from the predicate classes of the frame for artefacts, some information types such as items containing information of a grammatical or linguistic nature, are added.

PREDICATE CLASSES	DATA
Lemma sign	uhadi
Language of origin	Xhosa
Pronunciation	(In the electronic database, the pronunciation is supplied by a sound file.)
Translation equivalent	musical bow
Short everyday description in the sense of Wiegand (1992: 235)	The uhadi is a stringed instrument made of a wooden bow and a metal string, with a calabash resonator attached to the bow.
Other names for the object	Although there may be small differences, the instrument has the following names in different cultures: Tswana: <i>segwana</i> Tsonga: <i>dende</i> Northern Sotho: <i>sekgapa</i> Swazi: <i>ligubu</i> Zulu: <i>ugubu</i>
Etymological information	The Xhosa also use the word <i>i-gubu</i> for the calabash resonator (just like the Zulu), but the instrument is called <i>uhadi</i> , which may be compared with <i>umhadi</i> ("a deep pit").
(*) Form, colour and measurements (*) Essential parts of the object Conditions under which the object originates/phases of production and later phases Objects used in the production phase Arrangement, functional roles and special features of the parts of the object Features of dismantling the object Other products of people which are used in the object	<ol style="list-style-type: none"> 1. The bow of the uhadi used to be made of wood of the <i>umbhangandhlela</i>, a small tree with a yellowish flower. The Thembu Xhosa in the Lumko district use the branches of the <i>ulizi</i> tree. Nowadays any type of suitable wood is used. The shaft of the bow is called <i>injikwe</i> or <i>intonga</i>. 2. The string is about 100-115 cm long. The string used to be made of twisted oxsinew, or of cowhair or horsehair. Nowadays it is made of brass wire (<i>icingo</i>), generally taken from an old bracelet. The string is called <i>usinga</i> or sometimes <i>ijijo</i>. 3. The beater is a piece of <i>dobo</i> ("coarse grass") or <i>umcinga</i> ("wheat straw"). The beater is called <i>umcinga</i>. 4. The resonator is a small or medium size calabash. The resonator is called <i>iselwa</i>. 5. Between the calabash and the bow, a piece of cloth is wound. <p>(NOTE: The text is accompanied by a sketch of the uhadi and the animation</p>

	<p>makes it possible for the user to see how the bow is played and to hear what it sounds like. The different parts are also shown clearly.)</p>
<p>(*) Additional discernible features</p>	<p>To play the uhadi, it is held to the front of the player, with the opening of the calabash towards her. The string is struck with the beater. The overtones can be heard clearly. Then the calabash is pressed against the player's chest, and struck again. This time the higher overtones will be damped.</p> <p>The player alternatively plays the open string and then stops the string between her thumb and forefinger. This results in two <i>fundamental tones</i> which can be achieved, and therefore two sets of overtones are possible on the uhadi.</p> <p>(NOTE: Here a video clip is added showing a woman playing the uhadi. In addition, a button will be available where the user can go to a synoptic article on <i>polyphony</i> in Xhosa music. This article will explain the use of polyphony in Xhosa music, and can be accessed from other articles such as <i>umqangi</i>, a mouth bow, also using overtones and polyphony techniques.)</p>
<p>(*) Other exceptional features; Applicability of the features</p>	<p>When the uhadi is played, <i>overtones</i> can be heard. Some players play more than one rhythm at a time and some play the instrument while singing. Some even dance while they are singing and playing. Tuning of the instrument is important for the overtones to sound clearly.</p> <p>(NOTE: Here the term <i>overtones</i> will be a link which acts as cross-reference to the synoptic article <i>overtones</i>. In this synoptic article, theoretical aspects of the Xhosa use of overtones are discussed and notation examples are shown to illustrate the explanation. This information will also be accessed from other articles such as <i>umngqokolo</i>, a Xhosa singing style which makes use of overtone singing.)</p>

Distribution of the object Availability of the object	(Here a window opens to an appendix, which shows the distribution of the instrument on a map of South Africa.)
(*) Activities in which the object functions/plays a role (*) Role(s) of the object in such activities Meaning of the object for the people using it	The uhadi is usually played by women or girls, generally after the evening meal, rarely in daytime. It is used for accompanying the voice and the player usually sits while performing. It is preferable that the calabash comes in direct contact with the skin of the player in order for the overtones to sound more clearly.
(*) Similar objects/Classification of the object/Superordinate whole of which the object is a part	(NOTE: Here the user is referred to an appendix in which all Xhosa instruments are listed according to the Hornbostel-Sachs classification system to see the relationship with other stringed instruments within Xhosa culture, but also within an international classification system.)
Theories in which the object plays a role	
Information published about the object: citations and references	Bibliography: 1. Dargie, D. 1987. <i>Techniques of Xhosa Music</i> . Unpublished Ph.D. thesis. Grahamstown: Rhodes University. 2. Dargie, D. 1988. <i>Xhosa Music</i> . Cape Town: D. Philip. 3. Hansen, D.D. 1981. <i>The Music of the Xhosa-speaking People</i> . Unpublished Ph.D. thesis. Johannesburg: University of the Witwatersrand. 4. Kirby, P.R. 1968. <i>The Musical Instruments of the Native Races of South Africa</i> . Johannesburg: Wits University Press.

The following notes are important with regard to the article:

- (a) Cross-references to synoptic articles dealing with topics such as (i) the traditional ways of learning to perform, by observation and imitation, and (ii) the origin of all musical bows (hunting bows) will also be possible.
- (b) The database is aimed at high school pupils. It is therefore important to make use of nonstandardised articles and to use "everyday language" in the descriptions. Every technical term or foreign word that is used, will

eventually be accessible in its own right as a lemma so that the user may become acquainted with its meaning and use.

- (c) An attempt is made here to present the data in a linear way. Because of the hypertext nature of the database, the data types are of course accessible in different ways, from different angles. For example, users can access data about the *uhadi* by means of the list of topics, which would be an index in an alphabetical arrangement. They could also access the article by means of a path which runs through the instrument classification (the *uhadi* is part of the Xhosa stringed instruments). In addition, users could access *uhadi* through other articles on stringed instruments, through synoptic articles, such as the ones about the use of overtones, or polyphony, or appendices. Bergenholtz, Tarp and Wiegand (1999: 1788) claim that the user-friendliness of the lexicographical work is much better when users not only access appendices through single articles, but also access single articles through appendices. This can be done by adding buttons called "related articles" to the page of the appendix and listing the single articles there.
- (d) No attempt is made to present the dictionary article in a tree-structured graph as Wiegand usually does, because not all predicate classes have received lexicographical item classes yet. This still has to be done in further research.
- (e) This database is still in its initial stages. It is planned that once a sufficient number of articles has been completely worked out, the database will be tested and scrutinised by potential users for their commentary and feedback. This will be done by means of user protocols (Ripfel 1990, and Ripfel and Wiegand 1986).

In conclusion: In the ethnomusicological database an attempt is made to account for object-constituting meaning knowledge as well as semantic knowledge and special-field encyclopedic knowledge. Because Konerding has ascertained that his system of frames is viable, lexicographers will be wise to take into account these frames in the planning of new dictionaries for special purposes, learner's dictionaries, special-field dictionaries and special-field encyclopedic dictionaries. Much work is still to be done — lemmas of all the noun classes, as well as other types of word classes, will have to be worked on.

Notes

1. Cf. Smit (1997) for a general exposition of Wiegand's metalexical approach. The present article explores the application of frames to special-field lexicography in more detail.
2. Cf. Wiegand (1994: 122) for a description of this type of lexicographical work.

3. Languages for Special Purposes.
4. Wiegand (1984: 22) uses the so-called "theory of meaning-as-use", in which the meaning of an expression is determined by the "usual" rule for the use of that particular expression. Therefore, to know the meaning of an expression, is to know "how to act linguistically with it" (Wiegand 1984: 22). Wiegand's (1992: 237-243) well-known discussion of the word "Kayak" illustrates this issue.
5. Wiegand (1988: 744, note 23) illustrates this as follows: the genuine purpose of a pronunciation dictionary can be identified by its use to obtain information on pronunciation by means of the items in it. Consequently, the users fulfil the goals of the dictionary. A pronunciation dictionary has only one genuine purpose, and therefore belongs to the type of mono-informative dictionaries.
6. Wiegand (1988: 766) explains that linguistic items in a dictionary article are, for example, phonetic items, morphological items on inflection, items on gender, items on parts of speech, items on syllabification, items on verb valency, and many more.
7. According to Wiegand (1988: 767), encyclopedic items in a dictionary article convey information to the user about the object to which one refers by means of the lemma sign in a usual context, as well as knowledge about the social context within which the object is used. He (Wiegand 1988: 768, note 40) further notes that users can also often derive encyclopedic knowledge about objects from lexicographical examples.
8. It is difficult to find an English translation for "Allbuch". According to Wiegand, the French use an expression equivalent to "encyclopedic dictionary". For lack of a better term, this expression is also used in this article to refer to what Wiegand denotes as "Allbuch".
9. Also, cf. Wiegand (1998: 47-58) for a discussion of these three types of special-field dictionaries.
10. Cf. Wiegand (1985b: 86, 1989: 573) and Konerding and Wiegand (1994).
11. Cf., for example, Minsky, M. A Framework for Representing Knowledge. Haugeland, John (Ed.). 1997. *Mind Design II. Philosophy, Psychology, Artificial Intelligence*: 111-142. Cambridge MA / London.
12. Konerding's (1993: 147ff) approach is based, amongst others, on Hellwig's text theory, in which a set of questions can be used to determine the so-called "relevant silent knowledge" ("relevantes stillschweigendes Wissen") (Konerding and Wiegand 1994: 105) that can also be characterised as "object-constituting meaning knowledge" (Konerding and Wiegand 1994: 123). The works by Hellwig to which Wiegand and Konerding (1994: 105) refer are: (a) Hellwig, P. 1984a. Grundzüge einer Theorie des Textzusammenhangs. Rothkegel, A. und B. Sandig (Eds.). *Text — Textsorten — Semantik: Linguistische Modelle und maschinelle Verfahren*: 51-79. Hamburg, and (b) Hellwig, P. 1984b. Titulus oder Über den Zusammenhang von Titeln und Texten. *Zeitschrift für Germanistische Linguistik* 12: 1-20.
13. Balmer, T. and W. Brennenstuhl 1986. *Deutsche Verben: Eine sprachanalytische Untersuchung des deutschen Verbwortschatzes*. Tübingen.
14. Cf. Wiegand (1992: 257-258) for examples of these questions.
15. Cf. also Wiegand (1992: 253-267) for a detailed discussion on how one allocates item classes to specific predicate classes.
16. Wiegand (1977: 104-110) gives an excellent example of information which might be included for the word *Weberei* ("weaver's trade"). Historical facts about the weaver's trade could be

included as well as special vocabulary used in the trade, literature on the trade, etc. One of his ideas is that one could include "lexicographical narration", in which a systematic description could be given of the construction of a weaver's loom, together with an illustration.

17. Cf. Wiegand (1992: 263-264) for an example of this procedure.
18. Cf. also Konerding and Wiegand (1994: 138).
19. These would probably be illustrative examples to enhance linguistic competence.

Bibliography

- Bergenholtz, H., S. Tarp and H.E. Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Hoffmann, L., H. Kalverkämpfer and H.E. Wiegand (Eds.). *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research: 1762-1832*. Berlin: De Gruyter.
- De Vale, S. 1990. Organizing Organology. *Selected Reports in Ethnomusicology* 8: 1-34.
- Kammerer, M. 1999. Zur framebasierten lexikographischen Bedeutungsbeschreibung von substantivischen Lemmazeichen. *Lexicographica* 15: 229-263.
- Konerding, H.-P. 1993. *Frames und lexikalisches Bedeutungswissen. Untersuchungen zur linguistischen Grundlegung einer Frametheorie und zu ihrer Anwendung in der Lexikographie*. Tübingen: Max Niemeyer.
- Konerding, H.-P. und H.E. Wiegand. 1994. Framebasierte Wörterbuchartikel. Zur Systematisierung der lexikographischen Präsentation des Bedeutungswissens zu Substantiven. *Lexicographica* 10: 100-170.
- Pilegaard, M. 1994. Bilingual LSP Dictionaries User Benefit Correlates with Elaborateness of "Explanation". Schaefer, B. and H. Bergenholtz (Eds.): *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern: 211-228*. Tübingen: Narr.
- Ripfel, M. 1990. Wörterbuchbenutzung bei Muttersprachlern. Untersuchungsbericht über eine Befragung erwachsener muttersprachlicher Sprecher zur Wörterbuchbenutzung. *Lexicographica* 6: 237-251.
- Ripfel, M. and H.E. Wiegand. 1986. Wörterbuchbenutzungsforschung. Ein kritischer Bericht. *Germanistische Linguistik: Studien zur neuhochdeutschen Lexikographie* VI: 490-520.
- Smit, M. 1997. *Wiegand's Metalexigraphy as a Framework for a Multilingual, Multicultural, Explanatory Music Education Dictionary for South Africa*. Unpublished D.Litt. thesis. Stellenbosch: University of Stellenbosch.
- Wegner, I. 1989. Lexikographische Definition und Frame-Theorie im allgemeinen einsprachigen Wörterbuch. Hausmann, F.J., et al. (Eds.). 1989-1991. *Wörterbücher: Ein internationales Handbuch zur Lexikographie / Dictionaries: An International Encyclopedia of Lexicography / Dictionnaires: Encyclopédie internationale de lexicographie: 893-899*. Berlin / New York: Walter de Gruyter.
- Wiegand, H.E. 1977. Einige grundlegende semantisch-pragmatische Aspekte von Wörterbucheinträgen. Hyldgaard-Jensen, K. (Ed.). *Kopenhagener Beiträge zur germanistischen Linguistik: 59-149*. Copenhagen: Universitetsforlaget.

- Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Ed.). *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*: 13-30. Tübingen: Max Niemeyer.
- Wiegand, H.E. 1985a. Fragen zur Grammatik in Wörterbuchbenutzungsprotokollen. Ein Beitrag zur empirischen Erforschung der Benutzung einsprachiger Wörterbücher. Bergenholtz, H. und J. Mugdan: *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch, 28-30.6.1984*: 20-98. Tübingen: Max Niemeyer.
- Wiegand, H.E. 1985b. Eine neue Auffassung der sog. lexikographische Definition. K. Hyldgaard-Jensen and A. Zetterten (Eds.). *Symposium on Lexicography II. Proceedings of the Second International Symposium on Lexicography May 16-17, 1984 at the University of Copenhagen*: 15-100. Tübingen: Max Niemeyer.
- Wiegand, H.E. 1988. Was eigentlich ist Fachlexikographie? Munske, H.H., et al. (Eds.). *Deutsche Wortschatz: Lexikologische Studien*: 729-790. Berlin: De Gruyter.
- Wiegand, H.E. 1989. Die lexikographische Definition im allgemeinen einsprachigen Wörterbuch. Hausmann, F.J., et al. (Eds.). 1989-1991. *Wörterbücher: Ein internationales Handbuch zur Lexikographie / Dictionaries: An International Encyclopedia of Lexicography / Dictionnaires: Encyclopédie internationale de lexicographie*: 530-588. Berlin / New York: Walter de Gruyter.
- Wiegand, H.E. 1992. Elements of a Theory towards a so-called Lexicographic Definition. *Lexicographica* 8: 175-289.
- Wiegand, H.E. 1994. Zur Unterscheidung von semantischen und enzyklopädischen Daten in Fachwörterbüchern. Schaefer, B. and H. Bergenholtz (Eds.). *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 103-132. Tübingen: Narr.
- Wiegand, H.E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband*. Berlin: De Gruyter.

Theoretical Challenges to Practical Specialised Lexicography*

Sven Tarp, *Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark*

Abstract: During the last decades, part of lexicography has developed into an independent science with an elaborated system of theories based on the users' needs and corresponding dictionary functions. This also applies to specialised lexicography. Many lexicographers, however, still refuse to view their discipline as an independent science and treat it as a subdiscipline, i.e. of linguistics, or they consider specialised lexicography as something different from lexicography in general (which is the case with part of the so-called terminographers). Both interpretations have unfortunate consequences for lexicographic theory and practice. The great challenge now is to strengthen lexicography as a science, and to apply its system of scientific theories to the practical planning and compilation of specialised dictionaries. A look at existing dictionaries shows that there is still a lot of work to be done. Due to economic, editorial and other limitations, a great deal of pragmatism is often practised when compiling a concrete specialised dictionary. Lexicographic pragmatism might be necessary, but in order to be successful it must be guided by theory.

Keywords: COMMUNICATION-ORIENTATED FUNCTION, CULTURAL INFORMATION, DICTIONARY, ENCYCLOPAEDIA, ENCYCLOPAEDIC INFORMATION, GENUINE PURPOSE OF A DICTIONARY, GRAMMATICAL INFORMATION, KNOWLEDGE-ORIENTATED FUNCTION, LEXICOGRAPHIC FUNCTION, LEXICOGRAPHIC PRAGMATISM, LEXICOGRAPHIC SKILLS, LEXICOGRAPHY, LINGUISTIC SKILLS, LSP DICTIONARY, LSP LEXICOGRAPHY, PRACTICAL LEXICOGRAPHY, SEMANTIC INFORMATION, SPECIALISED DICTIONARY, SPECIALISED LEXICOGRAPHY, SUBJECT-FIELD SKILLS, TERMINOGRAPHY, TERMINOLOGY, THEORETICAL LEXICOGRAPHY, USER CHARACTERISTICS, USER NEEDS, USER SITUATIONS, USER-ORIENTATED LEXICOGRAPHY

Opsomming: Teoretiese uitdagings vir die praktiese gespesialiseerde leksikografie. Gedurende die laaste dekades het 'n deel van die leksikografie in 'n onafhanklike wetenskap ontwikkel met 'n uitvoerige stelsel van teorieë gebaseer op gebruikersbehoefes en korresponderende woordeboekfunksies. Dit is ook van toepassing op die gespesialiseerde leksikografie. Baie leksikograwe weier egter steeds om hulle dissipline as 'n onafhanklike wetenskap te sien en hanteer dit as 'n subdissipline, d.w.s. van die linguistiek, of hulle beskou gespesialiseerde leksikografie as iets anders as algemene leksikografie (wat die geval is met 'n deel van die sogenaamde terminograwe). Albei interpretasies het ongelukkige gevolge vir die leksikografiese teorie en prak-

* This paper was presented as a keynote address at the Fifth International Conference of the African Association for Lexicography, held at the University of Stellenbosch, Stellenbosch, 3-5 July 2000.

tyk. Die groot uitdaging is nou om die leksikografie as 'n wetenskap te konsolideer, en om sy stelsel van wetenskaplike teorieë op die praktiese beplanning en samestelling van gespesialiseerde woordeboeke toe te pas. 'n Kykie na bestaande woordeboeke wys dat daar nog baie werk is wat gedoen moet word. Weens ekonomiese, redaksionele en ander beperkings word 'n groot mate van pragmatisme toegepas wanneer 'n konkrete gespesialiseerde woordeboek saamgestel word. Leksikografiese pragmatisme mag nodig wees, maar om suksesvol te wees moet dit deur teorie gelei word.

Slutelwoorde: KOMMUNIKASIE-GEÏORIËNTEERDE FUNKSIE, KULTURELE INLIGTING, WOORDEBOEK, ENSIKLOPEDIË, ENSIKLOPEDIËSE INLIGTING, WARE DOEL VAN 'N WOORDEBOEK, GRAMMATIESE INLIGTING, KENNIS-GEÏORIËNTEERDE FUNKSIE, LEKSIKOGRAFIESE FUNKSIE, LEKSIKOGRAFIESE PRAGMATISME, LEKSIKOGRAFIESE VAARDIGHEDE, LEKSIKOGRAFIE, TAALKUNDIGE VAARDIGHEDE, TSD-WOORDEBOEK, TSD-LEKSIKOGRAFIE, PRAKTIËSE LEKSIKOGRAFIE, SEMANTIESE INLIGTING, GESPELIALISEERDE WOORDEBOEK, GESPELIALISEERDE LEKSIKOGRAFIE, VAKGEBIEDVAARDIGHEDE, TERMINOGRAFIE, TERMINOLOGIE, TEORETIESE LEKSIKOGRAFIE, GEBRUIKERSEIENSAPPE, GEBRUIKERSBEHOEFTE, GEBRUIKERSITUASIES, GEBRUIKER-GEÏORIËNTEERDE LEKSIKOGRAFIE

0. Introduction

Is lexicography a subdiscipline of linguistics? No, it is not. Although I would have preferred that many dictionaries, due to their poor quality, were considered products of linguistics and not lexicography.

Is terminography different from lexicography? No, it is not. Although I would have preferred that it was so due to the poor quality of a large number of terminographical products.

It is, of course, not my intention to offend anybody personally. In fact, I respect the efforts made by lexicographers or terminographers who produced the works which I criticise. But at the same time, I believe that it is my duty as a lexicographer who considers his discipline to be a scientific discipline, to be frank. The truth is that a large number of dictionaries, and especially LSP dictionaries as they are unfortunately called, do not live up to the standards of modern lexicographic works. And the truth is also that one of the reasons for this situation, if not the main reason, is that lexicography is not viewed as an independent science with its own independent system of lexicographic theories and its own independent lexicographic practice. This is the purpose of my paper, i.e. to argue that lexicography in general and specialised lexicography in particular, inclusive of terminography, must be viewed as an independent science. Only on this scientific basis will it be possible to raise the standard of lexicographic works immediately and produce the large number of high quality dictionaries needed in the present era of globalisation.

1. Lexicography and linguistics

Although the first dictionaries produced in China and India several thousand years ago were specialised dictionaries, during the last centuries it has mainly been academics with a linguistic background who have made the so-called LGP dictionaries. This is quite understandable because there is a growing need of a description and standardisation of the national languages and the creation of tools to facilitate bilingual communication.

However, this situation has had a number of negative consequences for lexicographic theory and practice. Many linguists have transferred their own theoretical instrumentarium to lexicography. In this way, they have been considering lexicography to be a subdiscipline of linguistics (see figure 1). This has meant that their analysis and studies of lexicography genetically were unable to go right to the root of lexicography and to discover its very essence. So all their studies of lexicography, which from the point of view of linguistics may be very scientific, leave much to be desired from the point of view of modern lexicography and are, as such, mostly very superficial.

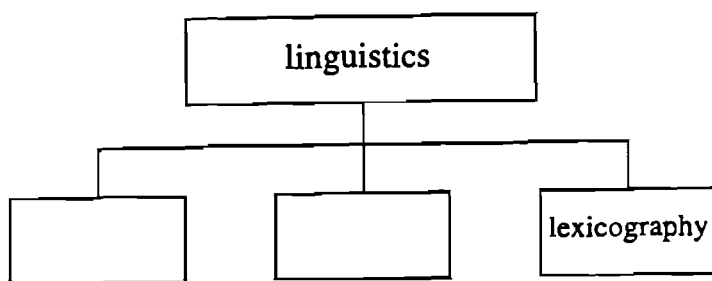


Figure 1: The linguistic approach to lexicography: Lexicography as a subdiscipline of linguistics

But is it not correct to view lexicography as a subdiscipline of linguistics? Is it not correct that dictionaries — as Landau (1984) once said — are about words, which are one of the basic subjects of linguistic study, and that dictionaries and lexicography must thus be considered a subdiscipline of linguistics? No, it is not. The following article from a specialised dictionary will clearly show this:

gastrulation During early embryogenesis, the invagination and reshaping of the cells of the blastoderm that results in differentiation into ectoderm, endoderm and mesoderm.

Text example 1: Dictionary article from Glick (1997)

Here we have an article from a dictionary that is conceived for people who have interest in learning more about biochemistry and molecular biology. Of course, the entry is a word and the explanation — or definition if you wish — is composed of words. But if this should be sufficient to make this dictionary a subproduct of linguistics, then every science in this world would be a subdiscipline of linguistics, because all sciences have to use words and language in order to express themselves. And this is absurd. In fact, if you can speak, write and spell English correctly, no special linguistic knowledge is needed to conceive and produce the above-mentioned dictionary. On the contrary, a profound knowledge of biochemistry and molecular biology is absolutely necessary to write the quoted explanation. So, if you should follow the logic of those who are normally claiming that lexicography is a subdiscipline of linguistics, in the particular case of the above-mentioned dictionary, this dictionary should be considered a subproduct of biochemistry and molecular biology. And the logical conclusion would be that at least this part of lexicography, i.e. biochemical and molecular biological specialised lexicography, should be viewed as a subdiscipline of biochemistry and molecular biology. And this is also absurd, because then lexicography would be something very schizophrenic: at the same time a subdiscipline of a large number of different and independent scientific disciplines (see figure 2).

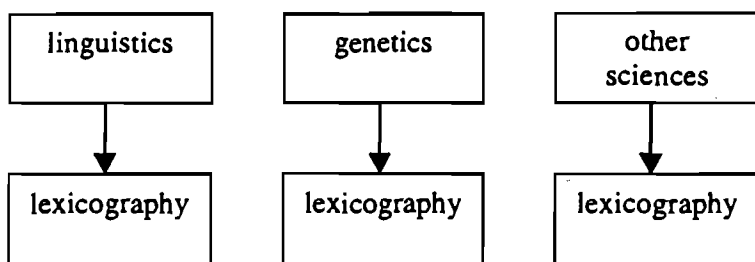


Figure 2: The schizophrenic model of lexicography: Lexicography viewed as a subdiscipline of linguistics, genetics, etc.

I think that the growing interest in specialised lexicography is making it quite clear that you cannot consider lexicography to be a subdiscipline of linguistics as many so-called LGP lexicographers have done due to their own linguistic background. The only solution is to view lexicography as an independent science, just like linguistics, genetics, biochemistry, molecular biology, etc. All of them are constituted by an independent set of knowledge elements, but at the same time they include a number of subsets of knowledge elements which they have in common with other disciplines.

In this way, linguistics and lexicography must be viewed as two independent disciplines where linguists sometimes — but only sometimes — make

use of lexicography for their work and where lexicographers sometimes make use of linguistics for their theoretical and practical work. This relation is shown schematically in figure 3:

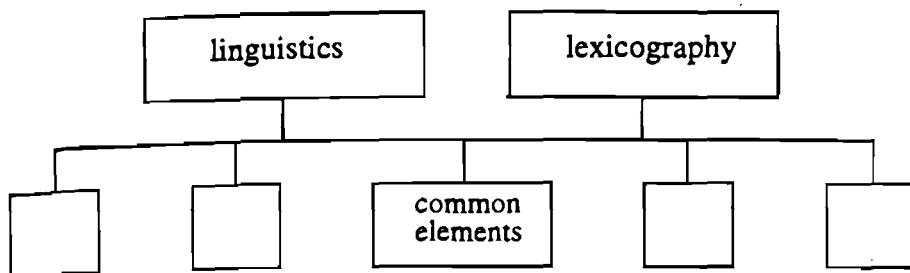


Figure 3: The modern lexicographic concept of lexicography: Linguistics and lexicography as independent disciplines with a subset of common elements

2. An independent science

Every independent science has its own basis and subject field: the subject of linguistics is language, the subject of biology is life, the subject of genetics is genes and their constituents, etc. What then is the subject of lexicography? It is dictionaries, or to be more specific: lexicographic reference works. It is of course true that these works need knowledge about language, biology and genetics in order to fulfil their objectives. But this only shows that in its very essence, lexicography is an open science with a great interdisciplinary vocation.

Genes and life are preconditions for the existence of human beings. And language is a property without which man would not be man. Lexicographic reference works, however, are human-made products. And as any other product of this sort, they are produced in order to satisfy certain human needs. That is why Wiegand (1988) called lexicographic reference works utility products that are produced for a specific purpose. In this connection, Wiegand introduces the concept of the "genuine purpose" of dictionaries in opposition to their "nongenuine purpose" — for example to kill the neighbour's cat with a heavy *Webster*. The genuine purpose of a lexicographic reference work is, according to Wiegand, "that it can be used to obtain information from its lexicographic data about the respective subject of the reference work (daß sie benutzt werden können, um aus ihren lexikographischen Daten Informationen über den jeweiligen Gegenstand der Nachschlagewerk zu gewinnen)".

This definition is, in my opinion, very precise, although it is (and of course must be) of a very general character. Wiegand (1988) himself tries to make a

more detailed typology of different kinds of dictionaries based on a distinction between semantics and encyclopaedics, but more about that later on.

3. Lexicography and terminology

Terminography is normally viewed as a subdiscipline of terminology. What is the subject field of terminology? Although there has been given many definitions of terminology, I have not succeeded in finding any exact definition answering this question. However, a qualified suggestion would be the registration and organisation of knowledge in general. If this is the case, terminology clearly distinguishes itself from lexicography, although it is at the same time a fact that specialised lexicographers also use a great part of the methodology that many terminologists consider a speciality of their discipline. In this field, the two disciplines have a lot in common.

However, when it comes to terminography, i.e. the presentation of the products of terminology in the form of different sorts of reference works, this is clearly the domain of lexicography. Whether or not we are dealing with the products of genetics, linguistics or terminology, when we have to present these products in lexicographic reference works, we are entering the special subject field of lexicography. That is why I consider terminography a subdiscipline of lexicography. And that is also the reason why I find it quite acceptable to review and criticise terminographic works on the basis of the principles developed by specialised lexicography.

4. The users and their characteristics

When you consider dictionaries as utility products made with the purpose of satisfying certain human needs, then two concepts are essential for lexicography as an independent science: the *user* and his/her *needs*. When a lexicographer conceives, produces or reviews a dictionary, he/she must always determine the corresponding or expected group of users and their needs. If this is not done, and that is unfortunately the case with many dictionaries, then lexicographic work cannot be done on a scientific basis. A concrete group of users can be homogeneous or heterogeneous. But in any case you must determine the *characteristics of the users*, first of all their language competence:

- Which language is their mother tongue?
- At what level have they mastered their mother tongue?
- At what level have they mastered a foreign language?

The above competencies are relevant for all monolingual and bilingual dictionaries, inclusive of LGP dictionaries. For a number of these dictionaries it is also important to know of the users' general cultural and encyclopaedic knowledge:

- What is the level of their general cultural and encyclopaedic knowledge?

For specialised dictionaries you also have to know the users' competence in the respective subject field:

- At what level have they mastered the special subject field in question?

But this information is not enough. You must also know the users' LSP competence within the given subject field:

- At what level have they mastered the corresponding LSP in their mother tongue?
- At what level have they mastered the corresponding LSP in the foreign language?

All these competences have to do with what I call the basic or *primary characteristics* of the user group. But in order to produce a high quality dictionary that is user-friendly, the lexicographer must also know the users' general experience of dictionary use, which can be called their *secondary characteristics*.

5. User situations

However, these characteristics of the user group constitute only one aspect of what should be taken into account when conceiving a modern dictionary. You must also determine or know in which *situations* the user is going to use the dictionary and with what *purpose*.

Basically, there are two different kinds of general user situations. Firstly, the user consults the dictionary in order to facilitate an existing or future communication. Secondly, the user consults the dictionary in order to obtain knowledge about a special subject, or in very rare cases, in order to learn and study a foreign language. These two types of user situations can respectively be called communication-orientated and knowledge-orientated (Tarp 1998).

6. User needs

When a user consults a dictionary, this is done in order to achieve information that allows him/her to solve a concrete problem. The dictionary must be able to meet the needs that arise in such a situation. Often the needs are very simple and can be covered with only one or a few lexicographic data. In other situations the needs are very complex and can only be met with a combination of different sorts of lexicographic data. The user might need:

- information about the native language,

- information about a foreign language,
- a comparison between the native and a foreign language,
- information about culture and the world in general,
- information about the special subject field,
- a comparison between the subject field in the native and foreign culture,
- information about the native LSP,
- information about the foreign LSP, and
- a comparison between the native and foreign LSP.

All these needs are primary user needs, because they are the needs that give birth to the consultation of the dictionary. There are, however, also other kinds of lexicographic user needs, which can be called the secondary user needs. These are the needs that arise when a dictionary is used:

- general information about lexicography and dictionary use, and
- information about the specific dictionary and how to use it.

In order to meet all these primary and secondary user needs, the lexicographer, when planning a dictionary, must know both the primary and the secondary characteristics of the user group. It is, of course, not the same to meet the LGP needs of a seven-year-old child who is still learning his/her mother tongue and the needs of an adult. It is not the same to meet the needs of a novice and a specialist within a given subject field. It is not the same to meet the needs of an experienced dictionary user and a person who is just consulting a dictionary for the first time.

7. Lexicographic functions

When you know the group of users, the users' characteristics, the user situations and the needs of the users, then you can determine the so-called *lexicographic functions* of the dictionary. A lexicographic function can be defined as "the endeavour and ability of the dictionary to cover the complex of needs that arise in the user in a particular user situation" (Tarp 1998).

The functions of a dictionary can be subdivided into communication-orientated and knowledge-orientated in correspondence with the respective user situations. The most important communication-orientated functions are:

- to assist the reception of texts in the native language,
- to assist the production of texts in the native language,
- to assist the reception of texts in a foreign language,
- to assist the production of texts in a foreign language,
- to assist the translation of texts from the native language into a foreign language, and

- to assist the translation of texts from a foreign language into the native language.

The knowledge-orientated functions are:

- to provide general cultural and encyclopaedic information,
- to provide special information about the subject field or the discipline, and
- to provide information about the language (e.g. when studying a foreign language).

Of course, all this must be worked out in a more detailed way when you deal with a concrete dictionary. For instance, if the dictionary in question is treating a subject that has developed in a different way from country to country or from culture to culture, then the dictionary must also include functions in order to give the user information on the subject in both the user's own country or culture and in the foreign country or culture. This could for example be the case with legal dictionaries, where the legal systems differ very much from country to country, or from culture to culture. This leads to the following more detailed functions that are certainly very relevant for a great number of specialised dictionaries:

- to provide information about the legal system in the native country,
- to provide information about the legal system in the foreign country, and
- to compare the legal systems in the native and foreign countries.

Here I would like to add a comment. When we talk about specialised dictionaries, we of course expect that the users have a high native-LGP competence. You can also expect that a translator with a linguistic background has reached a high foreign-language competence (if that is not the case, he/she should not be working as a translator). But if the dictionary is also conceived to serve experts in the respective subject field, you cannot necessarily expect such a foreign-LGP competence. Hence, in order to meet the needs of such users, you have to give more detailed general information about the foreign language. However, when it comes to LSP, many translators will most certainly lack the corresponding competence both in their mother tongue and the foreign language. And neither can the expert in the subject field be expected to have this LSP competence in both languages. There are many cases where experts do not have such a foreign-LSP competence. But there are also cases where they do not have it in their native language. That is for instance the case with many Danish geneticists who have only mastered the corresponding LSP in English. So you have to include more detailed information about the LSP in both languages in any case in order to satisfy the needs of both the translator and the

subject-field expert. And when it comes to knowledge of the subject in question, then the translator is generally not expected to have this knowledge, part of which is necessary in order to produce a qualified translation. But neither will a semi-expert, e.g. a student of the subject field, have sufficient knowledge of the subject field in question, and he/she might use a dictionary in order to obtain this knowledge. And even the expert, who is supposed to know everything about the subject field, will in many cases in real life have to consult a dictionary in order to confirm his knowledge or the correct use of a concrete term. Thus, in order to meet the different needs of these users, the dictionary must include the necessary special information about the subject field.

As I have already mentioned, all these considerations must be worked out in a much more detailed way when preparing a concrete dictionary dedicated to a concrete group of users. However, I hope that I have showed that the only way to reach a scientific conclusion of what should be included in a dictionary is to base this conclusion on an analysis of the user, the user's characteristics, the user situations, the user needs and the corresponding lexicographic functions.

I have already quoted Wiegand's definition of the "genuine purpose" of a dictionary considered as a utility product. But on the basis of the above considerations, I would like to suggest an alternative definition, which I believe is much more operational when planning a concrete dictionary. This definition of the "genuine purpose" of a dictionary is as follows:

The dictionary covers this or that area and is conceived to assist users with these or those characteristics in this or that situation in order to solve problems of this or that sort (Tarp 1998).

8. Priority and pragmatism

The functions are the very essence of lexicography. The functions constitute the leading principle of all dictionaries. Everything in a dictionary is to a greater or lesser extent influenced by its respective functions. Neither the contents nor the form of a dictionary can be conceived without taking the functions into account. It is possible to outline an integrated model of a lexicographic reference work on the basis of the functions.

For this reason, it is not recommended to have an abstract discussion on whether or not certain lexicographic data or items should be included in dictionaries, for instance information about culture in LGP dictionaries and about professional jargon in specialised dictionaries. Such a discussion must, in order to be scientific, always be related to the respective function or functions of the dictionaries in question. Should, for example, irregular flexion forms be included as independent entries in a monolingual English learners' dictionary, i.e. the imperfect form "went" of the verb "go"? In a dictionary for text reception, it certainly should be done, because you cannot expect the users, espe-

cially at beginners' level, to be able to relate this form to the entry "go". In a dictionary for text production, however, there would be no need to include "went" as an independent entry. In such a dictionary, it should figure as a morphological flexion form together with "gone" under the entry "go". On the other hand, should "went" and "gone" be included as flexion forms under the entry "go" in a learners' dictionary for text reception? Yes, this would be recommended as a means of identification that ensures the user that he/she has found the correct entry. This means that "went" and "gone" should be included as flexion forms under "go", both in learners' dictionaries for reception and for production, but for different reasons and with different purposes.

From the standpoint of user-orientated lexicographic theory, the best thing would always be to compile a dictionary with only one function and dedicated to only one type of user. In this way you will get the most accurate quality product designed to solve the special problems of a homogeneous group of users.

In practice, however, this is only possible in very few cases, normally when you are conceiving a LGP dictionary for a language that has a large number of speakers. And in specialised lexicography it is seldom the case. Thus, you have to combine various functions and a heterogeneous group of users. In order to do so, you have to analyse in detail what categories of lexicographic data are needed to cover each function for each different type of user. It is possible to combine all these categories in one and the same dictionary, but this frequently gives rise to problems and contradictions that make the dictionary less user-friendly, i.e. of lesser quality. And this is certainly not the objective of modern user-oriented lexicography.

Of course, as a means to solve this problem, you could omit certain data every time they create contradictions. But such a solution could also give rise to new problems, because it would result in a heterogeneous dictionary with a fluctuating quality where each type of user will sometimes be able to find what he/she is looking for and sometimes not. Another solution, which I prefer, and which is less pragmatic, would be to establish a ranking of functions and user types, giving first priority to some of them, second priority to others and third priority to still others. That means at least that you are sure that you are making a homogeneous quality product that meets the functions and serves the user types that you regard as most important for this particular dictionary. For the second and third categories of functions and user types, the dictionary may not be perfect, but it provides at least some kind of assistance to the users.

It is, however, not always possible to avoid pragmatism. Often you have to reduce the requirements of a planned dictionary due to financial limitations or the policy of the publishing house. But when you are pragmatic, it must not be in a pragmatic way. Pragmatism might be necessary in concrete cases, but it must be guided by theory in order to get the optimal results within the given limits. It is quite acceptable to say that it is not possible to meet all the demands of a linguistic translator (for instance, collocations) but that he/she is better off

with a dictionary which at least provides equivalents than with no dictionary at all. But this should not lead to "lexicographic" products, as this article taken from an English-Danish specialised dictionary:

analyst analytiker, konsulent, planlægger

Text example 2: Dictionary article from L&H (1992)

A native English-speaking translator who does not know anything about the subject field in question and the corresponding Danish LSP, would not know which of the three equivalents he/she should choose. Instead of helping the user, the dictionary is creating new problems that the user must solve before coming to a correct translation. In fact, the article is more dangerous than helpful, as the user has a 66 percent risk of choosing the wrong word. And these words are not at all synonymous: translated into English the three Danish words mean *analyst* (analytiker), *adviser* or *consultant* (konsulent) and *planner* (planlægger) respectively. The only native English speakers who could use this article would be those that already know the three Danish words and just need a reminder.

9. What skills are needed?

What skills are needed to conceive and produce a specialised dictionary? And what skills are needed to review it? These are basic and serious questions that have to do with the very quality of modern specialised lexicography.

Although lexicography is conceived as an independent science, it has, as already mentioned, a broad interdisciplinary vocation. Thus, a combination of several skills are often needed in order to produce a dictionary. And that is even more true when we are speaking about specialised dictionaries. In fact, I know only of one type of dictionary that can be produced with only one skill. That is a dictionary of lexicography where the lexicographer is a specialist in both dictionary making and the subject field in question. All other kinds of dictionaries need a combination of several skills. But what are these skills?

First of all, lexicographic skills. As I have argued in this paper, you cannot successfully engage in lexicographic work if you do it on the basis of other sciences or disciplines. You must have a specialised knowledge of lexicography. The role of the lexicographer is to conceive, plan, guide, direct and supervise the lexicographic project. If there is no person with special lexicographic skills, then the dictionary will not achieve the needed quality and will also often take much longer to finish, as the methods chosen would not always be the most suitable.

A second kind of skill needed in many lexicographic projects is linguistic knowledge. That applies to both LGP dictionaries and many specialised dic-

tionaries where you have to give general information on language. In bilingual dictionaries, linguistic skills are needed in both languages.

A third kind of skill is general cultural and encyclopaedic knowledge. This is the case both with general knowledge-orientated dictionaries and with some general communication-orientated dictionaries where at least some general cultural and encyclopaedic information is necessary in order to assist the user.

When dealing with specialised dictionaries, a fourth category of skills is necessary. That is knowledge about the subject field in question. And if the dictionary is communication-orientated, then a fifth category is also needed, i.e. knowledge about the LSP in question in all the languages covered by the dictionary.

All these skills can very seldom be found in one and the same person. The lexicographer often has both linguistic and lexicographic skills. But even that is not always the case, especially where bilingual dictionaries are concerned. And when it comes to knowledge about the subject field, it is very seldom that the lexicographer also has that knowledge.

But what exactly is the role of the subject-field expert? For which concrete tasks is he/she needed? The expert is needed to make a classification or systematisation of the subject field in question, to select the texts for the dictionary corpus, to select entries, equivalents and collocations and to write definitions, explanations and a special introduction to the subject field, if any. But all these tasks cannot be accomplished by the expert alone, but must be carried out in close co-operation with and under the professional guidance of the lexicographer.

So the conclusion is that a lexicographic project must from the very start be seen as co-operation between various specialists within very different fields of knowledge. This requires a capacity for planning and co-operation with other specialists. If the lexicographic project is not conceived in this way, then the first step towards an inferior product has already been taken.

And the same applies to the review of a specialised lexicographic reference work. A review or criticism of a concrete dictionary cannot be done only by a lexicographer, a linguist or an expert, because then the review would easily be one-sided and would not reflect the many-sidedness of the dictionary. Only a lexicographer can judge the lexicographic quality of a given dictionary, only a linguist the linguistic quality and only a subject-field expert the corresponding subject-field quality. This means that to do a professional and qualified review of a given dictionary, this must also be undertaken co-operatively by specialists with different skills.

10. Review

What is the present situation concerning specialised dictionaries? The answer is very simple: there are pathetic, bad, average, good and excellent dictionaries, but the general situation is that the lexicographic quality leaves much to be de-

sired. I will choose just two examples in order to illustrate what I mean. The first is an article from an English-German/German-English dictionary of business, commerce and finance:

bank 1. *n* (*bk*) BANK Bank *f*, Bankinstitut *nt*, Geldinstitut *nt*, Kreditinstitut *nt*; 2. *vt* BANK money einzahlen, zur Bank bringen; 3. ♦*vi* ~ with BANK ein Konto underhalten bei

Text example 3: Dictionary article from Dictionary of Business (1997)

This is a very typical dictionary article in a specialised dictionary. It possesses a number of good qualities. We are informed about the word class (although indirectly), about the gender and about the transitivity or intransitivity of a verb. A collocation is also given. But the article also creates some serious problems. Unfortunately, we are not informed for whom the dictionary is conceived, but a qualified guess would be both native German- and English-speaking translators, among others. They would surely have needed more collocations in order to produce a correct translation, especially the native English-speaking ones. This is, however, not the main problem. The article provides three equivalents that are not synonymous, and the native English-speaking user would find it very difficult to choose the right one. But even worse, we are not informed about the big differences between the British, American or South African banking systems and the German one. Without this important information, the translator could in a concrete situation very easily produce a wrong and very misleading translation. In this particular case, it would have been very beneficial if the dictionary had been conceived and produced on the basis of an analysis of the users and their needs.

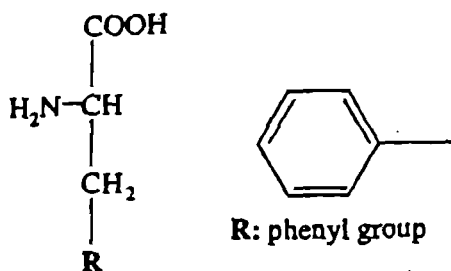
The second article on which I will comment, is from an English-Spanish dictionary of gene technology. This dictionary was conceived to serve as a tool for a wide range of persons who want to know more about gene technology and "whose background may be in medicine, technology, pharmacy, biology and/or journalism", and, in addition, was also produced to assist students of gene technology in their studies and to serve translators translating texts from the field of gene technology. The dictionary is the result of a co-operation between lexicographers, Danish and Cuban experts in gene technology and specialists in the corresponding English and Spanish LSP. Let us look at an article from this dictionary, for instance "amino acid" (see text example 4). This article, organised in paragraphs, contains an English entry and a Spanish equivalent, word class (although indirectly) and gender. It also includes an explanation together with an illustration in order to make it more understandable. Then there is a text example from an original document and a number of useful collocations. Finally, there are references to other entries and even to a 42 page systematic introduction to gene technology. In other articles in this dictionary there are also synonyms and antonyms. I think it is an excellent dic-

amino acid *aminoácido* m

Amino acids are organic compounds containing an amino group, $-NH_2$, and a carboxylic acid group, $-COOH$. Twenty different amino acids are the building blocks of peptides and proteins.

An average-sized protein consists of about 300 amino acids connected by peptide bonds. The possible number of amino acid sequences in a 300 amino acid protein is about 10^{90} . Thus the number of proteins that could be made is virtually unlimited.

Many important biomolecules are peptides, i.e. consists of a few amino acids.



● Moreover, we saw that in every one of the zipper and DNA-binding region were separated by exactly six amino acids.

▲ each ~ is encoded by *cada ~ está codificado por*; ~s are incorporated into *los ~s están incorporados a*; be translated into an *~ ser traducido a un ~*

→ amine; amino acid abbreviation; carboxylic acid group; § 2; § 12

Text example 4: Dictionary article from Kaufmann, Bergenholtz et al. (1998)

tionary that really meets the different needs of a very heterogeneous group of users.

11. An irrelevant discussion

At the beginning of this paper, I stressed that the understanding of lexicography as a subdiscipline of linguistics has had a number of negative consequences for theoretical and practical lexicography. There has been a pro-

nounced tendency among many linguists who have practised lexicography at a theoretical or practical level to transfer their linguistic theories, principles and methods to lexicography. Of course, this is mainly a problem in LGP lexicography, which is not the topic of this paper. However, there is at least one discussion that must be considered to be foreign to specialised lexicography. That is the ongoing discussion on the difference between semantic and encyclopaedic items in dictionaries, e.g. in the definitions or explanations. There is a long tradition of discussion on this problem. And it has mainly been influenced by linguistics and, to a lesser degree, philosophy.

Haiman (1980) takes a philosophic point of departure and declares that dictionaries are encyclopaedias, thus making no distinction whatsoever between semantic and encyclopaedic knowledge. On the contrary, Landau (1984) makes such a distinction and declares that "dictionaries are about words and encyclopaedias about things". In this way, he excludes encyclopaedias from the world of dictionaries, but unfortunately he does not tell us how this distinction between words and things is to be made. The answer to this question can be found in Wiegand (1988 and 1994) who in a very detailed way argues how it is possible to distinguish between semantic and encyclopaedic items in lexicographic definitions. Rossenbeck (1994) doubts whether it is possible to make such a clear distinction as Wiegand proposes and Bergenholtz (1994 and 1998) declares outright that it is impossible.

I do not have the intention to participate in this discussion as a lexicographer. It might be interesting from the point of view of philosophy or linguistics whether or not a distinction can be made between semantic and encyclopaedic knowledge and whether or not such a distinction can be made in a very clear way. But from the point of view of modern lexicography, it is, at best, completely irrelevant. I say "at best", because a lot of people wasted a lot of time participating in this discussion instead of developing other aspects of lexicography. And I say "irrelevant", because from the point of view of user-orientated lexicography, the important concern is whether or not the definition or explanation meets the user's needs and not whether it can be analysed in different parts.

12. A strange method

The understanding of terminography as something different from specialised lexicography and exclusively related to terminology also creates big problems for the users. There is a tendency among a number of terminographers to transfer the principles and methods of terminology to lexicographic reference works. This makes them very user-unfriendly indeed. One such example is the Biotechnology Glossary that was published by the Commission of the European Communities in 1990 and prepared according to the principles of terminology. The Commission is one of the worldwide institutions that spends a considerable amount of money to produce lexicographic reference works.

An introduction to the Biotechnology Glossary was written and signed by the Translation Service, Directorate for General and Linguistic Matters, Terminology Unit, in Brussels. In this introduction, the purpose of the glossary is clearly explained:

This glossary was originally conceived as an aid to translators faced with technical texts relating to biotechnology. The aim was to cover, with limited resources, a very wide field. (Biotechnology Glossary 1990: 1)

So the function of the glossary was the translation of biotechnological texts between the Community's official languages. And the user group was constituted as translators who are usually recruited among linguists or experts in various subject fields (but not necessarily biotechnology). Let us look at one article, or whatever you would call it, in order to study the result:

- EN linearizing of DNA**
 RF Enzymology vol. 100, p. 79; Olivier-Ward (DF)
 DF Introduction of a single double-strand break in a covalently closed circular DNA molecule and so converting it to a linear molecule. The cleavage of a unique restriction site within a plasmid will linearize it.
- FR linéarisation de l'ADN**
 RF Biotechnologies, p. 182
- DE Linearisierung**
 RF Gene und Klone, p. 317
- IT linearizzazione del DNA**
 RF Biotec 5/86, p. 42
- NL linearisatie van DNA**
 RF Biotechnologie Teleac, p. 14
- DA linearisering af DNA**
 RF Kursus i gensplejsning, p. 29 + TAI
- ES linearización del ADN**
 RF A. Aguilar
- PT linearização do ADN(1); retilinização do ADN(2)**
 RF div.PT(1,2)

Text example 5: Article from Biotechnology Glossary (1990)

What catches the eye first of all is that there is no grammatical information. There is also no information on gender, morphology or collocations. This type of information, however, is essential when making a translation within a LSP unknown to the user. Of course, the shortage of collocations can be explained and even excused by the "limited resources" referred to, although the information on gender and morphology would have been relatively easy and cheap to

produce. What cannot be excused, however, is that there is only a definition in one of the languages, i.e. English. In other cases there is again only a definition in French. And in still other cases there are definitions in several languages. In this way, the usefulness of the dictionary is automatically limited.

This very strange principle is explained in the above-mentioned introduction:

Whilst the main purpose of this glossary was to provide equivalence in the Community's official languages, it was in some instances felt preferable to leave a concept without equivalent in a particular language rather than put forward a translation without original source. Explanations of the concepts used were also regarded as essential. Rather than attempt to establish definitions which would then have been translated into the various languages, we opted to include helpful examples (codes DF + NT) of the use of the terms taken from original documents, thereby enabling them to be considered from various standpoints. We hope that the results will prove useful to both translators and specialists. (Biotechnology Glossary 1990: 2)

This is, indeed, some kind of lexicographic masochism.

First of all, the translators who do not find any collocations or other grammatical information in their target language, and not even an equivalent, will hardly consider the glossary to be helpful and useful, as it does not meet their special needs in these cases.

Secondly, it might be a principle of part of terminology not to write definitions and instead go and look for them in original texts. This principle might be useful when the terminologists want to establish a knowledge database, but it does not serve any lexicographic purpose. The specialists referred to need a clear definition of the term and the translators, in many cases, also need a definition or explanation in order to produce a correct translation. But this dictionary does not help them, because when no "helpful examples" are found in "original documents", then it does not provide any explanations of the concepts, although they are "regarded as essential".

I do not want to discuss the principles of terminology. It is not my field. And I am perfectly aware that there are different opinions about this problem among terminologists. But I do not find it very wise to use these principles when you are working on a reference work, although it is disguised as a terminographical and not a lexicographic reference work. It is not the principles of an alien discipline that should decide how you make your reference work and what you include in it. This is only one example of the negative consequences of viewing terminography as something different from lexicography and then transferring the principles of terminology to a subject field where they do not belong.

13. Conclusion

Lexicography is living very well — at least as a practice. But in a certain sense, lexicography is suffering from a severe disease. The problem is purely genetic. Lexicography has shared genes with linguistics, philosophy, terminology, etc. Therefore, the offspring has often been hybrids that cannot survive in the long run — and especially not in a modern world that increasingly demands quality. In order to survive, lexicography must opt for a user-orientated perspective. This is the only way to put it onto its own feet as an independent science and to ensure a new generation of lexicographic utility products that meet the needs of the modern dictionary user.

Literature

- Bergenholtz, Henning. 1994. Faglige oplysninger i monolingvale betydningordbøger. *LexicoNordica* 1: 13-26.
- Bergenholtz, Henning. 1998. Das Schlaue Buch. Vermittlung von Informationen für textbezogene und textunabhängige Fragestellungen. Zettersten, A., J.E. Mogensen and V.H. Pedersen (Eds.). *Symposium on Lexicography VIII. Proceedings of the Eighth International Symposium on Lexicography May 2-5, 1996 at the University of Copenhagen*: 93-110. Tübingen: Max Niemeyer.
- Bergenholtz, Henning, Jette Pedersen and Sven Tarp. 1994. Basic Issues in LSP Lexicography. Bergenholtz, H., A.L. Jakobsen, B. Maegaard, H. Mørk and P. Skyum-Nielsen (Eds.). *Translating LSP Texts. OFT Symposium. Copenhagen Business School 11-12 April 1994*: 151-187. Copenhagen: OFT-projektet, Copenhagen Business School.
- Bergenholtz, Henning and Sven Tarp. 1995. *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Amsterdam: Benjamins.
- Biotechnology Glossary (1990) = *Biotechnology Glossary*. Commission of the European Community. London/New York: Elsevier Applied Science.
- Dictionary of Business (1997) = *German Dictionary of Business, Commerce and Finance. Wörterbuch für Wirtschaft, Handel und Finanzen English*. London/New York: Routledge.
- Glick, David M. 1997. *Glossary of Biochemistry and Molecular Biology*. Revised Edition. London/Miami: Portland Press.
- Haiman, John. 1980. Dictionaries and Encyclopaedias. *Lingua* 50: 329-357.
- Kaufmann, Bergenholtz et al. (1998) = Kaufmann, Uffe and Henning Bergenholtz in cooperation with B. Stumman, S. Tarp, L. de la Rosa Marabet, N. la Serna Torres and G. la Serna Miranda. 1998. *Encyclopedic Dictionary of Gene Technology. English-Spanish*. Toronto: Lugus.
- Landau, Sidney. 1984. *Dictionaries: The Art and Craft of Lexicography*. New York: Scribner Press.
- L&H (1992) = *L&H fagordbog, engelsk-dansk*. Copenhagen: L&H Fagordbøger.
- Rossenbeck, Klaus. 1994. Enzyklopädische Information im zweisprachigen Fachwörterbuch. Bergenholtz, Henning and Buchard Schaefer (Eds.). *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 133-159. Tübingen: Gunter Narr.
- Tarp, Sven. 1992. Enzyklopædiske oplysninger i tosprogede tekniske ordbøger. Ruth Vatvedt Fjeld (Ed.). *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.-31. mai 1991*: 522-531. Oslo: Nordisk forening for leksikografi.

- Tarp, Sven.** 1994. Funktionen in Fachwörterbüchern. Bergenholtz, Henning and Buchard Schaefer (Eds.). *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 229-246. Tübingen: Günter Narr.
- Tarp, Sven.** 1995. Wörterbuchfunktionen: Utopische und realistische Vorschläge für die bilinguale Lexikographie. Wiegand, H.E. (Ed.). *Studien zur zweisprachigen Lexikographie mit Deutsch II*: 17-51. Hildesheim/New York: Olms.
- Tarp, Sven.** 1998. Leksikografien på egne ben. Fordelingsstrukturer og byggedele i et brugerorienteret perspektiv. *Hermes* 21: 121-137.
- Wiegand, Herbert Ernst.** 1988. Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachlichem und enzyklopädischem Wissen. Munske, H.H., et al. (Eds.). *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern*: 729-790. Berlin/New York: Walter de Gruyter.
- Wiegand, Herbert Ernst.** 1994. Zur Unterscheidung von semantischen und enzyklopädischen Daten in Fachwörterbüchern. Bergenholtz, Henning and Buchard Schaefer (Eds.). *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 103-132. Tübingen: Gunter Narr.

Challenges Encountered in the Compilation of an Advanced Shona Dictionary

Moreblessings Busi Chitauro-Mawema, *Institutt for Nordistikk og Litteraturvitenskap, Seksjon for Leksikografi, Oslo, Norway*

Abstract: This paper highlights the challenges encountered by the African Languages Lexical (ALLEX) Project (at present the African Languages Research Institute (ALRI)) in Harare, Zimbabwe, which is in the process of compiling an advanced Shona dictionary (ASD). Its forerunner is the general Shona dictionary, *Duramazwi ReChishona* (1996). The ASD is intended to be a comprehensive reference work, which will serve as a resource for more advanced users, especially those at higher secondary and tertiary education levels. The most important challenges have been in the areas of headword selection and the treatment of geographical/individual variation. The matters discussed here show the conflict between usage, i.e. popular acceptance, and (orthographic) norm, a problem often experienced in young literary languages subject to heavy foreign influence. This paper looks at: (a) the limitations of the current Shona orthography, the selection and codification of international vocabulary, and the presentation of variants and synonyms in the dictionary, and (b) the solutions suggested, and/or the ongoing debate on the topics.

Keywords: HEADWORD, COMPILATION, DICTIONARY, GENERAL DICTIONARY, ADVANCED DICTIONARY, INTERNATIONAL VOCABULARY, VARIANT, VARIATION, SYNONYM, CROSS-REFERENCE, IMPLICIT CROSS-REFERENCE, EXPLICIT CROSS-REFERENCE

Opsomming: Uitdagings teëgekome in die samestelling van 'n gevorderde Sjona woordeboek. Hierdie referaat laat lig val op die uitdagings teëgekome deur die African Languages Lexical (ALLEX) Project (tans die African Languages Research Institute (ALRI)) in Harare, Zimbabwe, wat besig is met die samestelling van 'n gevorderde Sjona woordeboek (ASD). Die voorloper daarvan is die algemene Sjona woordeboek *Duramazwi ReChiShona* (1996). Die ASD is bedoel om 'n omvattende naslaanwerk te wees wat as 'n hulpmiddel vir meer gevorderde gebruikers sal dien, veral dié in die hoërskool of op universiteit. Die belangrikste uitdagings was geleë in die lemmakeuse en die hantering van geografiese / individuele variasie. Die sake wat hier bespreek word toon die konflik tussen gebruik, d.w.s. populêre aanvaarding, en (ortografiese) norm, 'n probleem wat dikwels voorkom in jong geskrewe tale onderworpe aan sterk vreemde invloed. Hierdie referaat beskou: (a) die beperkings van die huidige Sjona ortografie, die seleksie en kodifisering van die internasionale woordeskat, en die aanbieding van wisselvorme en sinonieme in die woordeboek, en (b) die voorgestelde oplossing en/of die voortgaande debat oor hierdie onderwerpe.

Slutelwoorde: SOEKWOORD, KOMPILASIE, WOORDEBOEK, ALGEMENE WOORDEBOEK, GEVORDERDE WOORDEBOEK, INTERNASIONALE WOORDESKAT, VARIANT, VARIASIE, SINONIEM, KRUISVERWYSING, IMPLISIETE KRUISVERWYSING, EKSPLISIETE KRUISVERWYSING

1. Introduction

This paper looks at some aspects of the work of the African Languages Lexical (ALLEX) Project¹ (at present the African Languages Research Institute (ALRI)) which is in the process of compiling an advanced Shona dictionary (ASD). Its forerunner is the general Shona dictionary, *Duramazwi ReChishona (DRC)* (1996). The ASD is to be a comprehensive reference work for more advanced users, especially those at higher secondary and tertiary education levels. Faced with this huge task, the ALLEX team has had to deal with areas that were either only slightly explored or not explored at all in the *DRC*. Consequently, a number of challenges have come up which need to be addressed before the ASD is completed. These have been the headword selection and the presentation of variants and synonyms in the dictionary. This paper highlights some of the problems encountered, the solutions suggested, and/or the ongoing debate on the topics.

The questions touched on in this paper are not unique to Shona. They stress the conflict between usage and norm, a problem often experienced in young literary languages subject to heavy foreign influence. Much has already been written on theoretical debates and discussions on similar lexicographic questions in other languages (see e.g. Haugen 1976 and Sinclair 1987). In this paper I present the practical problems that the team faced and the solutions that were suggested, without getting too deeply involved in theoretical issues and debates implied by the choices made. I do, however, refer to and draw parallels with the Scandinavian experience between the 16th and the 20th centuries.

2. Headword selection

Headword selection has had implications for the current Shona orthography, as indicated below.

2.1 The problem

Working towards producing the ASD has shown that the absence of certain letters and/or digraphs, like *l*, *th* and *rh* in the standard Shona orthography, compared with what is found in the corpus and in everyday spoken Shona, points to a well-justified argument to include headwords containing such letters in the dictionary. Haugen (1976) reports that Scandinavian grammarians

battled with these same questions back in the 17th century. Commenting on the heated debates that ensued between traditionalists and radicals, Holberg (Epistel 415, cited in Haugen 1976: 395) says, "there are no wars as bitter as those fought by grammarians".

The question of orthography has been discussed in the general planning and training meetings and workshops and certain recommendations have been made which need to be pursued. The following options have been explored:

- (a) Avoid including headwords containing these letters and /or digraphs.
- (b) Systematically replace *l* with *r* in headwords as is customary with adopted words. It has worked for e.g. *bhora* (< ball), *firimu* (< film) and *bhurara* (name referring to a pesticide < Nguni *bulala*, to kill).
- (c) Delete *l* in adoptives as has occurred in *sauti* (< salt) and *musoja* (< soldier).
- (d) Continue to ignore words with *th* and *rh* as has been happening already or systematically replace the *th* with *t* and *rh* with *r*.
- (e) Include adopted headwords containing these letters with the letters intact, as is shown in the following indigenous and adopted examples in which *l*, *th* and *rh* occur both in the corpus and in everyday Shona:
 - (i) Shona dialect, e.g. *kudhla* (standard Shona *kudya*) (to eat) and *kutyla* (standard Shona *kutya*) (to fear),
 - (ii) Shona proper names, e.g. *Sithole* and *Mlambo*, and
 - (iii) adoptives, e.g. *kukala* (to colour), *kudhila* (to deal), *mudhila* (a dealer), *ndombolo* (a Zairean dance), *yelo* (yellow), *bhuluu* (blue), *losheni* (lotion), *kalenda* (calendar), *themomita* (thermometer), *bhethi* (birth certificate), *thiyeta* (theatre), *thiyori* (theory), *rhumba* (type of music) and *rhumha* (a frightening thing).

A closer look at the occurrence of these items in the corpus leads to the realization that these words are now part of the everyday Shona language and that it would be unwise to continue to ignore their existence. Besides, there is a limit to which "foreign" letters can be replaced with local ones, as there is an aesthetic side to both written and spoken language that cannot be ignored. It should be pointed out also that dictionaries almost always seem to carry words that are not consistent with current orthography (see e.g. *Kristu* (< Christ) in *DRC*).

2.2 Decision taken

Our position, therefore, was that while the standard orthography needs to be respected, the corpus needs to be respected as well, as it should not happen that community norms prevent one from hearing about *themomita* (< thermometer) because its spelling has a *th* in it. The honest editor is true to the language as it is used. It may be that the team will be criticized for violating the

integrity of the language by introducing strange letters and/or digraphs to the language in what will be interpreted as a normative work. But, we also have to take other factors into account. Orthographic revisions have been made before and can still be made now. There have been three official orthographies for Shona so far, and we already know that dictionaries help with acceptability. It was therefore agreed that the following was to be done:

- (a) Include adopted headwords containing these letters with the letters intact (as *Kristu* was in *DRC*, even though it contains two consonant clusters which are prohibited by the standard orthography), and with etymological information included in the entry as well as an explanation in the front matter.
- (b) Present the problem to the speakers of the language. Explain that this is not a problem peculiar to Shona only but that other languages go through similar problems and describe how these communities handled the problem. The presentation could be in the form of a seminar or an advert in the newspapers that would solicit responses and suggestions.
- (c) Request the Shona Language Committee to recommend their use. One of the Committee's roles, among many others, are to examine and discuss language issues raised by users and make the necessary recommendations to the education ministry.

3. International vocabulary

From the scope of headwords selected, the ASD will be a different dictionary from *DRC*. This will also be achieved through including international vocabulary (IV) in the dictionary. IV often refers to technical words which carry specific, unchanging and unambiguous senses in the contexts in which they occur and are used internationally. Most of these words are encountered in scientific and technical subjects taught in schools, where part of the mastery of each discipline entails the mastery of the concepts in it. If the ASD is to be promoted, it must serve as a practical tool for teachers who want to improve teaching in Zimbabwe by teaching and writing textbooks and other materials in Shona. Teachers would eventually be teaching these subjects in at least Shona and Ndebele and would, therefore, need a source to consult.

3.1 The problem

While lexicography entails a codifying function, lexicographers do not like to perform this role because it forces them to be more prescriptive than they desire. How then were we to determine the headwords in the present circumstances? While Shona is undoubtedly a very rich language in other areas, it has had to borrow a lot on matters that "relate to material culture and technology" (Chimhundu 1983: 341). The questions that needed to be deliberated on were:

- (a) Should we continue to adopt and adapt the international headwords from English and other foreign languages?
- (b) Should we use indigenous headwords (where they exist) or should we coin new ones (where they do not exist)?
- (c) Should the headwords be restricted to those that are actually being used or should the team go beyond these, and if so, how far beyond?

In all these deliberations we had to remember that consistency is central to science and technology. While we would like things to be uncomplicated and restrict IV headwords to those that are actually used, as lexicographers we still have the responsibility to show the scientific and technical professions that indigenous languages can be used in these circumstances.

Again, the Scandinavian developments of the 17th and 18th centuries shed light on this challenge. Haugen (1976) writes that the concern of lexicographers and grammarians at that time was for the "purity" of their languages as they battled with foreign influence and the question of whether or not to coin new words, to revive dying words or to use foreign ones. While the general policy in the rest of Scandinavia was to promote indigenous words over adoptives as much as possible, Iceland's position was an outright "no loans from other languages" (Thorlaksson 1612).

3.2 Decision taken

We agreed that where good indigenous words existed for the technical/ scientific concepts that needed naming, it would be best to promote those words over adoptives. Shona already has a good tradition of naming in which it observes the behaviour of a thing/concept and then gives it a name that qualifies that observation. Our feeling was that this tradition should continue through new coinages, which would include new Shona words as well as semantic extensions of already existing ones. With coinages we had to face the possibility that speakers might not use them and would opt to stick to adopted forms with which they were more familiar. For this reason, we would only coin new words that we believed were likely to be easily accepted.

We made these decisions fully cognisant of arguments emphasizing the difficulties of using indigenous words and coinages. We were aware that the use of indigenous words and coinages could raise problems, since these were words that existed in the language and hence had meanings attached to them already. As a result, they could be ambiguous in some cases, e.g. *denderedzi-gama* (whose meaning we have extended to mean semicircle in geometry) which can be broken down as:

denderedzi-	+ -gama
(circle)	(not full)

"Not full" does not necessarily mean "half full". This means that ambiguity may arise where the same word may be used to mean e.g. a quarter or half full. Another illustration comes from the terms for "High Court" and "Supreme Court" which have been developed in Shona as *dare guru* (lit. big court) and *dare guru repamusoro* (lit. big court of top) respectively. However, it can be difficult to differentiate their respective uses when both courts are sometimes referred to just as *dare guru*. Further ambiguity results as any higher court may also be called *dare guru*. Parliament is called *dare guru (reParamende)*, when *dare* now refers to an assembly. Some users feel that the introduction of adoptives from English could diminish this potential for ambiguity. For example, *koti* (< court) is regularly used and accepted by some as part of the Shona vocabulary. Also, a number of IV words are used in Shona speech, perhaps through code-mixing and code-switching, but they do not appear in writing, e.g. the adoptives *mainasi/mainazi* (< minus) which are still very often used, even though *-bvisa/-bisa* (minus/subtract) has been well-integrated (see Appendix A for more examples). Drawing on the Scandinavian example again will show that attempts to replace the Graeco-Latin terminology with indigenous equivalents succeeded only in creating occasional synonyms for the Graeco-Latin terms. Translators tried with varying degrees of success to find equivalents, many of which became standardised while the others survived as synonyms. The team therefore decided to enter the parallel terms, so that the matter can resolve itself over time, when one term might gain favour above the other.

Our argument for the use of these indigenous terms is that IV words and technical terms do not just establish themselves. They come from an authoritative source. The dictionary should set the norm by conventionalizing the use of some of these IV terms. An example of a country that seems to have succeeded in implementing the use of indigenous terms is Iceland, where all school subjects are taught in Icelandic. The language policy there is to preserve the structure of Icelandic, thus eliminating any foreign influence. Government supports the structure-preserving policy strongly by funding full-time terminographers and providing facilities to ensure that the work is done successfully. Everyone has been sensitized to the policy, and they all strive to actively contribute to the development of Icelandic terminology. Ambiguity, however, continues to be a problem.

On the other hand, the argument in favour of adopting terms is that even those languages that are now said to be developed, have also borrowed terminology from others. For example, nearly all the morphemes used in IV words are Greek or Latin, though Shona speakers have encountered them through English. Consequently, in naming some of these scientific concepts we had to concede that while it would be preferable to use our own indigenous words, consistency across the world in the sciences is also needed. Therefore, adoptives would be selected where no indigenous words existed or where a coinage would run the risk of being rejected completely.

Appendix A shows a list of some of the IV that has been proposed for the ASD. The list covers, among others, biological, chemical, and mathematical terminology. The list of language and literature terminology has been relatively easy to compile because the Government has decreed that Shona and Ndebele as medium for the teaching of these languages.

4. Regional/dialect variation

Another challenge concerns the need to capture regional/dialect variation fully, and to present this variation in the dictionary in a manner that accommodates all speakers from the Shona community. A rich standard language should present choices to its speakers regarding dialects, constructions and grammatical forms. But, how are these to be presented in a dictionary? The team decided that commonly used variants would be entered as the main headword, i.e. carry the definition, while the less commonly used variant would be cited and cross-referenced to the main variant (see *DRC*), and also "from the form that is marked as regards register to the unmarked form and so on" (Svensen 1993: 195). The same would apply to synonyms.

Cross-referencing would be carried out with a functional view of saving space. In this way one and the same concept would be described sufficiently under only one of the forms.

The other function of cross-references would be to guide the user to the correct place in the dictionary where the meaning and examples can be found. Such presentation would, for instance, be helpful in a case where a user may be faced with an irregular inflexion and have difficulty in identifying the headword.

The third function of cross-references is to prevent information from being lost. In a case where a certain form of the headword may have been separated and given its own entry, e.g. for editorial reasons, cross-referencing would help to link the two.

4.1 The problem

These functions give us two types of cross-references that Svensen (1993) refers to as explicit and implicit cross-references. Explicit cross-references are expressed as an instruction as in *ONA* (see) and *TARISA* (cf.) in the ASD. This type of cross-referencing means the word is treated elsewhere from the place where it is found. The user has to obey the instruction in order to obtain information or to find additional material. The implicit cross-reference is usually part of a statement about the word, e.g. "pl." or "abbr. of".

Sometimes the difference in form may be so great that the cross-referenced words are synonyms. Cross-referencing may also be to a particular sense only, rather than to the whole entry. Contractions and abbreviations may also be

treated as synonyms and cross-referenced to one another. In contractions and abbreviations the complete form will function as an explanation.

In the ASD we decided to use mainly the explicit type of cross-referencing.

Cross-referencing in the dictionary will, however, be complicated by the complexity of the Shona-speaking community where variation carries regional political connotations. The term Shona which came in use after Doke's publication of his report on the unification of the Shona dialects in 1931, is a collective term for a group of five main, mutually intelligible varieties. These are Zezuru, Karanga, Manyika, Ndaou and Korekore. Standard Shona is mainly made up of Zezuru and Karanga, in that order. Doke's report (Chimhundu 1983) confirmed the zones/tribes (sic) established by missionaries, and made the Zezuru dialect, the dialect of the capital, the norm for phonemic analysis. As a result, Zezuru gained more prominence than the other dialects in usage and acceptance. The ministerial directive of 1982 which allowed dialectical variation in writing, tried to reverse this state of affairs. "Controlled flexibility" was now to be allowed in spelling, thereby allowing dialectal variation. However, the fear expressed by some is that, despite the ministerial directive, some varieties may in practice continue to dominate at the expense of the others. Cross-referencing means that one commonly used form, presumably from the more dominant varieties, will be the main entry, while the less commonly used, most likely from the less dominant varieties, will continue to be considered of less weight. This means that varieties like Zezuru and Karanga will continue to enjoy the prestige that they have always enjoyed, compared to the other varieties like Manyika, Ndaou and Korekore. According to language experts² who have been in the teaching field for some time, Zezuru continues to enjoy the prestige that it has had since Doke (1931). They observe that firstly, Korekore is more or less Zezuru, and secondly, Manyika will tend to borrow more from Zezuru than Karanga, its nearest rival, because Zezuru has prestige, because of its historic association with the capital city, and also because in some ways it is easier to understand than Karanga.

4.2 Decision taken

While all these other factors may prevail, we as a team are still convinced that to save space, the explicit type of cross-referencing is the best. Besides, our dictionary is based on a fairly representative corpus in terms of regional coverage. Hence, where there is doubt, the selection shall be made on the strength of the corpus evidence after all other factors have been considered (see the *weti/mutundo* illustration below for more clarification).

The ASD presents four types of variation, and this is how they are treated (see also sample entries in Appendix B):

(Statistics indicating the frequencies of occurrences were extracted from the ALLEX Shona corpus of over 2 million running words. The form marked with an asterisk * is the main headword, the more commonly used term.)

4.2.1 Variants

(a) Regional variations

There is a fair amount of variation among dialects, both with respect to the stem-changing derivational processes and phonological alternation, e.g.

<i>*nzeve</i> has	353
<i>zheve</i>	19
<i>*tsuro</i>	316
<i>shuro</i>	24
<i>*-rwara</i>	864
<i>-gwara</i>	less than 74

Vowel alternations in prefixes like *na-/ne-/no*; *sa-/se-/so*; *-a-/e-/o-*; as well as in several of the verbal extensions like *-is-/es-*; *-ir-/er-*; *-w-/iw-/ew-* would also be treated systematically in ASD as variants, e.g.

<i>na-</i>	19 709
<i>*ne-</i>	41 831
<i>no-</i>	14 960

(b) Inflected forms

Irregular inflected forms are treated as main entries and cross-referenced to the regular uninflected form, e.g.

<i>guru</i> (big — cl. 5 nouns)	459
<i>huru</i> (big — cl. 9 and 10 nouns)	519
<i>*-kuru</i> (big — adjectival stem for the rest of the nouns)	8 019

(c) Adopted words

Alternate spellings and forms of adopted words may also result from individual/regional pronunciations of the adopted words, e.g.

bhazi/bhasi/bhesi (< bus)
shati/sheti (< shirt)

Besides the individual/regional variations, the older generation also seems to pronounce these adopted words differently from the younger generation. The older generation's pronunciation is more faithful to the spelling while the younger generation's is more faithful to the pronunciation of that word by the

English-speaking community, cf. *shati/sheti*. This situation seems to result from the way the language was acquired by these two generations. At the time the older generation acquired the language, they had more access to the printed than the spoken word, hence they pronounced words according to spelling, like the Shona. The younger generation, on the other hand, has had access to both. Furthermore, there was the Roman Catholic influence which pronounced /a/ in *man, the, bank*, etc. All these factors have led to a multiplicity of terms that will need to be provided for in the dictionary. Again, where there is doubt, corpus evidence will determine which form is to be the main entry.

Sometimes variants of adoptives result from plural and singular forms of the original word, e.g.

kiti/kitsi/katsi (< cat(s))

bhani/bhanzi (< bun(s))

Sometimes variant forms are indirect adoptives, e.g.

soseji/sochisi (< English sausage) / *sochisi* (< Nguni *isositshi*)

4.2.2 Synonyms

Again, in the interest of saving space, synonyms should be defined only when it is absolutely necessary. Otherwise a less commonly used form would be cross-referenced to the more commonly used one carrying the definition. Where in doubt, the strength of the corpus would also help determine the main headword, e.g.

-svikirwa (be possessed) 106

-sutswa 0

However, as we discovered, the problem is not always easy to deal with, especially in cases where users feel strongly for their own form, e.g.

mhuka / mututu / muhotwe (nosebleeding)

In cases like this, the standing decision is to define each one and cross-reference to one another, space permitting.

The general trend throughout the dictionary would be to promote indigenous words as much as possible. What we observed, however, is that where an indigenous and adopted headword compete, the indigenous word may not always turn out to be the one more frequently used. In such cases we may disregard frequency of use so that the indigenous word or both the indigenous and adopted word will carry the definition, examples, etc., e.g.

<i>-dhiraiwaha</i> (< drive)	35
<i>-tekenya</i> (drive)	13
<i>mutundo</i> (urine)	0
<i>weti</i> (< wet, i.e. urine)	23

In the *mutundo/weti* example, we witness a common feature of the language. This is the use of euphemism for taboo words, especially those referring to excretive functions or sex. In this illustration the corpus has 23 occurrences of *weti* and 0 for *mutundo*, yet *mutundo* is the main word. This example illustrates why we have to be careful when taking frequencies of occurrence as the ultimate answer to our problem. Other cultural and social factors need to be considered as well.

Contractions and full forms will be treated as synonyms. Some dialects use the contracted form while others use the full form, e.g. *chipotswa/chipotserwa* (casting a spell).

Individual variations, e.g. *mwana wangu/ mwanangu* (my child), must also be taken into consideration.

5. Conclusion

This paper tried to highlight some challenges that face the ALLEX team as it works on its second monolingual Shona dictionary. The lesson that the team has learnt from the Scandinavian experience is that the development and cultivation of indigenous languages is a gradual task, which may take centuries. For the Scandinavians it was only after three to four centuries that the established languages started stabilizing. In the ALLEX team we are content, for now, in the knowledge that we have laid a foundation of codification and cultivation of Shona on which later generations can build.

Notes

1. The ALLEX Project has operated as a tripartite co-operative venture in monolingual lexicography under the UZ/NUFU Agreement between the University of Zimbabwe, the University of Oslo and the University of Gothenburg.
2. This observation is drawn from consultations with Korekore and Manyika colleagues.

References

- ALLEX Project. 1999. *Draft Report of the 7th ALLEX Planning and Training Workshop held at the University of Zimbabwe, Harare, in February 1999.*
- Chimhundu, H. (Ed.). 1983. *Adoption and Adaptation.* D. Phil. thesis. Harare: University of Zimbabwe.
- Chimhundu, H. (Ed.). 1996. *Duramazwi ReChiShona.* Harare: College Press.

- Doke, C.M. 1931. *Report on the Unification of the Shona Dialects* carried out under the Auspices of the Government of Southern Rhodesia and the Carnegie Corporation. Hertford: Stephen Austin.
- Gronvik, O. 1999. *Strategies for dealing with International Vocabulary*. Presentation at the 7th ALLEX Planning and Training Workshop held at the University of Zimbabwe, Harare, in February 1999.
- Haugen, E. 1976. *The Scandinavian Languages: An Introduction to their History*. London: Faber and Faber.
- McEnery, T. and A. Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Sinclair, J.M. (Ed.). 1987. *Looking Up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London: Collins.
- Svensen, B. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making* (Translated from the Swedish by J. Sykes and K. Schofield). Oxford/New York: Oxford University Press.
- Thorlaksson, G. 1612. *Visnabok*.

Appendix A: International Vocabulary List

Indigenous IV (old, new coinages, semantic extensions)

nyaviri (allergy)
-pfudza nhumbu (abort)
-pfudzika pamuviri (miscarry)
-tadza pamuviri (abort) (uninduced)
-bva nhumbu (abort/miscarry)

denderedzi (circle)
pakatikati (centre)
-govanisa (divide)
-sanganisa (add)
-hwerengedza (add)
endereno (semantic extension)
enzane (equation)
mboni yedenderedzwa (centre of circle)
denderedzigama (semicircle)

-ongorora (examine/look critically at)

mumiriri wenyika (ambassador)

Adoptives

asidhi (< acid)
atomu (< atom)
metsi (< maths) (also popular even though there is a Shona term)
mainasi (< minus) (also popular even though Shona terminology exists)
mainazi (< minus) (also popular even though Shona terminology exists)
pulazi (< plus)
purasi (< plus)
-taimiza (< times/multiply)
-dhivaidha (< divide)
redhiyasi (< radius)
noti (< not/zero)
haidhirojeni (< hydrogen)
okisijeni (< oxygen)
dhayamita (< diameter)
dhinomineta (< denominator)
lita (< litre)
mililita (< millilitre)
mita (< metre)

milimita (< millimetre)

aleji (< allergy)

ikweta (< equator)

adhaputa (< adapter)

aluminiumu (< aluminium)

Appendix B: Cross-referencing

1. Variants

Regional variation

-gwara K itik. ONA *-rwara*.

-rwara [-*gwara*] K itik. Kana munhu achirwara muviri wake unenge uchirwadziwa nokuda kwedenda rinenge riri mumuviri make, kana kuti nokuda kwekukuvara kwaakaita.

nzeve [zheve] DK z 9>10. 1 *Nzeve* inhengo dzemuviri dzinowanikwa kumativi emusoro dzinoita kuti munhu anzwe. *Ane nzeve hombe dzakaramwa musoro.* 2 • Kana *nzeve yaro*va inenge yaputika mukati ichibuda urwa. 3 *Nzeve* rudzi rwehowa hunodyiwa, hwakatsvukuruka pamusoro asi hune danda jena.

zheve DK z 9>10. ONA *nzeve* 9.

Inflected forms

guru K- pr. ONA *-kuru*.

huru K- pr. Pava nenguva huru kubva zvatakawana kuzvitonga kuzere. ONA *-kuru*.

-kuru [*guru*, *huru*] K- pr. 1 Munhu mukuru kune mumwe ndiye anenge akatanga kuzvarwa, kana kuti ane makore akawanda pakuberekwa. FAN *-hombe*. 2 Chinhu chikuru ndechine muviri kana muturu muhombe, kana kuti chinozadza nzvimbo hombe, kana kuti chakawirira kuenda mude-ngadenga. FAN *-hombe*.

Adoptives

bhasi KD z 5. ONA *bhazi* 5.

bhazi [*bhasi*, *bhesi*] KD z 5. *Bhazi* zimota zihombe rine zvigaro zvakawanda zvinotakura vanhu vazhinji vari parwendo. FAN *dutavanhu* 5.

bhesi KD z 5. ONA *bhazi* 5.

2. Synonyms

Indigenous (regional variation)

mhuka DK z 9. 1 *Mhuka* chirwere chinokonzera kuti ropa ribude nomumhuno, kazhinji kana munhu achipiswa. FAN *mututu* 3, *muhotwe* 3. 2 *Mhuka* imota risingawanzorapika.

muhotwe [muhota] DDK z 3. *Muhotwe* chirwere chinoita kuti munhu abude ropa mumhuno, kazhinji kana kuchipisa. FAN *mututu* 3, *mhuka* 9.

mututu DKK z 3. 1 *Mututu* chirwere chinoita kuti munhu abude ropa mumhuno, kazhinji kana kuchipisa. FAN *mhuka* 9, *muhotwe* 3, *mutso* 3.

Adoptive vs indigenous word

-dhiraiuha D it. Kana uchi*dhiraiuha* motokari unenge uchiita kuti ifambe. FAN *-tyaira*, *-tekenya*. < Ru:drive.

-tekenya K it. Kana uchitekenya motokari unenge uchiita kuti ifambe. FAN *-tyaira*, *-dhiraiuha*.

Adoptive vs indigenous taboo word

mutundo D- z 3. *Mutundo* imvura inoburitswa nevanhu nezvimwewo zvipuka nekuzasi ichibva mudundira. FAN *weti* 9, *dope* 5.

weti D- z 9. *Weti* imvura inoburitswa nevanhu kana zvimwewo zvipuka nekuzasi ichibva mudundira. FAN *mutundo* 3.

Grammatical Terminology for the Teaching of African Languages at Tertiary Level*

C.T. Msimang, *Department of African Languages, University of South Africa, Pretoria, Republic of South Africa*

Abstract: The article addresses the question of the elaboration and modernisation of African languages to enable them to be used as language media for the teaching of African languages, particularly grammar. Whereas all institutions at tertiary level of education have undergone a paradigm shift from teaching African languages in English and Afrikaans to teaching them in their own medium, the necessary terms to ensure efficiency in this process are lacking. Scholars therefore need to focus on term creation and thereafter standardization of the terminologies. Some flaws in the creation of African terminologies in the past are highlighted and remedies suggested. It is recommended that statutory bodies such as the Council for Higher Education (CHE) or the South African Qualifications Authority (SAQA), or academic bodies such as the African Languages Association of Southern Africa (ALASA) could be entrusted with the task of the standardization of these terminologies.

Keywords: AFRICAN LANGUAGES, LANGUAGE MEDIUM, GRAMMATICAL TERMINOLOGY, TERM CREATION, STANDARDIZATION OF TERMINOLOGY, TEACHING AT TERTIARY LEVEL.

Opsomming: *Grammatiese terminologie vir die onderrig van Afrikatale op tersiêre vlak.* Die artikel behandel die vraag oor die uitbouing en modernisering van Afrikatale om hulle in staat te stel om as taalmedia gebruik te word in die onderrig van Afrikatale, veral taalkunde. Waar alle inrigtings op tersiêre opvoedkundige vlak 'n paradigmaterskuiwing ondergaan het vanaf die onderrig van Afrikatale in Engels en Afrikaans na die onderrig daarvan in hul eie medium, ontbreek die nodige terme om doeltreffendheid in hierdie proses te verseker. Geleerdes moet hulle daarom toespits op termskepping en daarna op die standaardisering van die terminologieë. Sommige gebreke in die skepping van die Afrikataalterminologieë in die verlede word belig en regstellings voorgestel. Daar word aanbeveel dat statutêre liggame soos die Raad op Hoër Onderwys (RHO) of die Suid-Afrikaanse Kwalifikasieowerheid (SAKO), of akademiese verenigings soos die Afrikatale-Vereniging van Suider-Afrika (AVSA) met die taak van die standaardisering van hierdie terminologieë toevertrou word.

Keywords: AFRIKATALE, TAALMEDIUM, GRAMMATIKATERMINOLOGIE, TERMSKEPPING, TERMINOLOGIESTANDAARDISERING, ONDERRIG OP TERSIÊRE VLAK.

* This paper was delivered at the ALASA Biennial Conference held at the University of South Africa, Pretoria, 6 July 1999.

The main purpose of this paper is to highlight certain problem areas in coining and developing grammatical terminologies in African languages. This problem is the result of a switch at tertiary education institutions from teaching the grammar of African languages in English or Afrikaans to teaching it through the medium of the African languages themselves. The paper will end up by suggesting ways and means of addressing the problem. As far as possible, examples will be drawn from the Zulu language.

Changing from an established tradition is not always easy. Nonetheless this particular change is welcome and is perhaps long overdue for a number of considerations. Firstly, this transformation is in line with the strong pedagogical view that the best medium in which to teach a person is his/her mother tongue. Therefore, for those students who learn an African language as their first language, the medium of teaching and learning should be that language itself. Secondly, in the South African education system the grammar of African languages at primary and secondary school level is taught in the medium of those languages. It subsequently becomes problematic for those students to change at tertiary level from an African to an English language medium. If that university student is a prospective teacher, he/she would, for instance, be learning Zulu in English to enable him/her to teach it better in Zulu — an anomalous situation.

While there is a consensus of opinion on the advantages of teaching the grammar of an African language in its own medium, there are also many limitations. Firstly, of course, there is the lack of relevant terminologies. Secondly, this approach restricts the number of potential scholars. Once African languages and literatures are in African language media, non-Africans have no access to such studies. The studies become too localised and are thus not exposed to international scholarship and critique. This also limits overseas study opportunities for such scholars. Thirdly, within the Bantu language family itself, it becomes difficult if not impossible to engage in comparative Bantu linguistics or grammatical typology. Correspondences between languages cannot be critically analysed without a bridge language. These, however, must be seen more as challenges than problems.

In an attempt to set the scene for coining grammatical terms in African languages, the approach of Louwrens (1994) will be used as a point of departure. Louwrens considers three broad areas in his approach. The first area comprises terms with an official status. These are the terms found in the fourth edition of the series *Terminology and Orthography*, 1988-1993, issued by the Department of Education and Training, Pretoria. Separate volumes were published on each of the languages then offered by the Department at different intervals. Whilst these are very important term banks, it should be borne in mind that they provide terminology only up to secondary level. Terms in English (or Afrikaans) introduced at tertiary level will not always have a counterpart in these African language lists. The second area comprises terms coined in a haphazard manner by grammarians who produce language manu-

als or other treatises in African language media. These terms are coined in order to fill in linguistic gaps where the official source does not provide the desired terms. Finally, there is an area where no terms are available.

The foregoing makes it very clear that language elaboration in African languages is far from adequate. Not only are there gaps, but even the available technical terms lack acceptability which is necessary for standardisation. Without elaboration and standardisation of African language terminologies the teaching of such languages in their own medium can never be effective. This is where official bodies and African language teachers should come in and address the defects and discrepancies. A few thoughts on such defects and discrepancies are apt at this juncture.

Perhaps one should start from the formal to the informal. This takes us back to the *Terminology and Orthography* series. The first edition of these appeared in 1957 under the auspices of the Department of Native Affairs. Only two publications were issued, one for Zulu and Xhosa and the other for the three Sotho languages. The second edition was published in 1962 under the auspices of the Department of Bantu Education. A separate publication appeared for each language, namely: Zulu, Xhosa, Tswana, Northern Sotho and Southern Sotho. The same Department issued the third edition of this series in 1976. Finally, the Department of Education and Training was responsible for the fourth edition between 1988 and 1993.

In her current research Van Huyssteen (1999) refers to criticism of this series of official terminology lists compiled by various Language Committees which later became Language Boards. Citing Cluver and other authorities, she (Van Huyssteen 1999: 2) states that critics "have questioned the legitimacy of terms listed" in these publications. Such critics "feel that terms in the African languages were coined artificially on legislative demand and were therefore imposed on the speakers". These critics concluded that this resulted in an "exclusive short-lived lexicon" which was unacceptable. They suggested that the creators use terms that already exist in society, thereby making them more acceptable to the users.

While this criticism contains some validity, it is generally short-sighted, biased and subjective. It does not seem to take cognisance of the fact that a statutory linguistic structure is essential for the standardisation of a language — in this case the Department of Education and Training and its Language Committees/Language Boards. The Department only provided enabling statutes and infrastructures. Once the Committees/Boards started functioning, they enjoyed a measure of autonomy and were not subjected to undue pressure from the government. They were staffed by outstanding educationists and social leaders with a recognised standing in the Black community. Although they were condemned by certain ideologists for serving in apartheid structures, their only aim was to help elaborate and standardise their languages which they loved, and they plied this business with enthusiasm, devotion and utmost

commitment. It is true that they made errors along the way, but their *bona fides* were always above board.

The process should not be dubbed illegitimate simply because of the intervention by the erstwhile government. Terminologies are accepted or rejected purely on their linguistic impact. Even pioneering lexicographers such as C.M. Doke and B.W. Vilakazi who compiled the *English-Zulu/Zulu-English Dictionary* before the birth of apartheid (Vilakazi actually died in 1947) furnished a number of unacceptable Zulu equivalents for certain foreign English concepts. A glaring example is their Zulu definition of culture, i.e. improvement or refinement of mind, manners, etc., by education and training (*Oxford Illustrated Dictionary* 1962: 196). They translated this term as: Ukukhanya okunjengokwabelungu (Light which resembles that of the white people). This definition is not only unacceptable, it is also misleading and insulting. It is misleading to claim that culture is equivalent to light. "Enlightenment" would have been a better choice. It is insulting to claim that only white people have a culture.

It is somewhat simplistic to demand that the creators of terms in the *Terminology and Orthography* lists should have consulted various speech communities and used terms that already existed in such communities. This shows ignorance of the fact that term creators were part of those Black speech communities and that they interacted with members of such communities using existing terms. This fact made consultation redundant.

Secondly, this charge does not consider the circumstances under which term creation often happens. A news reader over the radio or on the television can be cited as an example. This person receives the news bulletin in English or Afrikaans, one or two hours before broadcast time. He/She must translate and type out a Zulu, Xhosa, Venda, or Tsonga version and read it timeously. If he/she comes across a term that does not have an equivalent in available dictionaries and terminology lists, he/she must create one on the spot. When therefore will he/she find time to consult?

Thirdly, and perhaps more importantly, this charge fails to take cognisance of the fact that theories which are applied to analyse African data, including the grammars of African languages, are foreign to such data. These are usually imported from Europe or the United States of America. If concepts in such theories are introduced only at tertiary level of education, their counterparts in African languages will not be found in the official terminology lists which cater mainly for primary and secondary education level. In all probability, members of African speech communities at grass roots level shall not have been exposed to such concepts. Under the circumstances, one can hardly speak of consultation or use of existing terms. Reference has already been made to Louwrens' (1994: unpaginated Preface) categorisation of Northern Sotho terms for which counterparts were coined, English terms for which no Northern Sotho counterparts have been suggested and English terms for which Northern Sotho counterparts do exist. Against this scenario what chances are there for

the existence of popular counterparts in Northern Sotho, Southern Ndebele, Southern Sotho, Swati, Tsonga, Tswana, Venda, Xhosa or Zulu speech communities of the following grammatical concepts?

agglutinating languages
inflectional languages
isolating languages
transformational generative grammar (TGG)
lexical phonology
autosegmental phonology
pragmatics
discourse analysis
sociolinguistics
sociology of language
speech act theory
speech register
subcategorization
surface and deep structure meaning
synchronic linguistic study
diachronic linguistic study
stative meaning
stylistic meaning
cognitive meaning
connotative meaning
affective meaning
associative meaning
topicalisation
grammaticalisation
modality
isogloss
semantic bleaching
pronominalisation
zero morpheme
verbaliser
WH-Questions
category shifting
Sandhi rules
suppletion
anaphora
inchoative meaning

This is but a small sample of linguistic expressions. A more comprehensive linguistic term bank is to be found in David Crystal's book *An Encyclopaedic Dictionary of Language and Languages*. The 1992 edition of this collection comprises

2 750 entries in a 428 page volume. Yet even this figure is, by Crystal's own admission (on the inside of the cover jacket), not all embracing as he delineates his scope thus: "Linguistic here does not mean technical terminology of linguistic sciences, but language in a more everyday sense."

Obviously, African language counterparts of these and many other grammatical terms shall need to be coined through tutorials and workshops by university communities or by grammarians in their studies and libraries. This then takes us to the informal domain.

Louwrens (1994: unpaginated Preface) has already referred to the dangers of terminographers working in isolation. He states:

It should be emphasised that some of these "new" coinages are already unofficially employed particularly in school manuals which are used for mother-tongue instruction. This obviously results from the pressing need which is experienced by authors of such works for suitable official terms which cover the broad grammatical spectrum ... Unfortunately, grammarians who have in the past been involved in the coining of terms such as these have, in most cases, worked independently of one another, with the result that there is no uniformity and consistency with regard to either the terms which have been decided upon, or the way in which they are applied.

What Louwrens observed in Northern Sotho also occurs in Zulu, and causes confusion. A few examples of literary terms used by the Departments of African Languages at the University of Zululand and the University of South Africa illustrate this point.

English term	Zulu term	
	University of Zululand	University of South Africa
suspense	uheho	ilukuluku
anticlimax	ibohlo	upholavuthondaba
plot	uhlaka	isakhiwo

The moral lesson from this experience is that independent and individualistic term creation should be followed by standardisation.

The foregoing should not however blind us into thinking that those on Language Committees/Language Boards did not make mistakes. Although they were members of their communities, they did not always recognise the popular terms which had been created spontaneously by these communities and which could have facilitated their task and ensured ready acceptability of the term bank. Since most of these Committee/Board members were language

purists, they often frowned upon and ultimately rejected the borrowings used by people at grass roots level. This was a serious oversight which had two dire consequences for language elaboration. In the first place, this stultified language modernization. The surest way of killing a language is to maintain its pure form. Secondly, it led to some distortion of meaning where African language lexemes were used as neologisms (in extended or compounded forms) to express foreign concepts from an English or Afrikaans context.

However, grammarians are not unanimous on the significance of extending the meaning of lexical items so as to create technical terms. For instance, Louwrens (1994: unpaginated Preface) views this in a positive light with regard to Northern Sotho: "A more important reason why it was decided to create terms where Northern Sotho has no equivalents, is because it was felt that even if these new creations were not official, they put the novice who is a mother-tongue speaker of Northern Sotho in a better position to conceptualise the semantic content of the English counterpart." While there can be concurred with this view, it must be added that in Zulu at least, this theory has not always been upheld in practice. A few examples will help illustrate this point.

For many years certain Zulu sticklers for language purity refused to "bastardise" the Zulu language by accepting the adoptive *itafula* (from Afrikaans *tafel*) as an equivalent for *table*. Instead they imposed the word *ithala*. This completely distorted the meaning of the word *ithala*, since this piece of furniture is more of a shelf which is suspended from the roof at the back of a Zulu hut, and is used for storing objects, whereas the table is a piece of furniture with a flat top of wood, marble, etc. and vertical supports, used mainly for laying meals. Eventually common sense prevailed and *ithala* was discarded and *itafula* adopted in the modern Zulu term bank.

In grammar, however, there is quite a number of problematic terms. Take *indlela*, for instance. This word glosses as *path*, *way* or *road*, but in Zulu grammar it is used to denote *mood*. No stretch of the imagination can show compatibility between these two concepts. One needs to know the significance of *mood* in English first before one could accept *indlela* as its Zulu equivalent. Let us consider more examples from Kumalo (1987: 21-22):

English term	Zulu equivalent	Literal meaning of the equivalent
dissimilation	inhlukano	separateness
grammatical tone determiner	izwi lohlelo	word/voice of arrangement
	isilawuli	controller

Coinages created in this manner are normally self-explanatory if they refer to concrete objects, e.g. *umtholampilo* (clinic), *isiqedakoma* (cool drink) and *isiqedaphunga* (deodorant).

Coinage	Gloss
umtholampilo (clinic)	-thola = find; impilo = health i.e. a place where one's health is restored
isiqedakoma (cool drink)	-qeda = finish/quench; ukoma = thirst
isiqedaphunga (deodorant)	-qeda = finish; iphunga = (bad) smell

Where a coinage expresses an abstract notion, publicity especially through media could be the only way to facilitate acceptance of such a term. *Indlela* has been accepted as an equivalent of *mood*, for instance, owing to its wide publicity in the school environment.

It is pleasing to note that the Zulu Language Board gradually moved from a puritan standpoint and started accepting adoptives. In the 1993 edition of *Zulu Terminology and Orthography* one often comes across two Zulu equivalents for a single English term. The one equivalent is invariably an indigenous word and the other an adoptive from English or Afrikaans. Invariably the adoptive becomes more popular, which demonstrates the power of internationalisation over indigenisation:

English term	Zulu indigenous term	Zulu adoptive
refrigerator	isiqandisi	ifriji
radio	umsakazo	irediyo
script	ubhalo	isikriphthi
toy	isilibaziso sengane	ithoyi
television	umabonakude	ithelevishini
grammar	uhlelo	igrama
phonetics	ubhalojikelele	ifonethiki

Khumalo too has become a strong advocate of internationalisation in his later publications.

While the above adoptives give the impression that terms created by the people on the ground are borrowings, this is not necessarily the case. A cursory scanning of the lexical items created by people in road construction, for example, convinces one that these people are highly creative. For instance, they watch the movements of the caterpillar and then name it *isixaxa*, where the ideophone *xaxa* glosses as *short, staccato forward leap*. They also watch the heavy skidding of the rear wheels of a tractor and call this vehicle *ugandaganda*, where the ideophone *ganda* glosses as *heavy stampeding*. It is also interesting how the people on the ground combine borrowing and transliteration to arrive at new adoptives. Take the term *bathroom*, for example. The official equivalent thereof is *igumbi lokuhlambela* (*igumbi* = room/chamber; *ukuhlambela* = for washing). The Zulus who work in the compounds where

the chamber is so big that it is called a washing house, simply call it iwasha-wozi (washa = a borrowing from wash; wozi = a transliteration of house). Other interesting examples are umalaleveva for malaria fever (umalala = the one who sleeps; eveva = trembling) and ithuba-lakho-losizi for tuberculosis (ithuba = time; lakho = of you; losizi = of misery). In both cases (the one who trembles during his/her sleep, and your time of misery) the translation results in an accurate description in Zulu of the consequences of malaria fever and tuberculosis.

This was an attempt at highlighting initiatives that have been made in addressing the challenges of teaching African languages in their own medium. Reference has been made to Zulu examples so as to illustrate ways and means that are often employed to develop terminologies. Unsubstantiated criticism of some of these methods was discussed, while the validity of developments from the ground was acknowledged albeit with certain limitations where foreign theories not always accessible to people on the ground are involved. The need for publication and acceptability of the terms was emphasized as the necessary step towards standardisation. Standardisation at tertiary level, however, could be problematic since there is no overarching linguistic body for all colleges, technikons and especially the universities which enjoy a large measure of autonomy. Could a Quality Promotion Unit of SAUVCA (the South African University Vice-Chancellors Association) undertake such a standardising task particularly for the university membership? Could the necessary task be entrusted to a statutory authority such as PANSALB (the Pan South African Language Board) or a relevant SGB (Standards Generating Body) under the auspices of SAQA (the South African Qualifications Authority), or should this be left to scientific organisations such as ALASA (the African Language Association of Southern Africa)? Consensus on an acceptable choice needs to be sought from all interested role-players.

Bibliography

- Coulsin, J., C.T. Carr, L. Hutchinson, D. Eagle (Eds.). 1962. *Oxford Illustrated Dictionary*. London: Oxford University Press.
- Crystal, D. 1992. *An Encyclopaedic Dictionary of Language and Languages*. Cambridge: Blackwell.
- Department of Education and Training. 1993. *Zulu Terminology and Orthography*. Pretoria: The Government Printer.
- Doke, C.M., D.M. Malcolm, J.M.A. Sikakona and B.W. Vilakazi. 1990. *English-Zulu/Zulu-English Dictionary*. Johannesburg: Wits University Press
- Kumalo, M.B. 1987. Current and Proposed New Terminology in Language Description and Zulu Literature. *Educamus* 33(4): 21-24, April.
- Louwrens, L.J. 1994. *Dictionary of Northern Sotho Grammatical Terms*. Pretoria: Via Afrika.
- Van Huyssteen, L. 1999. *Research Methodology for the Testing of the Extent of Acceptability of Zulu Technical (Health) Terms and the Testing of Language Attitudes towards Zulu*. Ongoing Ph.D. research. Pretoria: University of South Africa.

Terminology Management at the National Language Service

Mariëtta Alberts, *National Language Service, Department of Arts, Culture, Science and Technology, Pretoria, Republic of South Africa*

Abstract: Terminology is a strategic resource in a multilingual country. It is the medium through which knowledge and information is disseminated. Through the use of correct, standardised terminology, effective scientific and technical communication skills are developed. A brief overview is given of terminology development in South Africa, with special emphasis on the work of the Terminology Division of the National Language Service. Aspects of present terminology practice such as terminology management, the nature of terminology and terminography, and work-flow procedure in a multilingual terminology office receive attention. To present training activities, initiatives and needs attention is also given.

Keywords: BILINGUAL, COMMUNICATION SKILLS, DISSEMINATION, DOCUMENTATION, DOMAIN, MULTILINGUAL, SOCIOLINGUISTICS, SOURCE LANGUAGE, STANDARDISATION, STANDARDISED TERMINOLOGY, SUBJECT FIELD, TARGET LANGUAGE, TECHNICAL DICTIONARIES, TECHNICAL LANGUAGE, TERMINOGRAPHER, TERMINOGRAPHY, TERMINOLOGIST, TERMINOLOGY, TERMINOLOGY DEVELOPMENT, TERMINOLOGY MANAGEMENT, TERMINOLOGY PRACTICE, TERMINOLOGY PRINCIPLES, TERMINOLOGY TRAINING

Opsomming: Terminologiebestuur by die Nasionale Taaldiens. Terminologie is 'n strategiese bron in 'n meertalige land. Dit is die medium waardeur kennis en inligting versprei word. Deur die gebruik van korrekte, gestandaardiseerde terminologie kan effektiewe wetenskaplike en tegnologiese kommunikasievermoëns ontwikkel word. 'n Kort oorsig word gegee van terminologieontwikkeling in Suid-Afrika, met spesiale klem op die werk van die Terminologieafdeling van die Nasionale Taaldiens. Aspekte van die huidige terminologiepraktyk soos terminologiebestuur, die aard van terminologie en terminografie en die werksvloeiprocedure in 'n meertalige terminologiekantoor ontvang aandag. Aan huidige opleidingsaktiwiteite, -inisiatiewe en -behoefte word ook aandag gegee.

Sleutelwoorde: BRONTAAL, DISSEMINERING, DOELTAAL, DOKUMENTASIE, DOMEIN, GESTANDAARDISEERDE TERMINOLOGIE, KOMMUNIKASIEVERMOËNS, MEERTALIG, SOSIOLINGUISTIEK, STANDAARDISERING, TERMINOGRAAF, TERMINOGRAFIE, TERMINOLOGIE, TERMINOLOGIEBESTUUR, TERMINOLOGIEONTWIKKELING, TERMINOLOGIEOPLEIDING, TERMINOLOGIEPRAKTYK, TERMINOLOGIESE BEGINSELS, TERMINOLOOG, TWEETALIG, VAKGEBIED, VAKTAAL, VAKWOORDEBOEKE

1. Introduction

The National Terminology Services (NTS) and State Language Services (SLS) of the Department of Arts, Culture, Science and Technology (DACST) merged on 1 April 1998 to form the new National Language Service (NLS).

With the merging of the two offices some aspects of the management of terminology and terminography changed. The office is now called the Terminology Division of the National Language Service.

The Terminology Division (TD) is the national office assisting the Government by providing terminology information in all the official South African languages. The TD documents, develops, standardises and publishes term lists covering various registers — the objective being to reduce ambiguity and misunderstanding, and thereby improving on the exactness of scientific and technical communication in the working environment.

The TD advises on national policy concerning technical language, establishes and evaluates norms for terminographic purposes, creates and implements standardised facilities and procedures to collect, document, systematise, standardise and disseminate terminological information for the different language groups in the Republic of South Africa, provides an information service on technical languages as well as on terminographic matters, co-ordinates, supports and facilitates projects for the development of technical languages, makes terminology facilities and products available to related information and communication systems and develops and manages the National Term Bank.

Originally resorting under the Science and Technology Branch of the Department of Arts, Culture, Science and Technology, the National Terminology Services (now Terminology Division) worked according to subject fields. Since then, the TD has been reintegrated under the National Language Service of the Arts and Culture Branch of the DACST and will henceforth also function according to language specific needs and will emphasise the development of the terminologies of the hitherto marginalised languages. The TD will prioritize projects with the objective of promoting scientific and technical communication in the African languages by enlarging and developing the various technical languages.

The centralised documentation of terms is important to avoid duplication and to promote standardisation. The TD acts as national facilitator recording the terminological input of various speech and subject communities in a central data bank. Terminological information can be obtained from this central office. In this way the TD serves the community and empowers them in communication.

The TD will become a clearing house for terminology work in South Africa, since the demand for terminology increases rapidly. Unfortunately the TD does not have the personnel to cope with this demand. Various aspects

concerning terminology and terminography management in the new NLS will receive attention in this article.

2. Brief Overview on Terminology Development in South Africa

The history of African language terminology in South Africa starts in 1928 with the formation of Language Committees — structures similar to language boards. These have changed over the decades in both structure and focus.

Since 1948 Language Boards have taken a fundamentally political outlook in composition, function and relation. These Boards were associated with "independent homelands" and similar political formations. Their political role was to present a picture of separate development, primarily of homeland structures. Some of these boards were accountable to the Chief Ministers and their Cabinets, some to the Departments of Education and Training and others to both the Ministers and the Departments. These Boards were funded mainly by the homeland governments and had representatives of these governments as members.

There were boards for every African language, though some would be duplicated for political purposes, as was the case in Transkei, Ciskei and Bophuthatswana (cf. Ntshangase 1996).

Translators started documenting English and Afrikaans terms on cards in the fifties, at the erstwhile Language Bureau of the Department of Culture (later the Department of National Education). Later, a Terminology Division developed within the Language Bureau. Terminologists were appointed to excerpt terms from documents and to systematise and to standardise the terminology of various subject fields. A variety of bilingual term lists and technical dictionaries were published.

Language bureaux for institutions such as the South African Defence Force (SADF), the Suid-Afrikaanse Akademie vir Wetenskap en Kuns (SAAWK), the South African Iron and Steel Industrial Corporation (IsCOR), Municipalities, South African Bureau of Standards (SABS), the South African Broadcasting Corporation (SABC), etc. that employed translators, also had a few terminologists to excerpt the terminology of the relevant institution. A Coordinating Terminology Board (COTERM) was formed in 1971 to co-ordinate the terminology endeavours of the various offices.

In 1976 the Terminology Division of the Department of National Education amalgamated with the Vaktaalbuuro of the SAAWK to form a new Terminology Bureau. This Bureau was later named the National Terminology Services (NTS) of the Department of Arts, Culture, Science and Technology (DACST) and the office became the official national office to document terminology. In April 1998 the NTS and the State Language Services of DACST amalgamated to form the National Language Service (NLS).

Although the NTS and its predecessors in the past primarily compiled bilingual, bidirectional technical dictionaries, it was realised that there was a

pressing need for terminology in the various African languages. The NTS therefore began research into word-forming principles in the indigenous languages. It was decided to start with the compilation of multilingual technical dictionaries in all national South African languages and with the addition of extended information such as definitions and explanations.

Since 1994, African language terminologists have been employed to document African language terminology in a variety of subject fields. Multilingual terminology lists such as *Weather Terms*, *Basic Health Terms*, *HIV/AIDS Terms*, *Building Terms*, *Election Terminology*, *Banking Terms*, *Commercial and Financial Terms* and *Computer Terms* are being compiled, as well as lists on subjects such as Indigenous Birds, Indigenous Mammals, Frail Care, Personnel Administration, Basic Physics, Basic Sewing, Basic Agriculture, Biology, Zoology, Sewing and Fashion, Health Education, and lists for various school subjects.

The Terminology Division of the NLS provides a base of knowledge for all language groups on National, Provincial and Local Government level and also for all other spheres of technical and scientific communication. The office serves the whole community. This includes all levels of communication in all registers in which terminology of the official languages is needed — from grass-roots level to the higher echelons of science and technology.

The Terminology Division of the NLS also works in close collaboration with outside multilingual projects such as the Legal Terminology Project of the Centre for Legal Terminology in African languages (CLTAL), the *Chemistry Dictionary* which is a combined effort by the SAAWK and the University of Pretoria, the revision of the 1967 *Arts Dictionary* by the SAAWK, terminology for the Correctional Services, etc. Work on a few outside bilingual projects such as the *Music Dictionary*, the Oceanography project, and dictionaries on Education, Physics, Geography, Dietetics, Statistics, Ophthalmology, Political Sciences and Mechanical Engineering will be completed soon. The Terminology Division also assists with international projects such as a multilingual project of the European Union as well as a project for the International Organization for Unification of Terminological Neologisms (IOUTN) that manages a World Bank of International Terms (WBIT)¹.

The Terminology Division of the NLS works in close collaboration with the Pan South African Language Board (PANSALB) and its related bodies. PANSALB is an independent statutory body appointed by the Senate in March 1996 in terms of the Pan South African Language Board Act. Its work is to monitor the observance of the constitutional provisions and principles relating to the use of languages as well as the observance thereof in and content of any existing and new legislation, practice and policy relating to language matters. It has a broad function and will ensure that politicians in Parliament are guided by a body with nonpolitical knowledge of relevant multilingual issues in South Africa (cf. LANGTAG 1996).

According to its Act, PANSALB has to establish provincial language councils and language-specific bodies. The PANSALB Act was revised in 1999 to incorporate the establishment and management of national lexicography units for each of the official South African languages.

The Terminology Division usually works from a specific source language (SL). In South Africa the source language is usually English and translation equivalents have to be provided in the other official South African languages. When providing or coining translation equivalents for SL terms, terminologists consult linguists and subject specialists. The NTS used to work in close collaboration with the old Language Boards, but at present the terminologists of NLS consult the language bodies already established by PANSALB.

3. Terminology Division

From 1 April 1998 the Terminology Division (previously National Terminology Services) functions as a subdirectoriate of the National Language Service of the Department of Arts, Culture, Science and Technology. It is the largest (and at the moment the only) organisation in South Africa devoting its time entirely to terminology work.

The mission of the Terminology Division (TD) is to assist in the development and modernisation of the technical vocabularies of the official languages of South Africa.

The aim of the TD is to support Government in the formulation, development, implementation and maintenance of a national policy and strategies concerning technical languages with a view to the promotion of scientific and technical communication in all South African language communities.

The objectives of the TD are to support the Government in the formulation, planning, development, implementation and maintenance of a national policy and strategies concerning technical languages with a view to the promotion of scientific and technical communication in all South African language communities, to render a terminology service for the advancement of the official languages at a national level by conceptualising, creating and standardising terminologies on behalf of all South African language groups, to develop the required structures and facilities needed to meet the requirements of the indigenous language communities, to facilitate and co-ordinate terminology development and other technical language projects and to develop, manage and maintain a computerised information-oriented terminological database (the National Termbank) to further the dissemination of technical language information.

Primarily bilingual (English/Afrikaans) technical dictionaries were compiled in the past, but at present there is a pressing need for multilingual dictionaries with additional information such as definitions and explanations in especially the various African languages, in various domains and at various levels of communication. The TD therefore aims to document technical and

scientific concepts and definitions and provide equivalents (e.g. by development or elaboration of terminology) in all the official languages, and in different domains — from grassroots level to the highest tertiary level.

The TD has the infrastructure, but unfortunately not enough personnel to carry out this task. It employs language practitioners (terminographers and terminologists) for all the official South African languages in order to produce the relevant terminology products.

The TD focuses on providing a service to the community. Clients of the TD includes all persons in South Africa concerned with education and training as well as interpreters, translators, copywriters and journalists, i.e. all people in career groups involved in communication in one way or the other. One important function of the TD is to provide a terminology information service.

The TD aims to document technical and scientific concepts in various subject fields and related domains. Translation equivalents are provided in all the official South African languages with a view to the promotion of the standardisation of technical terms of those languages. The terminology information is disseminated in various ways: by conventional means such as draft terminology lists, glossaries and technical dictionaries, or by electronic means such as CD-ROM and online (i.e. the National Termbank and, in future, the Internet).

Terminology lists are compiled according to priorities determined by needs assessments. The TD sends out questionnaires to a wide range of target groups to determine the varied terminology needs of the broad community.

Extensive research has been done on copyright issues affecting terminology in order to determine the legal position of the TD with regard to the products created by this office. It was found that the dictionaries and term lists produced by the TD are of an original nature and that the publication of these will therefore not constitute an infringement of the copyright of the sources consulted.

The MultiTerm Plus database system forms the basis of the restructured National Termbank which will enable the TD to disseminate terminological information electronically. All activities of development in the field of computer systems take place in accordance with standards as set by the International Standardisation Organisation (ISO). The unique demands of the African languages with regard to terminology processing are attended to in collaboration with the ISO.

Computer assistance from the TD is available to several projects inside the TD as well as to projects by individuals or institutions outside the NLS. The computer assistance consists of the alphabetisation and conversion of data sets, the writing of synonyms as headwords (cross-referencing), the retrieval of data, the printing of work and draft lists, and the provision of master copies for printing.

The TD is at present working on a variety of multilingual terminology lists on various subjects such as Geography, Physics, Basic Science, Statistics, Com-

puters, Meteorology, and a variety of lists concerning terminology related to the Building Industry, Political Sciences, Law, Public Administration, Mechanical Engineering, Education, Olympic Games, Mammals, Dietetics, Frail Care, Basic Health, HIV/AIDS, Agriculture, etc.

The TD provides guidelines on both terminographic and terminology matters by means of desktop publications. The predecessor of the TD — the National Terminology Services (NTS) — used to regularly update and publish a catalogue providing information about existing scientific and technical dictionaries on a regular basis. The NTS also published the catalogue *Language Resources*.

The NTS used to publish the newsletter *Termbroker* at regular intervals with the aim of distributing and co-ordinating terminology and terminography information for all indigenous languages.

The TD strives to serve the government and the general public by making terminological information available. This forms an important part of the community service rendered by the TD to empower people in scientific and technical communication.

The TD is increasingly acting in a supporting capacity through guidance, project facilitation and outside liaison. The TD aims to introduce the activities and facilities of the TD and also promotes the availability and advantages of the online National Termbank.

In order to increase public awareness of scientific and technical terminology, the TD has shown its terminological databank to the public at two recent Science and Technology exhibitions. The exhibitions also previewed the online National Termbank which is at present being developed and which will be available on the Internet. The TD also participated in a series of exhibitions for the Year of Science Road Show.

4. Terminology Management

The TD used to be divided into four subject-related sections, namely:

- Natural and Computer Sciences (i.e. Physics, Chemistry, Computer Science, Building Science, Statistics, Geology, Geography, Engineering, etc.),
- Human and Social Sciences (i.e. Personnel Administration, Psychology, Sociology, Legal Sciences, Political Sciences, etc.),
- Life and Agricultural Sciences (i.e. Biology, Agriculture, Zoology, Botany, Dietetics, Basic Health, Frail Care, HIV/AIDS), and
- Commercial Sciences (i.e. Financing, Banking, Accountancy, etc.).

Since 1994 a separate section for the development of the African Languages has been established and the sections Human and Social Sciences and Commercial Sciences have been merged to form a section Commercial and Human Sciences.

Two to three terminographers did group work for the respective projects in each subject-related section, such as the basic terminographical part of excerpting, documenting the source language (usually English), defining the concepts, etc. The terminologists of the African languages section then received a work list on a specific subject field (i.e. Biology) to provide the translation equivalents in his/her first language. They also translated source language definitions into the various target languages. This procedure implied that the African language terminologists were merely translating terms and definitions.

To correct this it was decided that at least two language practitioners would work together as team on a given project in future. One member of the team would be a terminographer (he/she could be a speaker of any of the official languages, but should be proficient in the source language) and the other member a terminologist (mother-tongue speaker of one of the target languages). They deal with all aspects concerning the project (e.g. excerpting, systematisation, documentation, defining, committee meetings (subject specialists and linguists), final editing, proofreading, front and back matter, etc.). They share the responsibilities for the project and meet deadlines.

Terminologists never work in isolation. They consult subject specialists when dealing with the concepts and terms of a specific subject field and also with linguists, mother-tongue speakers, and language committees when supplying translation equivalents or when coining terms.

5. A Model of the Work Flow within the Terminology Division

5.1 Types of requests for terminology

The TD decides on terminology projects in response to needs assessment studies and/or requests:

- The section Systems Development and Research of the TD conducts needs assessment studies in collaboration with various stakeholders and identifies client needs in the field of terminology, and/or
- The Head of TD receives a request from a client, institution or other user of the language for a terminology list in a specific subject field or domain.

5.2 The route² the project takes within the Terminology Division

5.2.1 The request is referred to the Service Co-ordinator (researcher of Systems Development and Research responsible for updating the project manager database on which all the terminology projects are listed) who

- liaises with the client in collaboration with the Terminology Co-ordinator (terminologist to co-ordinate terminology projects within the Termi-

CLIENT
 REQUEST FOR TERMINOLOGY LIST (X)
 (National project)

HEAD OF TERMINOLOGY DIVISION

⇐ X + SL + DEF + 10 TLs + DEF

⇓
 SYSTEMS DEVELOPMENT & RESEARCH
 (Needs assessment, determination of priority;
 registration of project, monitoring of progress)

⇑
 LANGUAGE INFORMATION PACKAGING
 (Print dictionary (DTP) = X + SL + DEF + 10 TLs + DEF)
 (e.g. multilingual explanatory dictionary, Term bank, Internet)

⇓

⇑

X

TERMINOGRAPHY SECTION
 (excerpt SL, define SL)
 (SL = English)

TERMINOLOGY SECTION
 (supply translation equivalents in 9 TLs)
 (supply definitions in TLs)

(Supply 1x TL (Afrikaans translation equivalent + definition))

(TLs = 9 Official South African languages)

X + SL + DEF + 1 TL + DEF

⇒

X + SL + 9 TLs + DEF

⇑

Natural Life Human & Commercial
 Sciences Sciences Sciences

Sesotho isiXhosa Xitsonga Tshivenda
 Sepedi isiZulu
 Setswana isiNdebele
 Siswati ⇓

⇓

CONSULTATION: SUBJECT SPECIALISTS
 LINGUISTS (Eng + Afr) LANGUAGE USERS

CONSULTATION: SUBJECT SPECIALISTS
 LINGUISTS LANGUAGE USERS

TERMINOLOGY SECTION NLS TRANSLATORS

⇔

TERMINOGRAPHY SECTION NLS TRANSLATORS

- nology Divisions) and with the relevant Section Head to establish the detailed specifications pertaining to the envisaged project and completes the project registration form during the interview with the client,
- finds out whether a similar project is running somewhere else and contacts the compilers who are responsible for the project to establish a relationship of sharing information and uniting the efforts to produce a terminology list,
 - writes a report on the feasibility of the project and submits it together with the completed project registration form to the management of the Terminology Division for consideration, and
 - adds the information pertaining to each project to the project database and updates the database at regular intervals.

5.2.2 The TD Management (the Head of the TD plus Section Heads) discusses the project in depth and compares it to the priorities determined by the latest national needs assessment study. The TD Management approves, outsources or rejects the project and puts it on a priority list once it has been approved.

Each section within the TD indicates whether it will be possible to meet the requirements and deadlines set for the project, namely:

- Systems Development and Research (SD&R),
- African Languages Section (ALS),
- Terminography Section (TS) (the subject-specific sections), and
- Language Information Packaging (LIP).

5.2.3 All the information pertaining to the project is documented on the project manager database (run by SD&R), whether the TD Management approves the project or not, together with the reasons for rejection. It serves as a future reference should the Terminology Division be in the position to work on the project at a later stage. SD&R will register the projects.

5.2.4 Should TD Management approve the project, the following steps are followed:

- The information is given to the relevant section to prepare the list in the source language (including the information requested by the client).
- If the source language of the request (e.g. documentation on a specific subject, etc.) is in English or Afrikaans, the project goes to the relevant subject-specific section to document the terminology and to supply definitions in the source language. A draft list then goes to the African Language Section for translation equivalents and definitions and will at this

stage also be provided with Afrikaans translation equivalents and definitions.

- If the source language is any one of the African languages, the project goes directly to the African Languages section to excerpt the terms, document them in the given source language and provide the terms with definitions in the source language. The terminologists working on the project also supply translation equivalents and definitions in English and then give the draft list to the other African language terminologists and the subject-specific section for verification and provision of the required translation equivalents and terminological information required for Afrikaans.
- The relevant Terminography Section forms a subject committee of specialists in the specific domain to assist with the terms and definitions if necessary. The terminographers working in the subject-specific sections may from time to time be asked to work across subject fields should the need arise.
- Each terminologist/terminographer also works on his/her own project. The compiler (terminologist/terminographer) excerpts the terminology, documents the terms and provides the terms with definitions. In addition the compiler supplies the English translation equivalents and definitions to each term in the source language. The compiler discusses the terminological information of his/her project with subject specialists and linguists before finalising the draft lists that will be submitted to the Section Head for approval. After completion of each individual list, the draft lists go to the terminologists/terminographers of the other languages to be provided with translation equivalents and definitions. With each terminologist/terminographer having his/her own project, they take charge of the management of such a project and are able to carry the project through the various developmental phases. This forms part of the in-house training of all terminologists/terminographers.
- On national projects, the Terminography Section (i.e. the subject-specific section) and the African Languages Section work on the list in tandem. As soon as the Terminography Section has completed a considerable number of entries (e.g. 50-100) in the source language, the head of the specific Terminography Section informs the co-ordinator of the African Languages Section.
- The work is allocated to the relevant African language terminologists to provide equivalents and/or provide or translate definitions. (All the languages may not be requested.)
- Whilst working on the list, the African language terminologists should either form an editorial committee or consult with other knowledgeable collaborators such as linguists or subject specialists with a knowledge of the specific language to assist them with their work and to verify the terminology.

- Progress is reported to the head of the specific section that co-ordinates the specific project on a weekly basis on a prescribed form.
- This information is submitted to the Service Co-ordinator to update the progress of the project on a monthly basis.
- After the completion of the work, the section co-ordinating the project ensures that the information on the database has been entered correctly and supplies LIP with the specifications for preparing printouts for proofreading.
- The responsible terminologists (of all eleven official languages) have to proofread their own languages and check for orthography, spelling mistakes, omission of diacritics or punctuation marks, etc.
- After this step has been completed, the section co-ordinating the project requests LIP to prepare a master copy of the document for the client.
- The document is handed over to the client who can disseminate the document as such or publish a dictionary.

5.2.5 Responsibilities

- The Section Head in the African Languages Section is responsible for all terminology projects and terminography work that have to be conducted by the mother-tongue terminologists in the nine official African languages.
- Section Heads in the Terminography Sections (subject-specific sections) are responsible for the work of the English/Afrikaans terminologists and terminographers.
- When a Terminography Section has the assistance of the African language terminologists to supply target language equivalents for the African languages for a national project, it is the responsibility of the relevant section head, or his/her deputy to arrange a meeting with the terminologists concerned on a specific time to request them to do the work and to discuss problems emanating from the list with them after consultation and in collaboration with the co-ordinator African languages.
- It is the responsibility of the Terminography Section Head concerned or his/her deputy to provide the terminologists with the information as to where to find the data (the name of the filter for the MultiTerm software package), and to provide a hard copy of the list each time the list has been updated or when a new list has been produced.
- The section Systems Development and Research is responsible for research and development studies related to the field of terminology.
- The section Language Information Packaging manages the National Termbank and the MultiTerm database. This section handles all publications related to terminology work.

6. Training in Terminology and Terminography

Since terminology is the vehicle for disseminating information, there is a dire need for trained terminologists and terminographers. Each subject field or domain (area of science, technology or work place) uses its own terminology. This terminology needs to be systematised, documented, and disseminated with a view to eventual standardisation. Only then will it be possible to communicate effectively. Terminologists can assist the subject specialists and language workers working in various subject fields or domains to document their terminology and to provide them with multilingual terminology lists.

One therefore not only needs trained terminologists to document terminology in the source language (usually English), but in order to facilitate proper communication one also needs African language terminologists to provide translation equivalents in the ten other official languages.

To my knowledge there is no institution or organisation in South Africa other than the NLS employing terminologists. In several language offices (local or provincial), translators or interpreters or other language workers sometimes document terms (institutionalised terms), but they are not formally employed as terminologists.

Throughout the country there is a gradual decline in the number of corporate language offices — and therefore also a decline in in-service training opportunities for potential terminologists. Without the support, supervision and assistance of experienced terminologists, new terminologists will find it extremely difficult to work at a standard that will satisfy the users of dictionaries and maintain the standards of the profession.

Four levels of terminology training can be distinguished:

- tertiary training in theory and principles of terminology,
- in-house terminology training,
- training of potential terminologists, and
- training of collaborators (language workers and subject specialists).

6.1 Tertiary training in theory and principles of terminology

Training in the theory and principles of terminology currently receives little attention where it forms part of the lexicography or translation modules at South African universities and technikons. The first postgraduate course in the principles and practice of terminology and terminography has been introduced in 2000 at the Department of African Languages, University of Pretoria.

The training of terminologists and terminographers should receive much more attention at tertiary level. The training should be done either in language departments (e.g. English, Afrikaans or African Languages) or in linguistics departments. "It should be an optional topic that students can take as part of their undergraduate studies," states a LANGTAG report (LANGTAG 1996: 82).

"Training should include basic lexicography, terminography and translation studies. Students can then specialise in their Honours studies."

6.2 In-house terminology training

All the terminologists at the TD have undergone in-house training. During this training, all aspects of terminology work are dealt with. The terminologist/terminographer is trained in theoretical as well as practical principles to carry a project through all the stages of the terminographical process.

6.3 Training of potential terminologists

A potential terminologist should have the opportunity to receive appropriate training before he/she enters the profession. It is sensible that such training should be a combination of tertiary training and intensive training given by a national terminology office such as the Terminology Division of the National Language Service.

6.4 Training of collaborators (language workers and subject specialists)

Apart from in-house training sessions the TD offers some training sessions to individuals and bodies outside the TD. The purpose of these sessions is to elucidate the principles underlying terminology work. The training also includes practical sessions. Part of the training involves the illustration of computer applications in processing terminology and terminography.

Trainees are language workers or subject specialists with a desire to know how to document and systematise the terminology they encounter on a daily basis. These collaborators can *inter alia* assist the national terminology office by

- identifying terms in their working environment,
- documenting these terms according to basic terminographical principles to be entered into the National Termbank,
- answering terminology queries,
- providing translation equivalents in the official South African languages according to basic word-forming principles,
- coining new terms,
- commenting on draft terminology lists, and
- verifying the translation equivalents supplied by the Terminology Division with first language speakers.

Due to a lack of personnel, the TD cannot meet all the terminology needs nationwide. The office aims to extend the training programme in the near future. Various workshops and winter schools (with terminology experts from

abroad as trainers) are planned to train collaborators in the principles and practice of terminology work. Such training will enable collaborators to work according to fixed terminological procedures and provide the National Term Bank with terminology related to their various specialist fields and domains. The TD will then be in a position to serve as a clearing house specialising in the documentation and dissemination of terminology in all official languages.

"The grassroots creativity of every speech community should be exploited and channels should be created to enable new terms at this level to be introduced into the central termbank," emphasises a LANGTAG report (LANGTAG 1996: 84). "For instance, schools, universities, newspapers, local administrations, etc. should be encouraged to send in their new terms to the Terminology Division."

7. Conclusion

Several factors influence the development of terminology in the African languages of South Africa. Most of these are sociolinguistic factors. In the past, terminology development in South Africa was hampered by a number of ideological, historical, and educational factors, such as the language policies adopted in the Republic of South Africa (RSA) (Mtintsilana and Morris 1988: 109). With English and Afrikaans as the official languages in the RSA up to 1994, the State actively supported terminology development in Afrikaans to try to ensure parity with English.

Since 1995, structural and policy factors have affected the terminology practice in South Africa. Some of the language bureaux have either been dissolved or scaled down to perform smaller tasks. The State has also accepted the responsibility to develop and promote African languages.

Government policy in language promotion before 1994 stipulated that the promotion of the African languages was the task of the former national states and the self-governing regions, and not that of the South African Government.

The terminologist working in a multilingual society such as South Africa today deals with conflicting situations. The African terminologies should be developed to enhance the multilingual heritage of the country, but there are too many languages to make this practical and economically viable. There are also very few trained terminologists and even less terminology posts available. There is at present only one training institution offering an official course in terminology. Aspects relating to terminology principles are dealt with in courses such as translation or lexicography, but terminology receives too little attention in these courses to be of any value. The Terminology Division of NLS offers in-house training. If South Africa wishes to solve its communication problems through multilingual terminology in a variety of subjects or domains, the Terminology Division of NLS will have to offer terminology training to people interested in this profession.

These matters emphasize the importance of the Terminology Division as an organisation working on a national basis to modernise and document the various technical vocabularies used in South Africa by compiling multilingual technical dictionaries and by providing terminology training to colleagues in the field of terminology.

Notes

1. The IOUTN has been affiliated to the United Nations as a nongovernmental organisation since 1987. The IOUTN affiliates philologists, linguists and specialists in various fields of science and technology who are interested in problems of transnationalisation and the dissemination of specialised terminology.
The main task of the IOUTN is to encourage the borrowing of specialised terminology from the language of its creators and to stimulate a consciousness in all countries that this action is in every nation's interest. The IOUTN publishes glossaries and vocabularies to facilitate the work of translators, interpreters, writers and workers in various domains. It aims to transfer all the latest specialised terminology to less developed and developing countries. The WBIT within the IOUTN engages in accumulating and storing international terminology and specialised neologisms using the materials supplied by more than 200 members from 40 countries all over the world. Incidentally, it is the first term bank in the world to hold a catalogue of multilingual international terms (Stoberski 1998: 6).
2. See work-flow diagram on p. 242.

Bibliography

- Alberts, M. 1983. *Aspekte van die rekenarisering van vakwoordeboeke*. M.A. Dissertation. Pretoria: University of Pretoria.
- Alberts, M. 1990. *'n Bepaling van Afrikaanse vakleksikografiese behoeftes*. D.Litt. et Phil. Thesis. Pretoria: University of South Africa.
- Alberts, M. 1996. *The Harmonisation of Nguni and Sotho: Terminological Perspectives*. Paper read at the Pre-Seminar Colloquium on the Harmonisation and Standardisation of African Languages for Education and Development, 11-13 July 1996, University of Cape Town, Cape Town.
- Alberts, M. 1997a. *Legalese: South Africa's 12th Official Language?* Paper read at the Seminar on Language in Court at the Former Venda Parliament (Government Buildings), Thohoyandou, 16-17 May 1997.
- Alberts, M. 1997b. Legal Terminology in African Languages. *Lexikos* 7: 179-191.
- Alberts, M. 1998. *The Harmonisation of Nguni and Sotho: Terminological Perspectives*. Prah, K.K. 1998: 229-242.
- Alberts, M. 1999. *Overview of Current Initiatives in Terminology*. Tutorial presented at the Tenth International Conference held by ALASA, 6 July 1999, University of South Africa, Pretoria.
- Alexander, N. 1989. *Language Policy and National Unity in South Africa/Azania*. Cape Town: Buchu Books.

- Calteaux, Karen. 1996. *Standard and Non-Standard African Language Varieties in the Urban Areas of South Africa*. Stanon Main Report. Pretoria: HSRC.
- Carstens, A. 1997. *Introduction to Terminology and Terminography*. Pretoria: University of Pretoria.
- Cluver, A.D. de V. 1974. *Voorstelle ter normering van die voorstukke van vakwoordeboeke*. Pretoria: UNISA.
- Cluver, A.D. de V. 1987. A Sociolinguistic Approach to the Study of Technical Languages. *Logos* 7(2): 13-30.
- Cluver, A.D. de V. 1989. *A Manual of Terminography*. Pretoria: HSRC.
- Cluver, A.D. de V. 1994. Preconditions for Language Unification. *South African Journal of Linguistics, Supplement* 20: 168-194.
- Cluver, A.D. de V. 1996. *Language Development*. Proceedings of a LANGTAG Workshop on Issues in Language Development: Focussing on the Future of the Apartheid Language Boards and the Harmonisation and Development of African Languages, Pretoria, 28 March 1996.
- Cluver, A.D. de V. n.d. *Preconditions for Language Unification*. Private paper.
- Felber, H. 1982. *Basic Principles and Methods for the Preparation of Terminology Standards*. Vienna: Infoterm.
- Felber, H. 1996. Worldwide Efforts to Establish Principles and Methods in Terminology and Terminography. *TermNet News* 54/55: 24-25.
- Fourie, D.J. 1994. The Rejection of Terminology: Observations from the African Languages. *South African Journal for African Languages* 12(1): 11-15.
- Galinski, C. and K.-D. Schmitz (Eds.). 1996. *TKE '96: Terminology and Knowledge Engineering*. Frankfurt/M: Indeks Verlag.
- LANGTAG. 1995. *Towards a National Plan for South Africa. A Report of the Language Plan Task Group*. Pretoria: Department of Arts, Culture, Science and Technology.
- LANGTAG. 1996. *Information on the Language Plan Task Group (LANGTAG)*. Pretoria: State Language Services, Department of Arts, Culture, Science and Technology.
- Louwrens, L.J. 1994. *Dictionary of Northern Sotho Grammatical Terms*. Pretoria: Via Afrika.
- Louwrens, L.J. 1997. On the Development of Scientific Terminology in African Languages: The Terminographer's Dilemma in a New Dispensation. *Lexikos* 8: 248-250.
- Msimang, C.T. 1994. Language Attitudes and the Harmonisation of Nguni and Sotho. *South African Journal of Linguistics, Supplement* 20: 147-167.
- Msimang, C.T. 1996. *The Nature and History of Harmonisation*. Proceedings of a LANGTAG Workshop on Issues in Language Development: Focussing on the Future of the Apartheid Language Boards and the Harmonisation and Development of African Languages, Pretoria, 28 March 1996.
- Mtintsilana, P.N. and R. Morris. 1988. Terminography in African Languages in South Africa. *South African Journal of African Languages* 8(4): 109-113.
- Nahir, M. 1984. Language Planning Goals: A Classification. *Language Problems and Language Planning* 8(3): 294-327.
- Nkondo, C.P.N. 1987. Problems of Terminology in African Languages with Special Reference to Xitsonga. *Logos* 7(2): 69-78.
- Ntshangase, D. 1996. *Working Document: Future of the Present Language Boards*. LANGTAG Subcommittee on the Development of (South) African Languages, 20 February 1996.
- Ohly, R. 1987. Corpus Planning, Glottoeconomics and Terminography. *Logos* 7(2): 55-67.

- Poulos, G. and L.J. Louwrens. 1994. *A Linguistic Analysis of Northern Sotho*. Pretoria: Via Afrika.
- Prah, K.K. 1998. *Between Distinction and Extinction. The Harmonisation and Standardisation of African Languages*. Johannesburg: Witwatersrand University Press.
- Sager, J.C. 1990. *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.
- SATI. 1996. *Bulletin* 6: 1-2.
- Stoberski, Z. 1998. *Multilingual Dictionary of International Terms*. Warszawa: World Bank of International Terms (WBIT) within the International Organization for Unification of Terminological Neologisms (IOUTN).

Lexicography in Gabon: A Survey¹

B. Nyangone Assam and P.A. Mavoungou, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa*

Abstract: This paper traces the historical development of lexicography in Gabon. Gabon, like most African countries, is multilingual. The recent inventories of languages spoken in Gabon are those established by Jacquot (1978) and Kwenzi-Mikala (1998). According to Kwenzi-Mikala (1997), there are 62 speech forms divided into 10 language groups or language-units in Gabon. These speech forms co-exist with French, the official language. In fact, in article 2 of paragraph 8 of the revised Constitution of 1994 the following can be read: "The Gabonese Republic adopts French as the official language. Furthermore, she endeavours to protect and promote the national languages." This constitutional arrangement naturally makes French the language used in education, administration and the media.

The survey of lexicography in Gabon that is presented here includes the linguistic situation in and the language policy of Gabon, the lexicographic survey itself, as well as the lexicographic needs of the different speech forms (including languages and dialects). Initially, the pioneers of Gabonese lexicography were missionaries or colonial administrators. Very little was done in this field by the Gabonese themselves. Although credit is to be given to these early works, there are a number of shortcomings regarding the linguistic as well as the metalexigraphic contents of dictionaries and lexicons produced during this period. In fact, the main weak point of those studies was the lack of tones in the written transcription of oral productions and orthographic problems. Furthermore, in those contributions, the theory of lexicography is largely unknown and lexicographic works are hardly ever based on authentic data corpora of the languages being described.

Keywords: SPEECH FORMS, LANGUAGE GROUP(S), OFFICIAL LANGUAGE, LANGUAGE POLICY, LEXICOGRAPHIC NEEDS, LEXICOGRAPHIC WORKS, DICTIONARY BASE, METALEXICOGRAPHY, LEXICOGRAPHIC TRAINING.

Abstrait: *Lexicographie au Gabon: Un état des lieux.* Le présent article retrace l'histoire de la discipline lexicographique au Gabon. Comme la plupart des pays africains, le Gabon est un état multilingue. Les classifications établies par Jacquot (1978) et Kwenzi-Mikala (1998) sont souvent citées comme les travaux d'inventaire les plus récents sur les parlers gabonais. Selon Kwenzi-Mikala (1997), le Gabon compte 62 parlers réparties en 10 unités-langues. Les dits parlers cohabitent avec le français qui faut-il le rappeler jouit d'un statut particulier. En effet, à l'article 2 du paragraphe 8 de la Loi fondamentale révisée de 1994, on peut lire ceci: «La République gabonaise adopte le français comme langue officielle. En outre, elle s'engage à protéger et à promouvoir les langues nationales». Cette disposition constitutionnelle fait naturellement du français le véhicule et la matière de l'instruction, la langue de l'administration et celle des médias.

L'état des lieux de la lexicographie au Gabon présenté ici prend en compte la situation et la politique linguistique au Gabon, l'état des connaissances lexicographiques à proprement parler ainsi que les besoins lexicographiques des différents parlers au regard de ce qui existe déjà. Les

premiers travaux à caractère lexicographique ont été réalisés essentiellement par les missionnaires et les administrateurs coloniaux. Très peu de gabonais étaient engagés dans la confection des dictionnaires et lexiques à cette époque pionnière. Enfin, il convient de souligner que le principal talon d'Achille de ces travaux de la première heure est à la fois linguistique et métaléxicographique. En effet, la majorité de ces travaux ne prennent pas en compte le système tonal des parlers décrits et s'inscrivent rarement dans un cadre théorique défini. Enfin, ils souffrent de nombreuses insuffisances orthographiques et sont rarement basés sur des corpus linguistiques.

Mots-clés: PARLERS, UNITÉ(S)-LANGUE(S), LANGUE OFFICIELLE, POLITIQUE LINGUISTIQUE, BESOINS LEXICOGRAPHIQUES, TRAVAUX LEXICOGRAPHIQUES, BASE DU DICTIONNAIRE, METALEXICOGRAPHIE, FORMATION LEXICOGRAPHIQUE.

Introduction

The survey of lexicography in Gabon presented here has a twofold objective:

- (a) to give an account of lexicographic works and studies in Gabon, and
- (b) to envisage the different lexicographic needs of the Gabonese languages.

However, before discussing lexicographic studies and the lexicographic needs, it is necessary to offer an explanation of the linguistic situation in and the language policy of Gabon.

An overview of the development of lexicography shows that there are imbalances between languages. Some are far more advanced than others. Therefore, the treatment of each group will vary in length according to the lexicographic data available.

1. Linguistic situation in Gabon

Like the majority of other African countries, Gabon is a multilingual country. It is, therefore, necessary to distinguish between multilingual states: those which have several or more dominant languages and those without dominant languages, except regionally. Gabon belongs to the latter group.

As far as classification is concerned, many studies exist on the Gabonese languages.

The inventories of Jacquot (1978) and Kwenzi-Mikala (1998 and 1998b) are regarded as the most recent classifications of the Gabonese languages. Jacquot's classification updates Guthrie's (1953) listing. It is a genetic or genealogical inventory. However, the classification proposed by Kwenzi-Mikala is geographical-administrative. Kwenzi-Mikala (1988: 57) initially grouped the 62 heritage speech forms ("parlers", including languages and dialects) in 8 language groups (a group of different speech forms that are mutually comprehensible) and thereafter (Kwenzi-Mikala 1998) in 10 language groups. In order to

establish these language groups or language units, Kwenzi-Mikala used the criteria of mutual intelligibility (also referred to as intercomprehension) and the opening greeting formality "I say that" (cf. Emejulu and Nzang-Bie 1999: 2). This classification² is as follows:

- the Mazuna group: Fan-Atsi, Fan-Make, Fan-Mvai, Fan-Ntumu, Fan-Nzaman and Fan-Okak;
- the Mye_ne group: Enenga, Ghalwa, Mpongwe, Nkomi, Orungu and Okoa;
- the Mekana-Menaa group: Akele, Ungom, Lisighu, Mbanwe, Metombolo, Seki, Tumbidi, Shake, Wumpfu and Lendambomo;
- the Mekona-Mangote group: Ikota, Benga, Shamayi, Mahongwe, Ndasha, and Bakola;
- the Membe (or Okande-Tsogho) group: Ghetsogh_o, Ghepinzi, Kande, Ghevhovhe, Ghehimbaka, Ghevhiya, Ebongwe and Kota-kota;
- the Merye group: Ghisira, Ghivharama, Ghivhungu, Yipunu, Yilumbu, Yisangu, Ngubi, Civili, Yirimba and Yighama;
- the Metye group: Yinzebi; Yitsengi, Yimwele, Yivhili, Liduma, Liwanzi and Yibongo;
- the Membre group: Lembaama, Lekanini, Lindumu, Lategh_e and Latsitseghe;
- the Mekana group: Bekwil, Shiwa (or Makina) and Mwesa;
- the Baka group: Baka.

Some criticism has been directed at this classification based on the sociolinguistic criteria of intercomprehension. Earlier, Hombert (1990: 30) asked himself: "S'agit-il d'une intercompréhension totale et immédiate? ou doit-on considérer qu'il y a encore intercompréhension si le locuteur d'un parler A a besoin d'une période d'adaptation (de quelques jours? de quelques semaines?) pour comprendre l'essentiel (la totalité?) des conversations des locuteurs d'un parler B?" More recently, Emejulu and Nzang-Bie (1999: 2) stated that "the intercomprehension criterion used confounds some crucial sociolinguistic historicity and does not systematically reflect the native speakers' linguistic intuition". Despite the shortcomings expressed above, Kwenzi-Mikala's classification serves as a useful source of reference for scientific works for two reasons. Firstly, it offers a solution to the problem of the enumeration of all the heritage speech forms of Gabon. After Johnston (1922: 138-144), Guthrie (1953: 55-73), Doke and Cole (1963: 63-76) and Jacquot (1978), Kwenzi-Mikala (1998) introduced a grouping of those speech forms in 10 language-units. Secondly, it is established according to a comparative underlying design. Other speech forms are still to be identified, especially those which are threatened with extinction.

As far as dialectology is concerned, Fan and Omyen_e are the sole Gabonese languages of which the dialectal status is more or less clearly established, having six dialects each. For the other Gabonese speech forms, dialectal

differences are still to be identified and mapped. This is the difficult task that is currently being carried out by the researchers of the *Laboratoire de phonétique et linguistique africaine* of the *Université Lyon 2* under the supervision of Prof. Jean-Marie Hombert using the framework of the *Atlas Linguistique du Gabon (ALGAB)*. The conclusion of this project will provide a clear picture of the language and dialect boundaries in Gabon.

2. Language policy of Gabon

According to Kwenzi-Mikala (1990: 123), the language policy of Gabon largely depends on the one inherited from the colonial era. In fact, in the revised constitution it is stipulated: "The Gabonese Republic adopts French as the official language. Furthermore, she endeavours to protect and promote the national languages" (1994, article 2, paragraph 8). With this official status, French serves as the sole medium of instruction in the national educational system, and as the language of business, in civil administration and the media.

For several years Gabonese languages have been sidelined. When the Government became aware of the importance of the national languages in the development of the country, it started to promote them. A seminar *Les États généraux de l'Éducation et de la Formation* was convened from 12 to 23 December 1983 in this regard. One of the recommendations of the workshop was the introduction of the Gabonese languages in the national educational system. However, no concrete policy ensued from this recommendation. Fourteen years later, this recommendation was stressed again during the *Table ronde sur les Recherches linguistiques et l'Enseignement des Langues au Gabon* that was held from 9 to 11 December 1997.

Recently, in April 1999, the *Séminaire sur la Standardisation de l'Orthographe des Langues gabonaises* was organised by the Ministry of National Education in Libreville.

National languages are used as a medium in the literacy program for adults and in broadcasting. The introduction of a postgraduate course at the *Normal sup* (College of Education) to train teachers of Gabonese languages during the academic year 1998-1999 is an indication that the educational system of the country is undergoing drastic changes. All the concrete actions that the Government has undertaken toward the development of the national languages will sooner or later allow the production of teaching materials enabling them to be used both as medium and subject of instruction in schools.

As far as lexicography is concerned, the *Séminaire sur la Lexicographie Bantu* held at CICIBA (Centre Internationale des Civilisations Bantu) in Libreville is an indication of some of the recent developments in this field. This seminar took place in 1997, and was attended by delegates from six African countries. During the seminar Dr D.J. van Schalkwyk, editor-in-chief and Dr F.J. Lombard, senior coeditor of the *Woordeboek van die Afrikaanse Taal (WAT)*, presented lexicographic training. Particular attention was paid to the

planning and management of lexicographic units as well as to practical aspects of lexicography (cf. Bureau of the WAT 1997-1998).

3. Survey of lexicography in Gabon

The survey of lexicographic works and studies in Gabon will be done according to the 10 language groups of Kwenzi-Mikala (1998a).³



- | | | |
|-------------------|----------------|-----------------|
| 1 Estuaire | 4 Ngounié | 7 Ogooué-Ivindo |
| 2 Ogooué-Maritime | 5 Moyen-Ogooué | 8 Ogooué-Lolo |
| 3 Nyanga | 6 Woleu-Ntem | 9 Haut-Ogooué |

3.1 Mazuna lexicography

Fan is the home language of 29 percent of the population (circa 427 000 people).

The speech forms of this group are mostly found in four of the nine provinces of Gabon, namely Estuaire, Moyen-Ogooué, Ogooué-Ivindo and Woleu-Ntem. Fan-Atsi is spoken in Ndjolé and Lambaréné; Fan-maké in Libreville and Kango. Fan-Mvai is found in Minvoul; Fan-Ntumu in Oyem and Bitam. Finally, Fan-Nzaman and Fan-Okak are spoken in the regions of Makokou-Ovan-Booué and Medouneu-Cocobeach-Mitzic respectively. Fan is a so-called cross-border language because it is also spoken in neighbouring countries, namely in Equatorial Guinea, Cameroon, São Tomé and the Republic of the Congo.

Mazuna lexicography starts with the publication of the bilingual *Dictionnaire fang-français* by R.P. Marling (1872). This lexicographic work is followed some years later by another bilingual dictionary, the *Dictionnaire français-fang* by R.P. Lejeune (1892). The work consists of 347 pages and has two sections. The first section contains an overview of Fan grammar; the second part is the dictionary itself. Important is the publication of the *Encyclopédie pahouine* by Largeau (1901), a colonial administrator. With its ±4 996 articles covering 699 pages, the book is the only encyclopaedia ever compiled in a Gabonese language. It is divided into two parts.

The first part is a cultural overview of the Fan people. Here the compiler does not only renew the question of the origin of the Bafan (speakers of Fan), but other anthropological issues regarding rituals, the value system and mythology are discussed extensively.

The second part, which contains the encyclopaedia itself, starts with a lexicological or lexicographic chapter (as it is called by the compiler). This chapter contains the user's guidelines and the minigrammar of the encyclopaedia. It provides the user with a relevant discussion on the pronunciation system, the orthography and the punctuation system used as well as some morphological information such as the parts of speech, word formation and the conjugation system of Fan that is abundantly illustrated by tables. Although credit is to be given to this work, it reveals a number of shortcomings. The writing system proposed by the compiler is very much along the lines of the orthographic traditions of French. However, there is a big difference between the sound systems of the African languages and those of the European languages. As a result, the transcriptions proposed are likely to be only approximate. Moreover, the grammatical content is somewhat outdated. These shortcomings can be remedied by taking note of some of the recent works on Fan grammar, namely those of Mba-Nkoghe (1981), Andeme Allogho (1980), Mba-Nzue (1981) and Ondo-Mebiame (1992) on Fan-Atsi, Fan-Nzaman, Fan-Mvai and Fan-Ntumu respectively. As far as the lexicographic section is concerned, the entries of the encyclopaedia are arranged in alphabetical order. They vary in length from short explanations of terms and concepts to a more comprehensive treatment of lexical items. Cross-references guide the reader to thematically related articles.

In the preface of the book it is clearly stated that the encyclopaedia is based on the dialects Maké and Atsi. However, the compiler does not mention either the primary or the secondary sources of his macrostructural elements. The encyclopaedia itself contains lexicographic data such as the lemma in French, the translation equivalent in Fañ, some contextual information, and indication of the plural.

According to Largeau (1901: 4) the value of the encyclopaedia lies in the fact that it can be used as a useful reference work by scholars from various fields, among others politics, philosophy, economics and ethnology.

The most important Fañ dictionary is that of the Genevan pastor Samuel Galley entitled *Dictionnaire fang-français et français-fang*, edited by Henri Messeiller and published in Paris in 1964. Like its predecessors, this work is a bilingual dictionary. It consists of ±13 925 articles covering 588 pages. It is the result of the input from the Société des Missions Évangéliques de Paris in Gabon through its station Talagouga in the region of Ndjolé. After 40 years in Gabon, Galley passed away in 1959 without having the opportunity to see his major work being published. The Fañ people are indebted to the Association française des Amis d'Albert Schweitzer for the publication of the Galley dictionary. It is based on the translation of the Bible in Fañ, another tremendous task that Galley undertook for the Fañ community of Talagouga. In contrast with earlier dictionaries such as the works by Marling (1872) and Lejeune (1892), the Galley dictionary is intended to meet the needs of both Fañ and French speakers because of its bidirectional nature. As far as the macrostructure is concerned, the lemmata are arranged alphabetically. The microstructure includes a lot of data, e.g. the parts of speech, class numbers, translation equivalents, meaning explanations as well as cross-references. Compare the article of the lemma *abĩnya*:

ABĨNYA (h) n. 4, pl. *mebĩnya*. Rameau vert, bout de branche avec feuilles vertes. On s'en sert pour fermer les barrages pour pêcher (*lôkh* h, pêcher). En faisant un paquet enveloppé d'écorces et en y mettant le feu, on obtient une épaisse fumée pour étourdir les abeilles et prendre le miel dans la ruche. Syn.: *ébi* h).

Another feature of the dictionary worth mentioning concerns the access structure: the use of typographical markers such as the presentation of the lemmata in bold small capitals as well as the running heads of the column which form the outer rapid access structure (Hausmann and Wiegand 1989: 329) make it easier for users to find the word for which they are looking. Apart from his lexicographic and evangelical activities, Pastor Samuel Galley was also involved in the production of didactic materials. His best-known teaching aid is the book *Nteni osu nteni Fañwe*, published in 1889 under the aegis of the Protestant station of Talagouga. The following is Mayer's (1990: 74) comment on the book: "(C'est) un opuscule de 14 pages centenaire! '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. Étrange illustration de couverture: minarets et chéchias composent un paysage sans doute exotique mais fantastiquement irréaliste."

The last work to be treated was compiled by Martrou (1924), a missionary of the *Congrégation des Pères du Saint-Esprit*. It is the bilingual *Lexique fan-français* consisting of 3 431 articles covering 137 pages. It is divided into two parts. Part I presents the outline of Fan grammar, while part II deals with the lexicon itself. The lexicon contains lemmata in Fan and French translation equivalents. The use of typographical markers such as the printing of the lemmata in bold capitals, as well as the running heads, have relatively improved the access structure of the lexicon. The orthography employed in the lexicon is based on the principles expounded by Ch. Sacleux in his book *Essai Phonétique*. The fact that the work contains information on stress in the lemmata shows the lexicographer's awareness of the necessity to give an account of the pronunciation of the language. According to Martrou, tonic accent in Fan is characterised by a global rise and the lengthening of the syllable. It is indicated by the stress mark (').

In conclusion, all the dictionaries available for Fan language are translating dictionaries and biased towards French. Furthermore, they reveal a number of shortcomings that need to be remedied. Most of these works do not mark tone or give the pronunciation of the language. To a greater or lesser degree, the earlier dictionaries lack a metalexigraphic foundation, but as far as these shortcomings are concerned, they are excusable if one considers that the compilers of these earlier works were neither linguists, nor trained lexicographers. The most important and most often stated shortcoming is that these dictionaries contain a lot of orthographic inconsistencies. For example, P. Alexandre (1961, as quoted by Kidda Awak 1990: 11) points out that the Fan language has acquired "trois transcriptions [catholiques] et deux [protestants] pour environ un million de locuteurs, sans parler des différences morphologiques dues à des calques catholiques sur l'allemand, le français et l'espagnol, et protestants, sur l'américain".

To a greater or lesser degree, these shortcomings also apply to the language group that is treated in the following section.

3.2 Myene lexicography

Omyene⁴ is the home language of 5 percent of the Gabonese population.

The speech forms of this language group are mainly found in three provinces of Gabon, namely Estuaire, Moyen-Ogooué and Ogooué-Maritime. Enenga is spoken northeast of Lambaréné, in the vicinity of lake Zilé. Ghalwa is found around lakes Onangué, Avanga and Ezanga, and on the island of Lambaréné. Mpongwe is spoken in Libreville as well as on the island of Pointe-

Dénis. Finally, *Nkomi* is in use southeast of Port-Gentil and around the lagoon Fernand-Vaz.

The first dictionary to be compiled in Myeng is the *Dictionnaire français-mpongwe* by R.P. Amable Delorme (1877). The dictionary covers 354 pages and was published under the auspices of the Missionnaires de la Congrégation du Saint-Esprit et du Saint-coeur de Marie. Lemmata are given in French and supplied with a Mpongwe translation equivalent as well as some examples. Except for suprasegmental and orthographic problems, the monodirectional feature of the work is the main weak point of the dictionary, that could only be used actively by speakers of French but passively by those of the Mpongwe dialect. This relation between active and passive language use is reversed in Gachon's *Dictionnaire mpongwè-français* published in 1881. The book covers 287 pages and consists of two sections. The first section presents Mpongwe grammar, while the second section contains the dictionary itself. Lemmata are arranged alphabetically and according to the word tradition. The dictionary has been a great contribution to Gachon's translation of the *Bible Gnango ine agamba mi re tendo pa gou'ejango j'agnambie ji felio né TESTAMAN NOUNGOU NI TESTAMAN GNONA gou'inongo gni mpongoue* published in 1891. The dictionary was republished in 1974.

The last contribution is Raponda-Walker's dictionary. It was published for the first time by the Imprimerie de la Libre Lorraine in Metz (1930-1934). It is monodirectional, with Mpongwe as source and French as target language. In 1961, 27 years later, the dictionary was published in the reverse direction French-Mpongwe by the Imprimerie Saint Paul in Brazzaville. This version of the dictionary has since been reprinted under the auspices of the Raponda-Walker⁵ foundation in 1995. The last version of the work comprises some 8 000 articles arranged alphabetically according to the full word tradition. The first part of the book is the dictionary itself, while the back matter or the second part of the dictionary contains an exposition of Mpongwe grammar. The dictionary articles include the following data categories: the lemma in French, the part of speech, a label, the Mpongwe translation equivalent, some examples as well as the use of typographical indicators. Compare the following example from this dictionary:

Communiquer V.A. (*transmettre*) nóza; pa. *Communiquer une lettre, une maladie, nóz'ezango ; nóze nkani. Communiquer un ordre, bol'okenge. (Correspondre). Communiquer par lettre, tendane yango; une nouvelle, myeze ntsango. (Aboutir) punda, pókósa. Le Fernand-Vaz communique avec l'Océan, Eliwe-Nkómi zi punda go ntsuwa. Ce chemin communique avec la plaine, mpónó yinó yi pókósa g'orove. Se communiquer, nógana. Le pian se communique, abukwè mi nógana.*

As far as the primary source of the dictionary is concerned, Raponda-Walker's work is based on examples of oral usage of Mpongwe which was collected over

many years from various sources. Earlier books on Mpongwe literature (especially dictionaries and the Bible), as well as some of the monolingual and bilingual dictionaries of French published during the period concerned, are most probably the secondary sources of the work. The compiler based the orthography used in the dictionary upon the set of symbols presented by himself (Raponda-Walker 1932). This alphabet is much along the lines of the *Practical Orthography of African Languages* published in 1928 by the International African Institute (see also Touré (1990: 56-57) and Kwenzi-Mikala (1998a: 219-220) for reviews on this matter).

From the foregoing exposition, it is clear that all the dictionaries available on the Omyene language have a common denominator: they are all translation dictionaries based on the Mpongwe dialect and were compiled by Catholic missionaries.

The next language group to be examined is less developed.

3.3 Mekana-Mena lexiconography

The speech forms of this language group are scattered in six provinces of Gabon, namely Estuaire, Haut-Ogooué, Moyen-Ogooué, Ngounié, Ogooué-Lolo and Ogooué-Ivindo. Nkele is spoken in Lambaréné and Sindara, Ungom north of Franceville, in Koulamoutou and around Mékambo. Lisighu and Metombolo are spoken in Lastoursville and Seki in Cocobeach. Shake is found in Booué and Lastoursville, and Tumbidi in Mbigou and Malinga. Lendambomo is spoken in Booué, Okondja as well as south of Mékambo. Finally, Wumpfu is found in Mbigou, Malinga and north of Franceville.

Except for word-lists at the end of linguistic works in this language group, nothing exists as far as lexicography is concerned.

3.4 Mekona-Mangote lexiconography

The speech forms of this language group are dispersed in five of the nine provinces of Gabon, namely Estuaire, Haut-Ogooué, Ngounié, Ogooué-Lolo and Ogooué-Ivindo. Ikota is spoken in Booué and Mékambo, Benga north of Libreville at Cap Esterias and Pointe Santa-Clara. Shamayi is found in Makokou and Okondja, Mahongwe and Bakola in Makokou. Finally, Ndasha is spoken in Mbigou, Mandjaye and Okondja. Moreover, the Bandasha are also localised in the Republic of the Congo, more precisely in Mossendjo and in Dolisie and in the region of Nyari in the commune of Moutamba.

At present, there exists only one lexicographic work in this group, namely *Lexique français-ikota* by R.P. Perron (1964), published under the auspices of the Mission Catholique de Makokou. An interesting fact is related to the primary sources of the work. The lexicon is based on the manuscript notes of R.P. Lamour. The articles are arranged in alphabetical order. They start with a French

lemma followed by a translation equivalent in Ikota, and some examples. The language group which follows, is far more developed.

3.5 Membe lexicography

The speech forms of this group are mostly found in three of the nine provinces of Gabon, namely Ngounié, Ogooué-Ivindo and Ogooué-Lolo. Ghetsoḡho is spoken in Mouila, Moabi and Sindara. Ghepinzi and Ghevhiya are found in Mouila and Ghevhoḡve in Koulamoutou. Ghehimbaka, Ebongwe and Kotakota are mostly found in Booué and Mimongo.

All the works available in Membe's literature touch the domains of linguistics and religion. These works, which contain a lot of information regarding the lexicon as well as the grammar of the speech forms of this group, can be used as a starting point for the compilation of new dictionaries. Studies that have dealt specifically with lexicographic issues in this language group include work by the following authors: Raponda-Walker (s.a.) and Bodinga-bwa-Bodinga and Van der Veen (1990). Raponda-Walker's contribution is a bilingual dictionary entitled *Dictionnaire getsogo-français*. The work contains more than 5 000 articles covering 237 pages. Lemmata in Ghetsoḡho are supplied with French translation equivalents as well as some examples. The author also compiled a reverse edition with French as source and Ghetsoḡho as target language. Bodinga-bwa-Bodinga and Van der Veen's paper "Plantes utiles des Evia" deals with plant names as well as their medicinal qualities. The work is much along the lines of the *Plantes utiles du Gabon* published in 1961 by Raponda-Walker and Sillans and contains some 608 Ghevhiya terms. The entries of the lexicon appear in three columns. The scientific names of the plants are given in the first column in alphabetical order. The Ghevhiya names appear in the second column. The last column deals with the therapeutic properties of the plants listed. This is illustrated by the following example:

Abrus precatorius	dì-ndende	Propriétés adoucissantes (feuilles), soins de la voix (chanteurs)
-------------------	-----------	--

Unfortunately, the authors propose no metalexicographic discussion on how these plant names should be treated in a dictionary of plant names for example. Finally, the collaboration of Bodinga-bwa-Bodinga and Van der Veen also includes the compilation of a bilingual dictionary: *Dictionnaire gevia-français* (in preparation).

3.6 Merye lexicography

The speech forms of this group are mostly found in four of the nine provinces of Gabon, namely Ngounié, Nyanga, Ogooué-Maritime and Ogooué-Lolo.

Ghisira is spoken in the regions of Fougamou, Mandji and Ndougou; Ghivharama in Setté Cama and Gamba. Ghivhungu is found in Mandji, Yetsou and Moabi. Yipunu is spoken in Mouila, Tchibanga, Ndendé, Mabanda and Moabi. Yilumbu is mainly spoken in Mayumba, Gamba and Setté Cama. Yisangu is found in the regions of Mimongo and Mbigou, and from Koulamoutou to Iboundji. Ngubi is spoken around the Iguéla lagoon, Civili in Mayumba and Ndindi, and Yirimba in Moabi. Yighama is found between Mayumba and Tchibanga. Some speech forms in this group are so-called cross-border languages, namely Yipunu, Yilumbu (also spoken in the Republic of the Congo and the Democratic Republic of the Congo) and Civili (also spoken in the Republic of the Congo and Cabinda in Angola).

The first lexicographic work that has to be mentioned in this group is the *Dictionnaire français-vili* by R.P. Marichelle (1900). The dictionary which covers 114 pages was published in Loango. It comprises a brief outline of the Civili speech form. The dictionary articles supply the following lexicographic data: a treatment unit in French, a Civili translation equivalent as well as some examples.

The next work published was the *Grammaire pounoue et lexique pounoue-français* by Borneau (1956). The work which covers 177 pages, contains two sections. The first section presents Yipunu grammar, while the second section contains the lexicon itself. The entries are arranged alphabetically. They provide the reader with lemmata in Yipunu as well as translation equivalents in French. As far as the primary sources are concerned, Borneau's work is based upon a series of papers that the author wrote from 1940 to 1952 in the *Journal de la Société des Africanistes*. A more recent article "Lexique" by Rittaud-Hutinet (1980) is very interesting with regard to the lexicographic contents. As far as the primary sources are concerned, the lexicon was initially compiled on the basis of a questionnaire of Joseph H. Greenberg. Then it was enlarged from other sources. The work contains a comprehensive introduction in which the compiler deals with issues regarding semantics (synonymy, antonymy, homonymy and figurative speech) and morphology (gender, concord and derivation) as well as the tonal system of the speech form. Another good feature of the work is that the lexicographer gives an account of the principle or method of arrangement according to which all the lexical items have been entered in the lexicon. Entries are presented in three columns. In the first column, Yipunu lemmata are given alphabetically under the initial letter of the stem of each word. The treatment units are accompanied by an indication of the class number and eventually by a label or usage marker. The second column contains translation equivalents as well as meaning explanations. The last column deals with semantic and morphological information such as synonyms, metaphors, compounds, etc. The following example illustrates the point in question:

-kanga 1/2 ami avec lequel on est en relation cf. -mba:tsi (C') 9/2 = ami
d'échanges et d'hospitalité

The lexicon itself contains some 945 articles and is divided into two sections which reflect the different tonal patterns of Yipunu. In one section the substantives are represented according to the following tonal classes: A, B, D, B/D, B' and B'/C' and in the other section the verbs are listed under the tonal classes A, B and A/B.

As far as Yipunu is concerned, the work of Kwenzi-Mikala is indispensable. Some of his contributions deal with semantic issues, such as borrowing, componential analysis and anthroponymy, to list a few. For example, his paper "Contribution à l'analyse des emprunts nominaux du yipunu au français" is a phonological and semantic analysis of Yipunu loan words based on a corpus of 70 words borrowed from French by the Gabonese speakers of Yipunu. In "Analyse sémique des termes dénotant les relations parentales en yipunu", the author analyses the kinship system and terms of Yipunu according to the model developed by the French structuralists in the field of semantics which is largely influenced by B. Pottier and A.-J. Greimas, among others. He organises his description on the basis of the following semantic axis: axis of generations; axis of sex; axis defining the linear character; axis defining consanguinity; and finally the axis defining the direct and lateral characters of the relation on the father's side.

Finally, his article "L'anthroponymie chez les Bapunu du Sud-Gabon" of 1990 deals with anthroponymic issues. Particular attention is paid to the relation between the function of the personal name and the status of the bearer. Moreover, the author is currently working on a corpus of traditional personal names including all the Gabonese tribes. From ongoing research work, some unpublished contributions on the speech forms of this language-unit should be mentioned. The first unpublished work is the *Lexique isangu-français* compiled by Naidailac (1992) and presented at the Université Lyon 2. The second contribution is Jean Blanchon's Yipunu dictionary in electronic form. It contains some 4 272 lemmata with their Protobantu reconstructions.

3.7 Metye lexicography

The speech forms of this group are mostly found in four of the nine provinces of Gabon, namely Haut-Ogooué, Ngounié, Ogooué-Ivindo and Ogooué-Lolo.

Yinzebi is the most widely spread speech form of the group and it is located in regions such as Mbigou, Lebamba, Malinga, Lastoursville, Koulamoutou, Pana-Iboundji and Mounana-Bakoumba. Yitsengi is spoken in Lastoursville-Moanda-Bakoumba, and Yiwele in Mbigou. Yivhili is found in Booué-Sindara, and Liwanzi in Lastoursville-Moanda. Liduma is spoken in Lastoursville and Yibongo in Moabi.

The first work in Metye lexicography is the lexicon *Vocabulaire français-aduma; vocabulaire aduma-français* by R.P. Dahin (1893). The author started with French lemmata and their translation equivalents in Liduma. Later, he reversed his macrostructural elements and started again, this time with Liduma lem-

mata. This makes the lexicon useful for speakers of both French and Liduma. As far as the primary sources of the lexicon are concerned, it seems that R.P. Dahin based his work on the *Catéchisme en langue adouma* that he wrote in 1891, as well as other literary materials.

The next lexicographic work to be published has a restricted macro-structure. It is the pocket dictionary of Yinzebi entitled *Petit dictionnaire bantou du Gabon: français-ndjabi; ndjabi-français* compiled by Muroi (1989). Despite segmental as well as suprasegmental shortcomings, the quality of the dictionary lies in its bidirectional feature. Another Yinzebi dictionary is being compiled by Nadaillac in collaboration with the ALGAB's team.

Finally, the most important dictionary of this group is certainly the *Dictionnaire ndumu-mbede-français* by A. Biton and J. Adam (1969). Biton and Adam's work is interesting because it does not only involve French and one of the Gabonese heritage speech forms as it is the case in most of the existing dictionaries, but it involves French and two Gabonese speech forms (Liduma and Lateghe). The primary source of this multilingual dictionary are probably the *Grammaire composée mbede-ndumu-duma* which Adam published in 1954. However, this grammar book has to be regarded as only one of the various sources of the author. In fact, after 52 years in Gabon where he passed away in 1981, Adam's contribution to Gabonese oral literature encompasses many articles on the different speech forms of Haut-Ogooué province.

3.8 Membere lexicography

The speech forms of this group are mostly found in Haut-Ogooué province. Lembaama is spoken in the region Franceville-Okondja-Akiéni. Lekanini is found in Franceville and Boumango. Lateghe is spoken in Akiéni-Lékoni and Latsitseghe in Lékoni. Lindumu is used in Franceville.

Membere lexicography starts with the publication of the bilingual dictionary *Dictionnaire français-ndumu et ndumu-français* by R.P. Biton (1907). The work covers 97 pages and is divided into two parts. The first part introduces the reader to the grammar of Lindumu, while the second part contains the dictionary itself. The first section of the dictionary starts with French lemmata. Then, the compiler reverses the languages in the second section of the dictionary. The present dictionary has served as a basis for the production of religious didactic materials such as the *Kateçism Ndumu* in 1962 and a few years later, in 1969, the *Dictionnaire ndumu-mbede-français et français-ndumu-mbede* compiled by Biton in collaboration with Adam. The characteristic feature of this dictionary is that it involves three languages (French, Lidumu (Ndumu) and Lateghe (Mbede)) and that the languages are reversed within one and the same dictionary. Furthermore, lemmata are excerpted from the lexical stock of the plant names of the region of Franceville. The next work to be examined is an interdialectal lexicon entitled "Vocabulaire comparé des principaux dialectes ayant cours en Haut-Ogooué" published by Castex in 1938. The work is an interesting contribution

to the field of dialectology (also called linguistic geography) and can be used as a valuable reference work for the compilation of a multilingual or pandialectal dictionary involving the different speech forms of the Membere group which presents a relative rich lexicographic insight in comparison to the last two groups of this survey, namely the Mekana and Baka language groups.

3.9 Mekana lexicography

The speech forms of this group are mostly found in the Ogooué-Ivindo province. Bekwil is spoken in Makokou and Shiwa (or Makina) in Booué. Mwesa is found in Mékambo.

As far as this group is concerned, a lot of work is still to be done. However, the contributions of Puech (1990), Yembi Bouka (1995) and Kwenzi-Mikala (1998a: 10-16) contain an outline of the grammar and the oral literature of Shiwa, Bekwel and Mwesa respectively. As far as lexicographic publications are concerned, nothing is available.

3.10. Baka lexicography

Baka is the sole speech form of this group which is spoken in Minvoul-Makokou-Belinga.

The available literature on this group includes the work of Mayer (1987) and Mayer and Voltz (1990) in the field of anthropology, among others. Lexicographic activities are still to be started.

4. Lexicographic needs of the Gabonese languages

The purpose of earlier dictionaries was basically to serve as reference works for European traders and French colonial administrators in their daily routine. Moreover, dictionaries have served as a source for the compilation of pedagogic or teaching materials, such as the translation of the Bible (for evangelisation purposes), grammar books and syllabi to teach the inhabitants to read and write.

As far as the lexicographic needs of Gabonese languages are concerned, there is a lack of dictionaries of various types as well as a lack of a dictionary culture. This does, of course, not mean that the lexicographic needs of Gabonese languages have to be derived from the full typological diversity of dictionaries available (cf. Gouws 1996: 100). Thus, there is a need for dictionaries aimed at a very specific target user. It is here that once again the comments of Gouws (1996: 100) can be recalled: "the first priority for any given language is to compile a dictionary belonging to a typological category already represented in that language but where the existing dictionary does not comply with the criteria set by the needs' analysis of the real users."

As it clearly appears in the above survey, most dictionaries available for Gabonese languages are translation dictionaries and biased towards French. Moreover, they lack important linguistic features such as tones as well as a metalexigraphic foundation. For the good of new dictionaries, the emphasis should be on the formulation of metalexigraphic criteria which can be applied in any translation dictionary (bilingual or multilingual). In this regard, Wiegand's General Theory of Lexicography⁶ can play a valuable role. For meeting the real needs as well as the reference skills of the intended target user, one also has to think about the improvement of the macro- and microstructure of existing dictionaries. And the first step to do so is the compilation of a database. Before embarking on any lexicographic activity, the lexicographer has to make sure that he or she has a representative corpus of the language which is to be described. In establishing a basis for a dictionary, one should preferably take into account all four constituents of Wiegand's general theory, namely a general section (also known as constituent theory A), an organisational section (also referred to as constituent theory B), a theory of lexicographic language research (also called constituent theory C) and a theory of lexicographic language description (also referred to as constituent theory D) (cf. Smit 1996).

For the purposes of this section, the emphasis will be on the constituent theories A, B and C.

As far as the constituent theory A is concerned, the historical overview of Gabonese lexicography has shown that one of the shortcomings of existing dictionaries regards the theory of lexicography or metalexigraphic. In other words, in those dictionaries metalexigraphic is hardly ever represented. For the good of new lexicographic activities, lexicographers will have to decide which structure is suitable to enhance the quality of the revised as well as the new dictionaries. Which one of these structures will fit the needs as well as the reference skills of the intended target user: an integrated, a nonintegrated or a semi-integrated microstructure?

Focus on decisions regarding the necessity to include lexical items as part of main lemmata or sublemmata should also be discussed. This is relevant because all the existing dictionaries seem to show a lack of sophistication due to the absence of nested and niched lemmata. In a niching dictionary, sublemmata which do not necessarily display a semantic relation are presented in a strict alphabetical order. Furthermore this internal ordering has to be maintained with regard to the preceding and following lemmata in the vertical ordering. In a nesting dictionary two levels have also to be identified. On the internal level, sublemmata occurring within a sinuous file are not arranged strictly according to the alphabet. On the external level, the horizontal ordering does not necessarily fit into the alphabetical arrangement of the vertical ordering. The deviation on alphabetical level is not only a space-saving device but also allows the lexicographer to give an account of morphosemantic relations between sublemmata (Hausmann and Wiegand 1989: 336).

Furthermore, as far as the dictionary articles are concerned, the average existing dictionaries only present the lemma (in the source language), morphological data, translation equivalent (in the target language) and examples. Other data categories such as indications of etymology and citations (which all existing dictionaries lack) could have improved the content of the articles of those dictionaries.

The organisational component⁷ is very important for existing as well as future dictionaries. It clearly states the mission of the dictionary project and gives a detailed discussion of all the structures involved in the preparation of a dictionary as well as an estimate of the time and capital needed for the project.

Finally, in the constituent theory C, Wiegand distinguishes two components, namely the theory of data collection and the theory of data processing. It is a well-attested fact that the compilation of a dictionary is practically impossible without an authoritative database. In establishing a database, a balance should always be maintained between oral and written sources. For the collection of spoken language, research has to be conducted at grassroots level through fieldwork. For the sake of representativeness and for the quality of language data, field workers should be well-equipped with recording material and their set of questionnaires should cover a wide range of subjects. Once the information is captured, transcripts of conversations, dialogues or interviews have to be computerised in the form of a database or in the form of a corpus. (Cf. chapters on "Language Databases, Corpora and Other Resources" and "Dictionary Data Input" in Bureau of the WAT 1999.)

For the collection of written material it is commonly accepted that African languages do not have a strong written tradition. However, some works do exist in Gabonese languages. In this regard, Hombert and Mortiers' (1990) bibliography gives quite an impressive picture of what has been written so far in Gabonese. Those works — which cover a wide range of subjects such as religion, ethnohistory, science, oral literature and literature dealing with dictionary research — are available in the Archives Spiritaines de Chevilly-Larue (Paris), the Archives Nationales and at CICIBA, in Libreville. The language material captured on paper (cards) and/or in electronic form can be valuable sources of data for the dictionary-making process. Electronic language data makes it easier for the lexicographer to execute quick queries and searches. A database is particularly useful for the compiler when it comes to statistical matter like the creation of frequency lists. Very often the inclusion of lexical items as main lemmata or as sublemmata is determined on the basis of the usage frequency of each lexical item. Finally, the most important advantage of electronic resources is their flexibility. A wide range of dictionaries can be obtained from the same database like translation dictionaries, special-field dictionaries as well as monolingual dictionaries which all Gabonese languages lack. As far as this point is concerned and in the light of Gallardo (1980: 61), one should not embark immediately on the tremendous task of compiling a comprehensive monolingual dictionary. The emphasis should first be on translation diction-

aries designed with an appropriate structure, making it possible for the lexicographer to supply lemmata with short meaning descriptions. Then due to the flexibility of a database, the short definition can be expanded to comply with the requirements of an explanatory monolingual dictionary for example.

It is also a well-attested fact that the preparation of a database may be costly and time-consuming. Due to the fact that most existing dictionaries on Gabonese languages have been the input of single lexicographers, future lexicographic products should be the result of team efforts.

Conclusion

The lexicographic survey of Gabonese languages above shows that most of the dictionaries available for the Gabonese heritage speech forms are bilingual dictionaries and biased towards French. Another feature is that the lemmas follow a vertical alphabetically arranged order. In other words, they are straight alphabetical dictionaries and lemmata are entered according to the word tradition (only Rittaud-Hutinet (1980) presents lemmas according to the stem tradition). Some of them present interesting lexicographic features such as guiding elements (the running heads of the column), outer texts giving an account of the principle of arrangement of lemmata, the user's guidelines and a mini-grammar for example. However, most of them reveal a number of shortcomings such as the absence of tones, orthographical inconsistencies as well as a lack of a metalexigraphic foundation. Furthermore, as far as dictionary purposes are concerned, most earlier dictionaries were compiled to help European traders and explorers to communicate in Gabonese languages. They were also useful aids for Christian evangelisation and French colonial administration. For the good of future dictionaries, the emphasis should not only be on the structure and the quality of the lexicographic data. Future dictionaries should also be compiled for a specific target group. In other words, the *genuine purposes* of the dictionary should be identified and provisions to fulfil the real needs and reference skills of the intended target user should be made (cf. Wiegand 1999: 299). This can only be achieved by consistent dictionary planning taking into account both organisational and metalexigraphic components. The lexicographic trend also shows an imbalance between speech forms. Some speech forms are far more developed than others. For others a lot of lexicographic work is still to be done. Therefore there is a necessity for well-trained lexicographers to bring all the Gabonese languages on par with other languages with a well-established lexicographic tradition.

Notes

1. We wish to express our profound gratitude to the Vice-Rector of Omar Bongo University, Dr Guillaume Moutou, the Dean of the Faculty of Arts and Human Sciences of Omar Bongo

University, Dr J.T. Kwenzi-Mikala and Prof. James Duplessis Emejulu, Director of Grelaco for all their efforts, without which we would not have had the opportunity to be trained by the Bureau of the *Woordeboek van die Afrikaanse Taal (WAT)* and to enrol as postgraduate students at the University of Stellenbosch. We must acknowledge that the original idea of this paper is from Dr J.C.M.D. du Plessis. We also want to thank Mr W.F. Botha for reading the article and making invaluable suggestions. Finally, we would like to thank Prof. R.H. Gouws, our promoter, for showing a real interest in all the lexicographic activities in which we are involved.

2. In this article we use the official Gabonese spelling and not the traditional semiphonetic spelling employed by Guthrie, Kwenzi-Mikala, and others. The underlined g, ɔ and n represents [ɛ], [ɔ] and [ŋ] respectively, while [y], [ʃ] and [β] are written gh, sh and vh respectively. The phonetic sound [a] is represented by the symbol ə in the Gabonese orthography. However, for phonological and practical reasons we have chosen to represent it by a, as in *Maḡe* and *Ghisira*.
3. The names of Gabonese tribes are based on the names the speakers of these tribes used for themselves or on administrative names inherited from the colonial era. Sometimes, the latter are the product of truncation. For example: The lexical items *pahouin* and later *pangwe* are both corrupted forms of the word *Faṅ* (also written *fang*) by speakers of *Mpongwe* based on their own ethnonym *Mpongwe*. The process can be explained as follows: *Faṅ* > *Mfaṅ* > *Mpaṅ* > *Mpaṅwe*, rendered as "pahouin" by the first French transcripts around 1845 (cf. Mayer and Voltz 1990: 45). The same line of argumentation holds true for the lexical item *Banzábi* (speakers of *Yinzebi*, *Nzabi* and *Njebi*) which comes from the wrong administrative orthography *Bandjabi*. In fact, speakers of *Yinzebi* called themselves *Munzeḡbi* (singular form) and *Banzeḡbi* (plural form) (cf. Mayer and Voltz 1990: 44 and 51).

In addition to the foregoing, the speech forms *Lindumu*, *Civili*, *Ntumu*, *Ghevhoḡhe*, *Ghepinzi*, *Yivhungu*, *Yilumbu*, *Ghisira*, *Lateghe*, *Lekaniṅi*, *Lembaama*, *Liwanzi*, *Ngubi*, *Ghalwa*, *Seki*, *Yisangu*, *Yitsengi*, *Maḡe*, *Yipunu*, *Liduma*, *Ikota*, *Ghetsoghḡ* and *Mbanwe* are sometimes associated with the following administrative or historical forms respectively: *NDoumou* (also *Ndumu*), *Vili* (also *Bevili*, *Fiotte*), *Ntoumou*, *Pove* (also *Bavove*, *Bapowe*), *Apindji* (*Pinji*, *Apinji*), *Voungou* (also *Bavungu*), *Baloumbou* (also *Lumbu*, *Ilumbu*), *Eshira* (also *Ghisir*, *Sira*), *Teke* (also *Bateke*), *Kaningi*, *Obamba* (also *Mbede*, *Mbere* and *Mbete*), *Bawandji* (also *Wandji*), *Nogové* (also *Ngové*), *Galwa* (also *Galoua*, *Gallois*), *Séki* (also *Sék-yani*, *Boulou*), *Yisangu* (also *Masangu*, *Massango*, *Sangu*, *Shango*), *Yitsengi* (also *Tsengui*, *Tsangui*, *Tsengi*), *Meke* (also *Maḡe*, *Makina*), *Pounou* (also *Pounoue* or *Punu*), *Adouma* (also *Aduma*, *Baduma*), *Kota* (also *Okota*, *Bakota*), *Tsogho* (also *Tsogo*, *Mitsogo*, *Getsogo*) and *Mbahouin* (also *Mbangwe*) (cf. Mayer and Voltz 1990: 50-51).

4. *Myeṅ* is the opening greeting formality, whereas *Omyeṅ* is the name of the language.
5. Mgr. André Raponda-Walker was the first Gabonese priest. He is the author of books and numerous papers on Gabonese literature. For example, in "Alphabet des idiomes gabonais" (1932), he presents a set of symbols according to which the Gabonese languages should be transcribed. Attention was paid to the principle of bi-univocity: one symbol for one sound. Fifty-seven years later, in 1989, the *Alphabet Scientifique des langues du Gabon (ASG)* was adopted. This alphabet which serves as reference for the transcription of the Gabonese languages is based on both the *International Phonetic Alphabet (IPA)*, revised version (Kiel

1989) and the "Africa" Alphabet published by the International African Institute (IAI). In "Néologismes sur les idiomes gabonais" (1933) he gives an account of the history of borrowing by some Gabonese languages (e.g. Fañ, Ghetsoghò, Ghisira and Civili) from Portuguese, English and French. "Dénominations astrales au Gabon" (1937) and "Enquête sur l'agriculture noire au Gabon et sur certaines techniques utilisant des produits végétaux" (1940) are terminologies or special-field dictionaries for astronomy and agriculture. The book *Plantes utiles du Gabon* (1961) which Raponda-Walker wrote in collaboration with Sillans gives a picture of the names and uses of Gabonese plants. Some 8 000 indigenous plant names, collected over a period of thirty years, as well as their scientific names and usage are listed. As far as dictionaries are concerned, Raponda-Walker is the compiler of *Dictionnaire mpongwè-français*, *Dictionnaire français-mpongwè*, *Dictionnaire getsogo-français*, *Dictionnaire français-getsogo* and *Dictionnaire étymologique des noms propres Gabonais*. The variety of his publications and his vast experience on oral tradition made him the first Gabonese scientist.

6. Wiegand's contribution to the field of metalexigraphy covers many aspects regarding theoretical models for the compilation of both monolingual and translation dictionaries as well as the history of lexicography and dictionary use. The study of Smit (1996), which is a dictionary plan for the compilation of a special-field (i.e. music) dictionary, contains a comprehensive discussion of Wiegand's metalexigraphy in part I, chapters 1 to 8.
7. We completed a training course at the Bureau of the Woordeboek van the Afrikaanse Taal (WAT) from 12 to 23 April 1999. The course covered the following themes: general and computer lexicography, and the planning and management of a lexicographic project. As a planning assignment, we compiled a framework for the revision of Raponda-Walker's *Dictionnaire français-mpongwè* under the guidance of Dr D.J. Van Schalkwyk, Editor-in-Chief of the Woordeboek van die Afrikaanse Taal (WAT). This modest work can be a useful tool for the revision of the proposed dictionary as soon as the project is established.

After the training at the Bureau, we also attended a series of lectures presented by Prof. R.H. Gouws in the Department of Afrikaans and Dutch at the University of Stellenbosch. During the course, the focus was on a general theory of lexicography (with special reference to Wiegand's theory). As far as dictionary typology is concerned, particular attention was given to equivalent relations in translation dictionaries.

Bibliography

- Adam, J. 1954. *Grammaire composée mbede-ndumu-duma*. Brazzaville: Institut d'Études Centrafricaines.
- Andeme Allogho, M.-F. 1980. *Esquisse phonologique du nzaman de Makokou*. Mémoire de Licence. Libreville: Université Omar Bongo.
- Biton, A. 1907. *Dictionnaire français-ndumu et ndumu-français, précédé d'éléments de grammaire*. Nantes: Imprimerie Bourgeois.
- Biton, A. 1962. *Kateçism Ndumu*. Rome: Société St Pierre Claver.
- Biton, A. and J. Adam. 1969. *Dictionnaire ndumu-mbede-français et français-ndumu-mbede. Petite flore de la région de Franceville (Gabon). Grammaire ndumu-mbede*. Archevêché de Libreville, Bar-le-Duc: Imprimerie St Paul.

- Blanchon, J. 1997. Les formes nominales de citation à préfixe haut en Pounou (Bantu B43). *Journal of African Languages and Linguistics* 18: 129-138.
- Blanchon, J. In preparation. *Dictionnaire informatisé du Pounou*.
- Bodinga-bwa-Bodinga, S. and L.J. van der Veen. 1990. Plantes utiles des Evia. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 8: 26-65.
- Bodinga-bwa-Bodinga, S. and L.J. van der Veen. In preparation. *Dictionnaire gevia-français*.
- Bonneau, J. 1956. *Grammaire pounoue et lexique pounoue-français*. Brazzaville: Institut d'Études Centrafricaines.
- Bureau of the WAT. 1997-1998. *Activities of the Woordeboek van die Afrikaanse Taal (WAT)*. Stellenbosch: Bureau of the WAT.
- Bureau of the WAT. 1999. *Course in Computer Lexicography*. Unpublished course notes.
- Castex (Dr). 1938. Vocabulaire comparé des principaux dialectes ayant cours en Haut-Ogooué. Essai de classification. *Bulletin de la Société des Recherches Congolaises* 26: 23-54.
- Dahin (R.P.). 1893. *Vocabulaire français-aduma; vocabulaire aduma-français*. Kempton: Kosel.
- Delorme, A. 1877. *Dictionnaire français-mpongwè*. Paris: Missionnaires de la Congrégation du St Esprit et du St-cœur de Marie.
- Doke, C.M. and D.T. Cole. 1963. *Contributions to the History of Bantu Linguistics*. Johannesburg: Witwatersrand University Press.
- Emejulu, J. and F. Pambo Loueya. 1990. *Yilumbu. Revue Gabonaise des Sciences de l'Homme* 2: 197-201.
- Emejulu, J. and Y. Nzang-Bie. 1999. *Linguistic Perspectives in Gabon*. Colloquium, Grand Forks, Summer Institute of Linguistics, University of North Dakota, July 20, 1999.
- Gachon, J.-B. 1881. *Dictionnaire mpongwè-français, précédé des principes de la langue mpongwè*.
- Gachon, J.-B. 1891. *Bible Gnango ine agamba mi re tendo pa gou'ejango j'agnambie ji felio né TESTAMAN NOUNGOU NI TESTAMAN GNONA gou'inongo gni mpongoue*. Sente-Marie yi Gabon. Freiburg im Breisgau: Herber B. Libr. Edit. Pontifical.
- Galley, S. 1889. *Nteni osi nteni Faiwe*. Talagouga: Imprimerie de la Mission.
- Galley, S. 1964. *Dictionnaire fang-français et français-fang, suivie d'une grammaire fang*. Neuchâtel: Henri Messeiller.
- Gallardo, A. 1980. Dictionaries and the Standardisation Process. Zgusta, L. (Ed.). *Theory and Method in Lexicography*: 59-69. Columbia: Hornbeam Press.
- Gouws, R.H. 1996. A Sequence for Meeting Lexicographic Needs. *Lexicography as a Financial Asset in a Multilingual South Africa*: 97-110. Pretoria: Department of Arts, Culture, Science and Technology.
- Gouws, R.H. and D.J. Prinsloo. 1998. Cross-Referencing as a Lexicographic Device. *Lexikos* 8: 17-36.
- Grimes, B.F. 1996¹³. *Ethnologue*. Dallas: Summer Institute of Linguistics, Inc.
- Guthrie, M. 1953. *The Bantu Languages of Western Equatorial Africa*. Oxford: Oxford University Press.
- Hausmann, F.J. and H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. (Eds.). 1989-1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie / Dictionaries. An International Encyclopedia of Lexicography / Dictionnaires. Encyclopédie internationale de lexicographie*: 328-360. Berlin: Walter de Gruyter.
- Hombert, J.M. 1990. Les langues du Gabon: État des connaissances. *Revue Gabonaise des Sciences de l'Homme* 2: 29-36.

- Hombert J. M. and A.-M. Mortier. 1990. Bibliographie des langues du Gabon. *Revue Gabonaise des Sciences de l'Homme* 2: 335-355. Mise à jour de la bibliographie publiée dans *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 1. 1984.
- Jacquot, A. 1978. Le Gabon. D. Barreteau. *Inventaire des études linguistiques sur les pays d'Afrique Noire d'expression française et sur Madagascar*: 493-503. Paris: CILF.
- Johnston, Harry H. 1922. *A Comparative Study of the Bantu and Semi-Bantu Languages. Volume II*. Oxford: Clarendon Press.
- Kidda Awak, Mairo. 1990. Historical Background, with Special Reference to Western Africa. Hartmann, R.R.K. (Ed.). 1990. *Lexicography in Africa*: 8-18. Exeter: University of Exeter Press.
- Kwenzi-Mikala, J.T. 1987. Contribution à l'inventaire des parlers Bantu du Gabon. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 2: 103-110.
- Kwenzi-Mikala, J.T. 1988. L'identification des unités-langues bantu gabonaises et leur classification interne. *Muntu* 8: 54-64.
- Kwenzi-Mikala, J.T. 1988a. Quelques remarques sur la transcription des textes oraux en langues africaines. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 3: 207-211.
- Kwenzi-Mikala, J.T. 1989. Contribution à l'analyse des emprunts nominaux du yipunu au français. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 4: 157-170.
- Kwenzi-Mikala, J.T. 1990. Quel avenir pour les langues gabonaises? *Revue Gabonaise des Sciences de l'Homme* 2: 121-124.
- Kwenzi-Mikala, J.T. 1990a. L'anthroponymie chez les Bapunu du Sud-Gabon. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 5: 113-120.
- Kwenzi-Mikala, J.T. 1998. Parlers du Gabon. A. Raponda-Walker: *Les langues du Gabon*: 217. Libreville: Éditions Raponda-Walker.
- Kwenzi-Mikala, J.T. 1998a. *Parémies d'Afrique centrale*. Libreville: Éditions Raponda-Walker.
- Kwenzi-Mikala, J.T. 1998b. Localisation des parlers du Gabon. A. Raponda-Walker: *Les langues du Gabon*: 215-216. Libreville: Éditions Raponda-Walker.
- 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.
- Marichelle, C. 1900. *Dictionnaire français-vili*. Loango.
- Marling (R.P.). 1872. *Dictionnaire fang-français*. New York.
- Martrou, L. 1924. *Lexique fān-français*. Paris: Procure Générale (des Pères du St Esprit) / Abbeville: Imprimerie Paillard.
- Mayer, R. 1987. Langues des groupes Pygmées du Gabon: un état des lieux. *Pholia. Revue du Laboratoire de Phonétique et Linguistique Africaine* 2: 119-127.
- Mayer, R. 1990. Histoire de l'écriture des langues du Gabon. *Revue Gabonaise des Sciences de l'Homme* 2: 65-91.
- Mayer, R. and M. Voltz. 1990. Dénominations ethnoscientifiques des langues et des ethnies du Gabon. *Revue Gabonaise des Sciences de l'Homme* 2: 43-53.
- Mba-Nkoghe, J. 1981. *Phonologie et classes nominales en fang (Langue bantoue de la zone A)*. Thèse 3^e cycle. Paris: Université Sorbonne Nouvelle.
- Mba-Nzue, N. 1981. *Esquisse phonologique du mvai (parler de Minvoul)*. Mémoire de Licence. Libreville: Université Omar Bongo.

- Mouguiama-Daouda, P. 1995. *Les dénominations ethnoichtyologiques chez les bantous du Gabon*. Thèse de doctorat. Lyons: Université Lyon 2.
- Muroni, J.-M. 1989. *Petit dictionnaire bantou du Gabon: français-ndjabi, ndjabi-français*. Paris: L'Harmattan.
- Naidailac, L. 1992. *Lexique isangu-français*. Lyons: Université Lyon 2.
- Naidailac, L. In preparation. *Lexique yinzebi-français*.
- Ondo-Mebiame, P. 1992. *De la phonologie à la morphologie du fang parlé à Aboumezok (langue Bantu A. 78)*. 2 volumes. Thèse de doctorat. Tervuren: Université Libre de Bruxelles.
- Perron, P. 1964. *Lexique français-ikota*. 2 volumes. Makokou: Mission Catholique.
- Puech G. 1990. Bekwel. *Revue Gabonaise des Sciences de l'Homme* 2: 127-128.
- Raponda-Walker, A. 1930-1934. *Dictionnaire mpongwè-français, suivi d'éléments de grammaire*. Metz: Imprimerie de la Libre Lorraine.
- Raponda-Walker, A. 1932. L'alphabet des idiomes gabonais. *Journal de la Société des Africanistes* 3(2): 305-314.
- Raponda-Walker, A. 1933. Les néologismes dans les idiomes du Gabon. *Journal de la Société des Africanistes* 3(2): 305-314.
- Raponda-Walker, A. 1937. Dénominations astrales au Gabon. *Bulletin de la Société des Recherches Congolaises* 24: 150-166.
- Raponda-Walker, A. 1940. Enquête sur l'agriculture noire au Gabon et sur certaines techniques utilisant des produits végétaux. *Revue de Botanique et d'Agriculture Appliquée* 24: 150-166.
- Raponda-Walker, A. 1961. *Dictionnaire français-mpongwè*. Brazzaville: Imprimerie St Paul. 1995². Classiques Africains.
- Raponda-Walker, A. s.a. *Dictionnaire getsogo-français / Dictionnaire français-getsogo*. Typescript.
- Raponda-Walker, A. and R. Sillans. 1961. *Plantes utiles: essai d'inventaire et de concordance des noms vernaculaires et scientifiques des plantes spontanées et introduites du Gabon*. Paris: Lechevalier.
- Rittaud-Hutinet, C. 1980. Lexique. Nsuka-Nkutsi, F. (Ed.). *Éléments de description du Punu*: 193-245. Lyons: PUL.
- Smit, M. 1996. *Wiegand's Metalexigraphy as a Framework for a Multilingual, Multicultural, Explanatory Music Education Dictionary for South Africa*. Unpublished D.Litt. thesis. Stellenbosch: University of Stellenbosch.
- Touré, A. 1990. L'écriture des langues africaines: évolution et principes méthodologiques. *Revue Gabonaise des Sciences de l'Homme* 2: 55-63.
- Van Wyk, E.B. 1995. Linguistic Assumptions and Lexicographic Traditions in the African Languages. *Lexikos* 5: 82-96.
- Wiegand, Herbert E. 1999. *Semantics and Lexicography. Selected Studies (1976-1996)*. Edited by Antje Immken and Werner Wolski. Tübingen: Max Niemeyer.
- Yembi Bouka, L. 1995. *Structures phonologiques et structures prosodiques (le modèle bekwel)*. Thèse de doctorat. Tervuren: Université Libre de Bruxelles.

Nuwe dinge in *Nuwe woorde* — 'n oorsig oor buitemikrostrukturele leksikografiese vernuwings

Herman L. Beyer, *Departement Germaanse en Romaanse Tale, Universiteit van Namibië, Windhoek, Namibië*

Abstract: *New Things in New Words* — an overview of nonmicrostructural lexicographic innovations. This review article focuses on the nonmicrostructural lexicographic innovations which are apparent in the dictionary under discussion. The emerging notion of user-friendliness in modern lexicographic practice does in no way imply that lexicographic innovations should be limited to the microstructure of dictionaries. Innovative approaches to the presentation of information in the user's guide (where a user's guide exists in a dictionary!) and other elements of the front and back matter in dictionaries play as an important role in making a dictionary more user-friendly as do innovations in the microstructure. *New Words/Nuwe woorde*, as the newest contribution to the collection of Afrikaans translation dictionaries, exhibits some remarkable nonmicrostructural lexicographic innovations. These innovations are highlighted, evaluated and criticised in this article.

Keywords: ACCESSIBILITY, BACK MATTER, DICTIONARY USER, FRONT MATTER, INFORMATION TYPES, LEXICOGRAPHER, LEXICOGRAPHIC INNOVATIONS, LEXICOGRAPHY, MACROSTRUCTURE, MEDIATESTRUCTURE, MICROSTRUCTURE, NONMICROSTRUCTURAL LEXICOGRAPHIC ELEMENTS, PRONUNCIATION GUIDANCE, TRANSLATION DICTIONARY, USER'S GUIDE, USER-FRIENDLINESS

Opsomming: Hierdie resensieartikel fokus op die niemikrostrukturele leksikografiese vernuwings wat in die woordeboek onder bespreking voorkom. Die toenemende bewustheid van gebruikersvriendelikheid in die hedendaagse leksikografiese praktyk beteken nie dat leksikografiese vernuwings tot die mikrostruktuur van woordeboeke beperk moet word nie. Vernuwende benaderings tot inligtingsaanbod in die gebruikerstoeligtig (as daar een in die woordeboek is!) en ander elemente in die voor- en agterwerk van woordeboeke speel 'n ewe belangrike rol in die gebruikersvriendelik maak van woordeboeke as vernuwings op mikrostrukturele vlak. *New Words/Nuwe woorde*, as die nuutste toevoeging tot die versameling Afrikaanse vertalende woordeboeke, vertoon 'n aantal besondere buite-mikrostrukturele leksikografiese vernuwings. Hierdie vernuwings word in hierdie artikel uitgelig en geëvalueer.

Sleutelwoorde: AGTERWERK, BUTTEMIKROSTRUKTURELE LEKSIKOGRAFIESE ELEMENTE, GEBRUIKERSLEIDING, GEBRUIKERSVRIENDELIKHEID, INLIGTINGSTIPES, LEKSI-

* Du Plessis, Madaleine (Hoofred.). 1999. *New Words, and previously overlooked ones/Nuwe woorde, en oues wat in die slag gebly het*. Kaapstad: Pharos.

KOGRAAF, LEKSIKOGRAFIE, LEKSIKOGRAFIESE VERNUWINGS, MAKROSTRUKTUUR, MEdIOSTRUKTUUR, MIKROSTRUKTUUR, TOEGANKLIKHEID, UITSPRAAKLEIDING, VERTALENDE WOORDEBOEK, VOORWERK, WOORDEBOEKGEBRUIKER

1. Inleiding

In 'n tyd dat Afrikaans onder verskillende tipes druk verkeer, is die verskyning van die woordeboek *New Words/Nuwe woorde* 'n baie tydige teken dat die taal lewenskragtig funksioneer en heeltemal in staat is om aan die eksponensieel toenemende taaleise van die inligtingsontploffing te voldoen. Hierdie siening word ondersteun deur 'n opmerking in die Voorwoord:

Dis 'n publikasie wat gegrond is op taal soos dit in die praktyk gebruik word. Die redaksie het sy werk deurlopend getoets aan die omvangryke databanke wat Pharos tot sy beskikking het. Slegs by hoë uitsondering, wanneer bronsoektogte doodgeloop het, is vertaalekwivalente agter lessenaars geskep.

Vanuit 'n metaleksikografiese oogpunt is dit ook verblydend om op te merk dat hierdie publikasie 'n aantal leksikografiese vernuwings bevat, wat tekenend is van die redaksie se kennisname van onlangse ontwikkelinge in die metaleksikografie. Dit is op hierdie vernuwings waarin hoofsaaklik in hierdie resensie-artikel kommentaar gelewer word, en die bespreking sal fokus op die buitemikrostrukturele elemente waarin leksikografiese vernuwings voorkom. Uitspraakleiding as mikrostrukturele element verdien egter besondere vermelding en word daarom by die bespreking ingesluit.

In die Voorwoord word gemeld dat hierdie woordeboek saamgestel is om Pharos se bestaande tweetalige handwoordeboeke aan te vul. Hiervan is die veertiende, uitgebreide uitgawe van *Groot woordeboek/Major Dictionary* (Eksteen 1997) (voortaan GW) en die eerste nuwe uitgawe van *Tweetalige Frasewoordeboek/Bilingual Phrase Dictionary* (Joubert 1997) (voortaan TFW) die nuutste. Veral eersgenoemde sal vergelykenderwys in die bespreking van enkele van die belangrikste buitemikrostrukturele leksikografiese vernuwings in *New Words/Nuwe woorde* (voortaan NW) betrek word.

Die teikengebruikers van NW is waarskynlik die gemiddelde gebruikers van GW en TFW, en meer bepaald waarskynlik taalpraktisyne in alle professies.

2. Gebruikersgerigtheid

'n Belangrike kriterium in die evaluering en saamstel van woordeboeke is die konsep van gebruikersvriendelikheid of gebruikersgerigtheid. Kromann, Riiber en Rosbach (1991: 2713) noem hierdie kriterium "the user aspect, which

involves consideration on the part of the lexicographer for the dictionary's target group, its needs and competence, and the types of user situations that occur." Omdat die woordeboek as die belangrikste houer van taalkundige inligting vir die gewone taalgebruiker beskou word (vgl. Gouws 1989: 1), behoort dié inligting op 'n toeganklike wyse vir die gemiddelde woordeboekgebruiker aangebied te word. Volgens Louw en Gouws (1996: 92) is gebruikersgerigtheid en woordeboektoeganklikheid van die eise wat toenemend aan leksikografe gestel word.

Lombard (1994) gebruik die metafoer van 'n driehoek om die verhouding tussen die drie belangrikste rolspelers ten opsigte van woordeboeke (d.i. die leksikograaf, die taalkundige en die woordeboekgebruiker) uit te beeld. In terme van hierdie verhouding verwag die taalkundige onder meer 'n akkurate aanbod van taalkundige inligting, terwyl oor die gebruikersaspek beweer word dat "lexicographers can be and often are guilty of creating ... tension [in this triangle — HLB], inter alia by including features in the dictionary which are pleasing to them but baffling to their users" (Lombard 1994: 205). Hierdie verskynsel kan vermy word deur meer aandag aan gebruikersvriendelikheid in die saamstel van 'n woordeboek te skenk. Hausmann (1989: 13, in Gouws 1999: 7) karakteriseer gebruikersvriendelikheid in woordeboeke dan ook tereg as die aanpassing van die leksikografie by die maatskappy. Vergelyk ook Van der Merwe-Fouché (1999) in hierdie verband.

Die gebruikersaspek geld dan ook as 'n belangrike maatstaf in die evaluering van die buite-mikrostrukturele vernuwings in NW.

3. Die niealfabetiese gedeelte van die woordeboek

Wiegand (1996: 135) beskou woordeboeke as draers van tekstipes of tekste. Met betrekking tot ander teksdraers, soos koerante en tydskrifte, word die ordening en struktuur van die tekste daarin deur historiese en pragmatiese praktyk bepaal. Daar bestaan dus geen onderlinge verband tussen die onderskeie tekste nie. Die verskillende tekste wat egter in 'n woordeboek verskyn (soos die voorwoord, die toeligting, die sentrale woordelys, die woordeboekartikels, e.s.m.) is nie slegs op 'n bepaalde manier georden nie, maar is ook funksioneel in terme van inhoud, en struktureel in terme van vorm aan mekaar verwant, betoog Wiegand. Hierdie kombinasie van tekste noem hy 'n makroteks, en die beginsels waarvolgens die makroteks gestruktureer is, moet in 'n metateks verduidelik word. Die term metateks verwys na die gebruikersgids of toeligting (Wiegand 1996: 136).

Binne Wiegand (1996:136) se teksbenadering kan die tekste in 'n woordeboek aan een van drie tipes behoort: die verpligte voorwerk (waaronder die metateks), die verpligte sentrale woordelys (die makro- en mikrostruktuur) en 'n opsionele agterwerk. Die verband tussen die verskillende tekste word deur 'n min of meer gedifferensieerde woordeboekinterne mediostruktuur bewerkstellig (Wiegand 1996: 135).

Binne hierdie raamwerk en binne die bestek van hierdie artikel (met gebruikersgerigtheid as belangrike kriterium) word leksikografiese vernuwings ten opsigte van die voorwerk en agterwerk van NW vervolgens bespreek, gevolg deur 'n beskouing van vernuwings ten opsigte van die medio-struktuur.

NW se voorwerk bestaan uit 'n Voorwoord, 'n "Key to entries" en 'n lys redaksionele afkortings.

3.1 Die "Key to entries"

Teen die agtergrond van die toenemende bewuswording van die belangrikheid van gebruikersvriendelike geleidsleiding is die uitvoerige "Key to entries" wat betref vertalende woordeboeke in Suid-Afrika 'n beduidende leksikografiese vernuwing. Hierdie verskynsel is 'n verbetering op die tradisionele geleidsleiding in GW, wat bloot uit die aangawe van 'n lys redaksionele afkortings en redaksionele simbole bestaan, en in TFW, wat soortgelyke inligting verskaf, hoewel dit in 'n skematiese formaat aangebied word.

Die toeligting onder die opskrif "Key to entries" word in die vorm van 'n skema bestaande uit voorbeeldartikels aangebied, waarin elkeen van die elemente waaruit artikels in die woordeboek sou kon bestaan, deur middel van tipografiese skadumerkers uitgelig en aan die hand van byskrifte verklaar word. Die lemma ("headword"), aard van homonimiese en polisemiese inskrywings, uitspraakleiding, meervoudsvorme, tempus, etiketaanwending en redaksionele simbole is ingesluit by die verskynsels wat, presies soos hulle in die mikrostruktuur verskyn, as voorbeeldmateriaal aangebied word. Dit maak die opneem van 'n aparte lys redaksionele simbole onnodig en lei die gebruiker meer effektief tot suksesvolle inligtingsontsluiting wanneer hy/sy die alfabetiese gedeelte van die woordeboek raadpleeg. Die verklaring van spesifieke inligtingstipes is ook makliker in dié formaat te vind as die gebruiker vanaf die mikrostruktuur terugverwys na die geleidsleiding om duidelikheid oor 'n bepaalde inskrywing te kry.

Die hoop word uitgespreek dat hierdie gebruikersvriendelike aanslag, wat 'n groot verbetering op die uiters gebrekkige geleidsleiding in GW is, dié verskynsel in al Pharos se toekomstige woordeboeke inlui.

Daar kom wel enkele geringe probleme en inkonsekwentheid in die "Key to entries" in NW voor, wat vervolgens kortliks uitgelig en bespreek word.

In die eerste plek is dit nie duidelik waarom die redaksie besluit het op 'n Engelse aanbod van die byskrifte tot die skema nie, veral gesien in die lig daarvan dat die waarskynlikste gebruiker een is wat vanaf Engels na Afrikaans vertaal, en as sodanig gewoonlik eerder Afrikaanstalig as Engelstalig is. Dié verskynsel is des te meer opvallend as die Voorwoord en lys van redaksionele afkortings, wat beide slegs in Afrikaans verskyn, vergelykend betrek word. Dit is te bevraagteken of hierdie gemengde taalaanbod in die voorwerk 'n gunstige indruk van reëlmatigheid by die gebruiker sal skep.

Kritiek sou ook teen die uitsluitlike gebruik van streng taalkundige terme in die "Key to entries" ingebring kon word. Hoewel die redaksie die waarskynlik tegnieser term *lemma* ten gunste van die meer algemene *headword* vermy het, kom suiwer taalkundige terme soos *homonym* en *collocator* sonder enige verduidelikings in andersins eenvoudig geformuleerde verklarings voor. Dit is duidelik dat die redaksie die rol van die woordeboek as houer van taalkundige inligting, wat akkuraat in taalkundige terme beskryf kan word, erken en toepas. Hoewel aanvaar sou kon word dat die waarskynlikste gebruiker van die woordeboek die taalpraktisyn is wat hierdie terme sonder moeite verstaan, loop sodanige inligtingsaanbod egter wel die risiko om die gemiddelde gebruiker wat reeds gebruikersgidssku is, te vervreem. Een manier waarop hierdie risiko vermy sou kon word, sou wees om verduidelikende parafrases tussen hakies ná sodanige terme in te voeg. Dit sou meebring dat 'n byskrif tot die skema wat tans

Homonyms are marked with superscript numbers

lui (p.^(x)), gewysig sou word tot

Homonyms (words with identical form but unrelated meanings) are marked with superscript numbers.

Op dié wyse word sowel die taalkundig geskoolde as taalkundig ongeskoolde gebruiker in hierdie belangrike gedeelte van die woordeboek geakkommodeer, en funksioneer die parafrase boonop as indirekte taalkundige skolingsgeleentheid ten opsigte van die taalkundig ongeskoolde gebruiker.

Die tipografiese merkers wat aangewend word om doeltaalsinonieme en doeltaalsinoniemparadigmas van mekaar te onderskei, word deur middel van die volgende gesamentlike byskrif by een voorbeeldartikel verklaar:

Target-language synonyms are separated by commas. Translations with distinctly separate meanings are separated by semicolons.

Dit is duidelik dat die redaksie hier wel die gebruik van die meer tegniese lekseem *polyseme* vermy het, wat die inligtingsaanbod meer gebruikersvriendelik maak.

Daar sou egter geargumenteer kon word dat die twee middele wat gebruik word eerder in twee afsonderlike voorbeeldartikels elk met sy eie byskrif verklaar word. Dit sou byvoorbeeld maklik kon gebeur dat die gemiddelde woordeboekgebruiker slegs die eerste drie kommas wat in die voorbeeldartikel uitgelig word as sodanig raaklees, en aanvaar dat die vierde uitgeligte leesteken ook 'n komma is. Verder sou dit ook kon gebeur dat die gebruiker slegs die eerste sin van die verduidelikende byskrif vlugtig lees, en nie die volledige inligting van die tweeslagtige byskrif inneem nie. Hierdie argument

verkry verdere geldigheid daardeur dat al die ander voorbeeldartikels telkens vir die verduideliking van slegs een tipografiese verskynsel in die mikrostruktuur aangewend word. Dit sou ook meer funksioneel wees om tipografiese verskynsels wat gewoonlik in die mikrostruktuur naby mekaar voorkom, juis apart in die gebruikersgids te verduidelik, aangesien elkeen van dié verskynsels 'n eiesoortige funksie met eiesoortige implikasies vir die gebruiker veronderstel.

Indien die inligting oor die onderskeid tussen doeltaalsinonieme en doeltaalsinoniemparadigmas dan apart aangebied word, maar saamgegroepeer is in die gebruikersgids, sou dit ook doelmatig kon wees om die inligting oor hoe homonieme van mekaar onderskei word hiermee saam te groepeer, sodat alle eksplisiete inligting op grond van betekenis- en ekwivalentdiskriminasie as 'n eenheid aan die gebruiker gebied word. Tans is die inligting oor die onderskeid tussen homonieme die tweede inligtingstipe wat in die gebruikersgids aange-
toon word, terwyl die inligting oor die aanduiding van sinonieme en poliseme sewentiende (tweede laaste) geplaas is. 'n Meer konseptuele organisasie wat betref die aanbod van inligting oor artikelstruktuur sou dus aanbeveel kon word.

Ten minste een tipografiese teken wat in die mikrostruktuur van die woordeboek verskyn, is nie in die "Key to entries" opgeneem nie, naamlik die gelykheidstekens (=). Dié teken kom onder meer in die volgende artikels voor:

- (1) **ad·min** *n.*, (*infml. afk.*) = **administration.**
- (2) **X·mas** *n.*, (*infml.*) = **Christmas.**
- (3) **li·cense** *n.*: ~ **plate** (*Am.*) = **numberplate.**

Die hantering van die lemmas hierbo suggereer 'n tipe kruisverwysing. Wat hier van belang is, is dat 'n redaksionele simbool aangewend word sonder dat dié simbool se betekenis in die gebruikersleiding verklaar word. Dit is nie duidelik of die simbool se gebruik enige makrostrukturele, mikrostrukturele of mediostrukturele implikasies vir die gebruiker (en die leksikograaf!) inhou nie. Boonop is dit nie 'n simbool wat gewoonlik in tweetalige woordeboeke soos GW verskyn nie. Hierdie verskynsel sal verder by die bespreking van die mediostruktuur aandag geniet (vgl. 4.2).

'n Volgende probleem by die "Key to entries" is die voorbeelde wat ten opsigte van 'n "idiomatic phrase" en 'n "usage example" onderskeidelik aangebied word. Uit die aangawe van dié voorbeelde kan die gebruiker nie aflei wat die verskil tussen die aanbod van 'n idiomatiese uitdrukking en dié van 'n gebruiksvoorbeeld is nie, aangesien die twee mikrostrukturele aangawes op dié vlak tipografies identies is. Vergelyk (4) as die gebruiksvoorbeeldaangawe en (5) as die aangawe oor idiomatiese uitdrukkings:

- (4) **mov·ie** *n.*: *go to the ~s* gaan flik.
- (5) **wob·bly** *n.*: *throw a ~*, (*Br. infml.*) die piep kry, die (aap/bobbejaan)stuipe kry, ontplof.

Die enigste aanduiding vir die gebruiker dat die lemma in (5) in 'n idiomatiese uitdrukking voorkom, is die feit dat die Afrikaanse ekwivalent in die vorm van 'n idiomatiese uitdrukking aangebied word. Die idiomatiese uitdrukking word op geen ander manier (soos d.m.v. 'n rubriekopskrif) van die gewone gebruiksvoorbeeld in (4) onderskei nie. Die toeligting bevat dus geen inligting oor hoe idiomatiese uitdrukkinge op watter manier ook al spesiale leksikografiese bewerking ondergaan nie, of dat brontaaluitdrukkinge van doeltaaluitdrukkinge (in plaas van bv. doeltaalparafrases) as vertaalekwivalente voorsien word nie. Dit is onvoldoende gebruiksheidsleiding.

Mikrostrukturele konvensies wat goed in die "Key to entries" verantwoord word, is die gebruik van die skuinsstreep en van hakies. Met betrekking tot die skuinsstreep word die volgende artikel as voorbeeld in die gebruikersgids aangewend:

(6) a·l*it*·er·ate *n.* tra*e*/onwillige leser

Daarnaas verskyn die volgende byskrif:

(7) Strokes separate alternative words: *trae leser* or *onwillige leser*.

Met betrekking tot die mikrostrukturele optrede van hakies word die volgende artikel as voorbeeld in die gebruikersgids aangewend:

(8) **h*o*, h*o*(, h*o*)** *tw.*, (*nabootsing v. laggeluid*) h*ô*, h*ô*(, h*ô*).

Die finale inskrywing in dié artikel ((, h*ô*)) is tipografies gemarkeer en dien as voorbeeld by die daaropvolgende byskrif:

(9) Letters or words in brackets are optional: *h*ô*, h*ô** or *h*ô*, h*ô*, h*ô**.

Die duidelike uitspel van die konvensies in (7) en (9) het tot gevolg dat die gebruiker maklik die inligting in 'n artikel soos dié in (5) korrek kan interpreteer en toepas om alternatiewe (10), (11) en (12) as vertaalekwivalente tot sy/haar beskikking te hê sonder dat dit vir die leksikograaf nodig was om elkeen afsonderlik aan te bied:

(10) die stuipe kry

(11) die aapstuipe kry

(12) die bobbejaanstuipe kry

Die relatiewe plasing van die hakies en die vorme wat daardeur in- of uitgesluit word, maak ook die ortografiese implikasies verbonde aan die alternatiewe vorme duidelik. Op dié manier is dit byvoorbeeld duidelik ten opsigte van (8) dat die komma na die tweede *h*ô** (slegs) nodig is indien 'n derde *h*ô**

bygevoeg word. Met betrekking tot die los of vas skryf van woorde is dit duidelik dat die plasing van die vorme tussen hakies in (5) (*aap/bobbejaan*) téénaan die daaropvolgende vorm (*stuipe*) impliseer dat die saamgestelde vorm as een woord geskryf word, dus *aapstuipe* en *bobbejaanstuipe* teenoor die foutiewe *aap stuipe* en *bobbejaan stuipe*.

Ander inligtingstipes wat duidelik aan die gebruiker in die toeligting gestel word, is uitspraakleiding en lettergreepaanduiding (vgl. die bespreking oor uitspraakleiding onder 4.1), woordklas, meervoude, tempus, die aanwending van die tilde, die dubbel- en enkelkoppelteken, die kruisverwysingsmerker, betekenisindikatorters (aanvullende semantiese kommentaar), etikette en kollokators.

Die vernuwende manier waarop die "Key to entries" in NW saamgestel is, verteenwoordig baanbrekerswerk in standaard Afrikaanse vertalende woordeboeke. Die leksikograaf gee nie alleen verantwoording van sy/haar mikrostrukturele struktuurplan nie, maar doen dit ook op 'n maklik verstaanbare en geredelik toeganklike manier. So 'n gebruikersvriendelike aanbod staan 'n groter kans om die gebruiker werklik in sy/haar naslaanaktiwiteit van hulp te wees deur die mikrostruktuur toeganklik te maak. Verder kan die eenvoudige aanbod ook die gemiddelde gebruiker se ignorering van die woordeboek se voorwerk tot 'n mate begin teenwerk, wat weer indirek kan bydra tot 'n uitbreiding van die gebruiker se inligtingsontsluitingsvaardighede deur kennisname van die inligtingstipes wat in 'n woordeboekartikel vervat word.

Dit is wel noodsaaklik dat alle beginsels en middele wat in die woordeboek geld en aangewend word, duidelik en volledig in die toeligting verantwoord word. By gebrek hieraan ly die woordeboek se toeganklikheid skade.

3.2 Die lys redaksionele afkortings

Die lys redaksionele afkortings in NW is verteenwoordigend van die tradisionele aanbod van sodanige afkortings in Afrikaanse tweetalige woordeboeke (vgl. GW).

Enkele tegniese tekortkominge bestaan in hierdie deel van die voorwerk. Die afkorting *v.* staan volgens die lys vir *verbum* (*werkwoord*), maar die afkorting word in die mikrostruktuur ook as afkorting vir die setsel *van* in semantiese kommentaar aangewend, soos wat uit die onderstaande uittreksel blyk:

- (13) *crä-que-lure n.*, (*netwerk v. fyn krakies/barsies in verf/vernis/glasuur*) krakeluur.

Die aanwending van die afkorting *verb.* vir *verbum* (*werkwoord*) in plaas van *v.* kan die probleem uit die weg ruim.

Ten minste een Afrikaanse redaksionele afkorting wat in die mikrostruktuur voorkom, verskyn nie in die lys redaksionele afkortings nie, naamlik die

afkorting *iem.* (vir *iemand*), soos uit die onderstaande uittreksel blyk:

- (14) *a·bun·dant·ly adv.: be ~ clear (to s.o.) that ...* (vir *iem.*) glashelder/alte duidelik wees dat ... [...].

Uit die uittreksel in (14) hierbo val ook die gebruik van die Engelse redaksionele afkorting *s.o.* (vir *someone*) op. Hier faal die lys van redaksionele afkortings in NW daarin dat geen Engelse redaksionele afkortings in die lys opgeneem is nie. In die lig van die vroeëre vraag na die rede vir die redaksie se besluit om die "Key to entries" se byskrifte in Engels aan te bied, wil dit tog begin voorkom asof die algehele aanbod van die voorwerk in NW by 'n heroerweging baat sou vind. Die Engelse aanbod in die "Key to entries" staan in skerp kontras tot die totale afwesigheid van die verklaring van Engelse redaksionele afkortings in die lys redaksionele afkortings. Dit lyk asof noukeuriger kontrole oor die aanwending en opname van redaksionele afkortings nodig mag wees.

'n Oplossing in hierdie verband mag 'n voorstel wees dat die lys redaksionele afkortings onderverdeel word in Afrikaanse teenoor Engelse redaksionele afkortings enersyds, en andersyds in redaksionele afkortings wat bloot ruimtebesparend optree (bv. *bv.*, *v.d.*) teenoor gebruiks-, vaktaal- en styletiekette (bv. *akr.*, *geol.*, *sl.*). Sodanige onderverdeling het die tweeledige voordeel dat die gebruiker in een oogopslag van die aard van die relevante inligtingstipes bewus gemaak word én dat dit die leksikograaf minder ruimte laat vir vergissing ten opsigte van duplisering of ander foute wat betref afkortings.

Soos vroeër opgemerk, is dit duidelik dat die redaksie van NW deeglik van die taalkundige inligtingsaspek van die woordeboek bewus is (vgl. ook weer Lombard 1994). Hierdie bewustheid word verder gemanifesteer in die aanbod van woordklas in taalkundig eenduidiger terme, soos *adjektief*, *adverbium*, *nomen*, *prefiks*, *preposisie*, *suffiks* en *verbum*. 'n Enkele potensieel verwarrende nalatigheid is dat beide die terme *verbum* en *werkwoord* apart in die lys verskyn, terwyl slegs *verbum* in die mikrostruktuur ter sprake is. Dié term word egter ook tussen hakies as werkwoord gekwalifiseer. Die taalkundig eenduidiger terme wat in die NW se lys redaksionele afkortings opgeneem is, is egter ook toeganklik vir die taalkundig ongeskoolde gebruiker deurdat die tradisionele woordklasaanduidings tussen hakies na die taalkundige terme volg. Vergelyk:

- (15) (a) adverbium (bywoord)
 (b) prefiks (voorvoegsel)
 (c) preposisie (voorsetsel)

Dit is duidelik dat die redaksie sowel die taalkundig geskoolde gebruiker as die taalkundige leek by hierdie inklusiewe aanbod van taalkundige inligting probeer akkommodeer. Dit verteenwoordig gewis 'n nuwe rigting in die Afrikaanse leksikografie — een wat prysenswaardig is.

3.3 Die agterwerk

In die agterwerk van NW verskyn 'n glossarium van algemene afkortings. Binne Wiegand (1996) se teksbenadering is die glossarium as deel van die agterwerk opsioneel; die afkortings sou dus ook op hul alfabetiese plekke in die makrostruktuur opgeneem en bewerk kon word. Die buitemikrostrukturele leksikografiese hantering van dié inligtingstipe verhoog egter die toeganklikheid daartoe, aangesien individuele items makliker in die glossariumformaat gevind kan word omdat hulle nie van uitgebreide bewerkings voorsien hoef te word nie. In GW verskyn soortgelyke glossaria aan die einde van die Afrikaans-Engels-deel en aan die einde van die Engels-Afrikaans-deel onderskeidelik. Aangesien NW nie 'n tweerigting vertalende woordeboek is nie, word die Engelse volvorme van die aangebode afkortings gegee, gevolg deur die Afrikaanse volvorme en die korresponderende Afrikaanse afkortings, waar van toepassing. Uit die aard van die woordeboek is 'n omgekeerde aanbod nie ingesluit nie. Vergelyk (16) hieronder:

- (16) (a) **prop.** proposition • aanbod aanb.
 (b) **Q, Q.** question • vraag

In (16)(b) is die alternatiewe afkortingsmoontlikhede vir *question* (Q óf Q.) duidelik, terwyl die afwesigheid van 'n Afrikaanse afkorting in dié geval impliseer dat daar nie in Afrikaans 'n afkorting vir *vraag* is nie. Ook hierdie tipe volledige inligtingsaanbod word ongelukkig nie konsekwent deurgevoer nie. Vergelyk die uittreksels in (17), waar geen Afrikaanse ekwivalent (volvorm óf afkorting) hoegenaamd aangebied word nie:

- (17) (a) **rec.** recipe
 (b) **rec.** record
 (c) **WP, w.p.** weather permitting

Die rede vir die onvolledige bewerking hierbo is nie afleibaar nie. In gevalle waar daar nie 'n Afrikaanse afkorting bestaan nie, word die Afrikaanse volvorm wel normaalweg gegee, soos in (16)(b) en in (18) hieronder:

- (18) (a) **VTR** video tape recorder • videobandopnemer
 (b) **w/o** without • sonder
 (c) **XL** extra large • ekstra groot

'n Vernuwing in die glossarium afkortings in NW is die verskynsel van etikettering van sommige afkortings. In hierdie instansie is NW 'n goeie voorbeeld van hoe die inligting in die toeligting in die voorwerk van 'n woordeboek ook vir die agterwerk van die woordeboek kan geld. Verder word 'n mate van kontinuïteit in terme van inligtingsaanbod tussen die alfabetiese deel van die

woordeboek en die agterwerk bewerkstellig. Hierdie verskynsel onderskryf Wiegand (1996: 135) se siening dat al die tekste binne die woordeboek op funksionele en strukturele vlak met mekaar verband hou. Vergelyk die geëtiketteerde afkortings in (19) hieronder:

- (19) (a) Pvt. (*Am. mil.*) private • weerman wrn.
 (b) RAM (*rek.*) random-access memory • lees-en-skryf-geheue LSG,
 RAM
 (c) TGIF (*infml.*) thank God it's Friday • dank Vader dis Vrydag DVDV

Dit is wel jammer dat geen melding van die agterwerk in die "Key to entries" (of desnoods in die Voorwoord) gemaak word nie. Hierdeur ly tekstuele eenheid binne die makroteks skade, en die gebruiker word nie ingelig dat sodanige agterwerk (en dus die betrokke inligtingstipe) wel bestaan nie.

4. Die alfabetiese gedeelte van die woordeboek

In hierdie afdeling sal die mikrostrukturele element van uitspraakleiding en die mediostruktuur in terme van leksikografiese vernuwings in NW oorsigtelik aandag geniet.

4.1 Uitspraakleiding

Daar bestaan uiteenlopende standpunte oor die aanbod van uitspraak-inligting in vertalende woordeboeke. Wells (1985: 45) benader die aanbod van uitspraak-inligting vanuit 'n pragmatiese perspektief: "There is little need for such [i.e. pronunciation — HLB] information in a dictionary of Spanish or Finnish, for example, since in all except the rarest cases the pronunciation of a word in those languages may be safely inferred from its spelling, given the knowledge of the relevant 'reading rules' which express the relationship between phoneme and grapheme, speech and writing." Uitspraak-inligting is wel nodig indien 'n taal soos Engels, waar die verband tussen ortografie en uitspraak nie so duidelik is nie, deel van die woordeboek uitmaak, voer Wells aan. Svensén (1993: 69) is meer uitgesproke oor uitspraakleiding en betoog dat "it has to be given in dictionaries intended for passive use". Die aard van die verhouding tussen ortografie en uitspraak is nie vir Svensén 'n oorweging nie; dit gaan oor die funksie van die woordeboek. Svensén (1993: 70-73) is egter nie baie krities oor hoe uitspraak-inligting aangebied word nie: dit kan volgens hom óf as deel van die lemma, óf apart van die lemma, óf as deel van én apart van die lemma aangebied word. Fonetiese transkripsie is dus nie 'n vereiste nie. Oor die opmerlike afwesigheid van fonetiese transkripsies in Afrikaanse vertalende woordeboeke handhaaf Sakurai en Kotzé (1997: 169) die standpunt dat die rede daarvoor voor die hand liggend is: "Dit word as onnodig beskou vir die woor-

deboekgebruikers, almal Suid-Afrikaners wat elke dag ouditief aan Afrikaans blootgestel word." Dié argument kan egter nie sonder meer aanvaar word nie. Dit sou waarskynlik akkurater wees om te beweer dat die meeste gebruikers van Afrikaanse vertalende woordeboeke Suid-Afrikaners is, en dat hierdie (en nie-Suid-Afrikaanse) gebruikers in verskillende grade (indien enigsins) ouditief aan Afrikaans blootgestel word, en boonop nie noodwendig elke dag nie.

Gouws (1989: 265) lewer die felste kritiek wanneer hy beweer dat uitspraak 'n inligtingstipe is wat nog nie met genoeg sorg en uitvoerigheid in Afrikaanse woordeboeke hanteer word nie: "Dit is veral in die vertalende woordeboeke dat uitspraak-inligting uiters onbevredigend aangebied word. Sowel die GW (*Groot Woordeboek/Major Dictionary* — HLB) as die TW (*Tweetalige Woordeboek/Bilingual Dictionary* — HLB) bied slegs die hoofklem van lemmas aan." Uit 'n taalkundige oogpunt bly 'n fonetiese transkripsie die suksesvolste manier van uitspraakleiding: "Dit skep ook die ideale terrein vir die toepassing van ander uitspraak-inligtingstipes soos klemaanduiding en sillabeverdeling" (Gouws 1989: 268). Ongelukkig is 'n stelsel van fonetiese transkripsie ook in NW afwesig.

Hoewel klemaanduiding in NW steeds ten opsigte van die geskrewe vorm van die lemma verstrekkend word, en nie volgens die klanklike weergawe daarvan nie (vgl. Gouws 1989: 266), tree NW wel daarin vernuwend op dat klemaanduiding geskied deur die onderstreping van die sillabe waarop die hoofklem in 'n meersillabige lemma val. Vergelyk die voorbeelde in (20) hieronder:

- (20) (a) **ab·sen·tee** *n.*: [...]
 (b) **cur·tain** *n.*: [...]
 (c) **vol·au·vent** *n.*: [...]

Hierdie wyse van klemaanduiding, hoewel nog nie die ideale vorm nie, is verkieslik bo die gebruik van klemstrepies, soos wat in GW aangewend word, aangesien die plasing daarvan verwarring by die gebruiker kan veroorsaak in die lig van die feit dat verskillende woordeboeke verskillende benaderings volg: sommiges plaas die strepie voor die sillabe wat die hoofklem dra; ander plaas die strepie daarna. Vergelyk klemaanduiding by die lemma *adult* in GW, wat effektief geen leiding ten opsigte van die korrekte uitspraak van dié item bied tensy die gebruiker seker is van die betrokke woordeboek se stelsel van klemaanduiding nie:

- (21) **ad'ult**

Deur die sillabe waarop die hoofklem val te onderstreep, word enige moontlike verwarring uitgeskakel. Hierdie praktyk getuig weer eens van die redaksie se sterk gebruikersvriendelike aanslag in NW.

Hoewel die aanduiding van lettergrepe (teenoor sillabes) nie eintlik tuis hoort onder uitspraakleiding in 'n woordeboek nie, kan dit hier genoem word

dat hierdie inligtingstipe glad nie in GW voorkom nie, terwyl ten opsigte van die NW van 'n volledige en welkome vernuwing op hierdie gebied sprake is. Lettergreepgrense word naamlik deur kolletjies binne die lemma aangedui (vgl. o.a. die voorbeelde in (20)).

Ongelukkig is dit nie die ideale taalkundige toestand dat lettergreepverdeling en hoofklemaanduiding ten opsigte van een en dieselfde weergawe van die lemma aangedui word nie. Die gebruiker kry in die afwesigheid van 'n fonetiese transkripsie ook steeds onvoldoende uitspraakleiding wanneer hy / sy voor 'n lemma soos *vol-au-vent* te staan kom. Vanuit 'n meer pragmatiese oogpunt moet die redaksie van NW egter gelukkig word met die vars en gebruikersvriendelike aanbod van tradisionele uitspraakinligting in Afrikaanse vertalende woordeboeke en die byvoeging van inligtingstipes wat tot dusver in geen standaard Afrikaanse vertalende woordeboek voorkom nie.

Indien NW egter as meer verantwoordbare houër van taalkundige inligting wil geld (soos wat die reeds aangeduide vernuwings suggereer), is die insluit van 'n stelsel van fonetiese transkripsie noodsaaklik. Dit is tog jammer dat die redaksie van NW nie ook hierdie vernuwende stap kon neem nie. Volgens Louw (1997: 85) sou fonetiese transkripsies nie die toeganklikheid of gebruikersvriendelikheid van die woordeboek aantas nie: "Die aanpassing wat die aanleer of gebruik van die fonetiese alfabet van die deursneegebruiker sal verg, is nie veel groter as wat die aanleer en gebruik van byvoorbeeld RD (*Reader's Digest Afrikaans-Engelse Woordeboek/English-Afrikaans Dictionary* — HLB) se sleutels sal verg nie. Verder word die bate van ondubbelsinnigheid toegevoeg ... Transkripsies volgens 'n fonetiese alfabet [is] die beste opsie, mits 'n deeglike, toeganklike gids gegee word verkieslik in die voorwerk." Boonop sou oortuigend geargumenteer kon word dat die teikengebruiker van NW ook (ten minste by geleentheid) Engelse verklarende woordeboeke raadpleeg. In die lig van die feit dat Engelse verklarende woordeboeke byna sonder uitsondering fonetiese transkripsies in uitspraakleiding aanwend, behoort sodanige gebruiker nie verlore te voel indien fonetiese transkripsies dan in 'n Afrikaanse vertalende woordeboek sou verskyn nie.

Omdat NW 'n eenrigting vertalende woordeboek is, sal by die insluit van fonetiese transkripsie oor die adresseringsprosedure daarvan besin moet word. Gegee dat die waarskynlikste gebruiker van NW Afrikaanssprekend is, soos afgelei kan word uit die huidige Suid-Afrikaanse sosiolinguïstiese konteks tesame met die algemene gerigtheid van die woordeboek, sou 'n stelsel van lemmatiese adressing voorgestel word.

4.2 Mediostruktuur

Volgens Gouws en Prinsloo (1998: 18) is een van die belangrikste eienskappe van 'n goeie woordeboek sy toeganklikheid. Toeganklikheid lei tot eenduidige inligtingsontsluiting op beide die makro- en mikrostrukturele vlakke. Toeganklikheid in NW word verhoog deur die implementering van 'n stelsel van eks-

plisiete kruisverwysing. Volgens Svensén (1993: 194) gebruik baie woordeboeke eksplisiete kruisverwysings op 'n lukraak wyse deur geen of gebrekkige leiding vir verwysingsmerkers te bied. In NW word die kruisverwysingsmerker → egter in die toeligting ("Key to entries") in die voorwerk opgeneem en verduidelik aan die hand van die volgende voorbeeldartikel:

(22) **ec.o.dis.as.ter** *n.* → **ecocatastrophe**.

Naas dié voorbeeldartikel verskyn die volgende byskrif:

(23) Cross-reference: Look for the information at the word that comes after the arrow.

Die gebruiker wat die toeligting raadpleeg, sal dadelik die betekenis van die pyltjie in 'n woordeboekartikel snap en die nodige inligting maklik bekom deur die kruisverwysingsverband op te neem. Vergelyk die volgende NW-artikels waarin eksplisiete kruisverwysing voorkom:

- (24) (a) **bald** *a.*: ~ **eagle** → **eagle**.
 (b) **eag.le** *n.*: **bald** ~ Amerikaanse adelaar.
- (25) (a) **so.pra.no** [...] ~ **recorder** (*Am.*) → **descant recorder**.
 (b) **des.cant, dis.cant** *n. & a.*: ~ **recorder sopraanblokfluit**.

Suksesvolle verwysing is tussen die verwysingsinskrywing in (24) (a) en (25)(a) en die verwysingsadres in (24)(b) en (25)(b) onderskeidelik bewerkstellig. Dit is 'n tegniek wat konsekwent in NW toegepas word, en wat 'n merkbare verbetering op GW se mediostruktuur verteenwoordig.

Daar kom egter een verwysingsmerker in die mikrostruktuur van NW voor wat nie in die toeligting verduidelik word nie, naamlik die gelykheidsteken (=) (vgl. (1), (2) en (3) onder 3.1). In hierdie gevalle bestaan die verwysingsadres nie, aangesien die items wat op die gelykheidsteken volg, nie in die makro- of mikrostruktuur opgeneem is nie. Hoewel die gebruik van die gelykheidsteken 'n geval van wisselvormlemmas suggereer in terme van Gouws (1989: 76) en Svensén (1993: 195), negeer die afwesigheid van die verwysingsadres dié aanname. Hier faal NW daarin dat daar uiteindelik geen Afrikaanse ekwivalent vir 'n Engelse lemma aangebied word nie, maar dat die gebruiker bloot na 'n niebestaande lemma verwys word. Dit wil selfs voorkom asof die Engelse wisselvorm wat op die gelykheidsteken volg as 'n tipe ekwivalent aangebied word. Die leksikograaf moet sistematies te werk gaan en kan nie van die gebruiker verwag om van sy/haar intuïsie gebruik te maak indien die aanwending van spesiale simbole in die mikrostruktuur nêrens in die toeligting verklaar word nie.

Aangesien NW as 'n aanvullende woordeboek ten opsigte van veral GW geld, sou geargumenteer kon word dat kruisverwysing deur die gelykheids-

teken 'n spesiale soort eksterne verwysingsadres (ekstrawoordeboek-verwysingsadres) aandui, naamlik 'n lemma wat nie in NW nie, maar in GW verskyn. Wiegand (1996: 141) maak in sy uiteensetting van mediostrukture binne die teksbenadering voorsiening vir intertekstuele mediostrukture, waaronder mediostrukture wat 'n netwerk tussen verskillende woordeboeke bewerkstellig, ressorteer. Binne die relatief spesiale funksie van NW sou dus tereg van sodanige intertekstuele mediostruktuur sprake kon wees, maar die feit bly staan dat dié gebrekkig verklaarde reëling die gebruiker in die duister kan laat en sodoende die toeganklikheid van die woordeboek skade berokken.

Afgesien van die bogenoemde tekortkominge lewer NW welkome vernuwings ten opsigte van die mediostruktuur in Afrikaanse vertalende woordeboeke. Die redaksie sou daarby kon baat om noukeuriger kontrole uit te oefen oor die presiese inhoud van die toeligtig in die voorwerk en hoe dit vergelyk met wat in die mikrostruktuur aangetref word. Uiteraard is die volledige verklaring van alle tipografiese verskynsels in die mikrostruktuur die ideaal.

5. Perspektief

Hoewel NW ongetwyfeld verdere leksikografiese vernuwings op mikrostrukturele vlak bevat, is in hierdie artikel hoofsaaklik gefokus op aspekte wat miskien nie altyd die aandag ontvang nie, en vanweë hulle relatiewe agtergesteldheid gevolglik die oorhoofse kwaliteit van die woordeboek deur beperkte toeganklikheid nadelig kan beïnvloed. Die redaksie van NW het bewys dat buitemikrostrukturele leksikografiese vernuwings 'n baie belangrike rol in die gebruikersvriendelik maak van 'n woordboek kan en behoort te speel.

Verwysings

Woordeboeke

- Bosman, D.B., I.W. van der Merwe en L.W. Hiemstra. 1984⁸. *Tweetalige woordeboek/Bilingual Dictionary*. Kaapstad: Tafelberg.
- Eksteen L.C. 1997¹⁴. *Groot Woordeboek/Major Dictionary*. Kaapstad: Pharos.
- Grobbelaar, P. (Red.). 1987. *Reader's Digest Afrikaans-Engelse Woordeboek/English-Afrikaans Dictionary*. Kaapstad: The Reader's Digest Association.
- Joubert, P.A. 1997. *Tweetalige Frasewoordeboek/Bilingual Phrase Dictionary*. Kaapstad: Pharos.

Ander bronne

- Gouws, R.H. 1989. *Leksikografie*. Pretoria en Kaapstad: Academica.
- Gouws, Rufus H. 1996. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 14-31.

- Gouws, R.H. 1999. *Die maatskaplike gerigtheid van die metaleksikografie in 'n meertalige samelewing*. Intreerede. Stellenbosch: Universiteit van Stellenbosch.
- Gouws, R.H. en D.J. Prinsloo. 1998. Cross-Referencing as a Lexicographic Device. *Lexikos* 8: 17-36.
- Hausmann, F.J. 1989. Die gesellschaftlichen Aufgaben der Lexikographie in Geschichte und Gegenwart. Hausmann, F.J. et al. (Reds.). 1989-1991: 1-19.
- Hausmann, F.J. et al. (Reds.). 1989. *Wörterbücher. Ein internationales Handbuch zur Lexikographie / Dictionaries. An International Encyclopedia of Lexicography / Dictionnaires. Encyclopédie internationale de lexicographie*. Berlyn/New York: Walter de Gruyter.
- Kroimann, Hans, Theis Riiher en Poul Rosbach. 1991. Principles of Bilingual Lexicography. Hausmann, F.J. et al. (Reds.). 1989-1991: 2711-2728.
- Lombard, F.J. 1994. Lexicographer, Linguist and Language User: An Uneasy Triangle? *Lexikos* 4: 204-214.
- Louw, Phillipus Adriaan. 1997. *Kriteria vir 'n standaard vertalende woordeboek*. Ongepubliseerde M.A.-skripsie. Stellenbosch: Universiteit van Stellenbosch.
- Louw, Phillip Adriaan en Rufus Hjalmar Gouws. 1996. Lemmatiese en nielemmatiese adressering in Afrikaanse vertalende woordeboeke. *Suid-Afrikaanse Tydskrif vir Taalkunde* 14(3): 92-100.
- Sakurai, Takashi en Ernst F. Kotzé. 1997. 'n Kontrastiewe beskouing van tweetalige woordeboeke in Suid-Afrika en Japan. *Lexikos* 7: 165-178.
- Svensén, Bo. 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- Van der Merwe-Fouché, Michele. 1999. *Gebruikersperspektief in die leksikografie*. Ongepubliseerde D.Litt.-proefskrif. Stellenbosch: Universiteit van Stellenbosch.
- Wells, J.C. 1985. English Pronunciation and its Dictionary Representation. Ilson, Robert (Red.). 1985. *Dictionaries, Lexicography and Language Learning*. Oxford: Pergamon.
- Wiegand, Herbert Ernst. 1996. A Theory of Lexicographic Texts: An Overview. *Suid-Afrikaanse Tydskrif vir Taalkunde* 14(4): 134-149.

Guide Structures in CD-ROM Dictionaries, with Specific Reference to the EWED and the ELHAT

Phillip Louw, *Bureau of the Woordeboek van die Afrikaanse Taal, Stellenbosch, Republic of South Africa*

Abstract: Much of the success of modern dictionaries depends on the accessibility of their macro- and microstructural information. In order to evaluate the methods which lexicographers employ to improve dictionary accessibility, the concept of guide structures was introduced by Hausmann and Wiegand (1989). Since the appearance of that article, various academics have written articles and read papers on guide structures in print dictionaries, but few studies have focussed on the tremendous potential of guide structures in electronic dictionaries.

In this article it will be shown that electronic dictionaries add dimensions to the implementation of guide structures not possible in print versions. The role of two of the guide structures in the transfer of information in CD-ROM dictionaries will be discussed. The access and mediostructures of the *Microsoft Encarta World English Dictionary* and the *Elektroniese Handwoordeboek van die Afrikaanse Taal* will be analysed and critically evaluated to illustrate the advantages of electronic dictionaries over print dictionaries and the urgent need for metalexical discussion of this publication medium.

Keywords: ACCESS STRUCTURE, ACTIVE CROSS-REFERENCES, CD-ROM DICTIONARY, DICTIONARY ACCESSIBILITY, ENCYCLOPAEDIC DICTIONARY, GUIDE STRUCTURES, INNER ACCESS STRUCTURE, INNER SEARCH PATH, INTERNET, MEDIOSTRUCTURE, MULTIMEDIA, OUTER ACCESS STRUCTURE, OUTER SEARCH PATH, PASSIVE CROSS-REFERENCING, USER-FRIENDLINESS

Opsomming: Gidsstrukture in CD-ROM-woordeboeke, met spesifieke verwysing na EWED en ELHAT. Die sukses van moderne woordeboeke hang grootliks af van die toeganklikheid van hulle makro- en mikrostrukturele inligting. Hausmann en Wiegand (1989) het die konsep van gidsstrukture ingevoer om die metodes te evalueer wat leksikograwe gebruik om woordeboektoeganklikheid te verbeter. Sedert hierdie artikel verskyn het, het verskeie akademiese artikels gepubliseer en referate gelewer oor gidsstrukture in gedrukte woordeboeke, maar min studies het gefokus op die geweldige potensiaal van gidsstrukture in elektroniese woordeboeke.

In hierdie artikel sal dit getoon word dat elektroniese woordeboeke geleentheid vir die implementering van gidsstrukture skep wat nie in gedrukte weergawes moontlik is nie. Twee van die gidsstrukture se rol in die oordrag van inligting in CD-ROM-woordeboeke sal bespreek word. Die toegang- en mediostrukture van die *Microsoft Encarta World English Dictionary* en die *Elektroniese Handwoordeboek van die Afrikaanse Taal* sal ontleed en krities geanaliseer word, om die voordele

van elektroniese bo gedrukte woordeboeke en die dringende behoefte aan die metaleksikografiese bespreking van hierdie publikasie-medium te illustreer.

Sleutelwoorde: AKTIEWE KRUISVERWYSINGS, CD-ROM-ENSIKLOPEDIE, EKSTERNE SOEKROETE, EKSTERNE TOEGANGSTRUKTUUR, ENSIKLOPEDIËSE WOORDEBOEK, GEBRUIKERSVRIENDELIKHEID, GIDSSTRUKTURE, INTERNE SOEKROETE, INTERNE TOEGANGSTRUKTUUR, INTERNET, MEDIOSTRUKTUUR, MULTIMEDIA, PASSIEWE KRUISVERWYSINGS, TOEGANGSTRUKTUUR, WOORDEBOEKTOEGANKLIKHEID.

1. Introduction

To many dictionary users the consultation of dictionaries — even reputable ones — can often be a labyrinthine affair. In order to evaluate the methods which lexicographers employ to lead the users through these labyrinths, the concept of guide structures was introduced by Hausmann and Wiegand (1989). This subset of structures is a key element of the structural theory, which, in turn, is a central part of Wiegand's textual theory for lexicographic texts. Since the appearance of that article, various academics have written articles and read papers on guide structures in print dictionaries, but few studies have focussed on the tremendous potential of guide structures in electronic dictionaries. However, Louw (1999a) has shown that the theory of guide structures can be applied without much difficulty to the study of electronic dictionaries. In fact, electronic dictionaries add dimensions to the implementation of guide structures not possible in print versions.

Two of the roles of the guide structures in the transfer of information in CD-ROM dictionaries will be discussed. The access and mediostructures of the *Microsoft Encarta World English Dictionary* (henceforth EWED) and the *Elektroniese Handwoordeboek van die Afrikaanse Taal* (henceforth ELHAT) will be analysed and critically evaluated to illustrate the advantages electronic dictionaries hold over print dictionaries and the urgent need for metalexical discussion of this publication medium.

2. Access structures in electronic dictionaries

Louw (1999: 109) describes the access structure as "the primary guide structure". It is through the use of the access structure of any dictionary that the necessary information is found. The macro- and microstructural content of a dictionary may be adequate or even excellent, but without an adequate access structure the user may very well be frustrated in his/her search. Gouws and Prinsloo (1998: 18) affirm this point: "One of the most important characteristics of a good dictionary is its accessibility. Accessibility leads to an unambiguous retrieval of the information presented on both the macro- and microstructural levels."

Louw (1999: 109) adds: "The access structure is, however, not a singular entity: It encompasses certain substructures, of which the outer and inner access structures are the most important." In the dictionaries under discussion, there is innovation in both the outer and inner access structures.

2.1 The outer access structure

2.1.1 The outer search path

In electronic dictionaries, as in print dictionaries, the user follows an outer search path which takes him/her to the desired article and is aided on this search path by the outer access structure. The outer access structure can be divided into a rapid outer access structure and a standard outer access structure.

2.1.2 The rapid and standard outer access structures

In print dictionaries, the rapid outer access structure usually consists of elements such as indexes, thumb indexes, running heads and any features which bring the user rapidly to the vicinity of the article in which the relevant information is presented. Louw (1999: 111), however, states: "It cannot guide the user along the last, most vital part of the outer search path. The final choice of the desired macrostructural element is made by means of the standard outer access structure." Most print dictionaries are monoaccessible, i.e. the standard outer access structure corresponds to the macrostructure. This standard outer access structure encompasses decisions made by the lexicographer on, for example, an initial alphabetical or a thematic listing of lemmata, as well as the consistent implementation of these decisions.

In electronic dictionaries, it is not always easy to recognise the boundaries between the rapid and standard outer access structure. In the ELHAT there is a traditional presentation similar to that of the print version of this dictionary, but also some innovative hybrid elements. In the EWED innovative methods are employed throughout, combining elements of the rapid and standard outer access structures into single items or texts.

2.1.2.1 ELHAT

At first glance, the ELHAT seems to provide a conventional outer search path. The opening "browse" screen provides several elements of the rapid outer access structure (see Fig. 1).

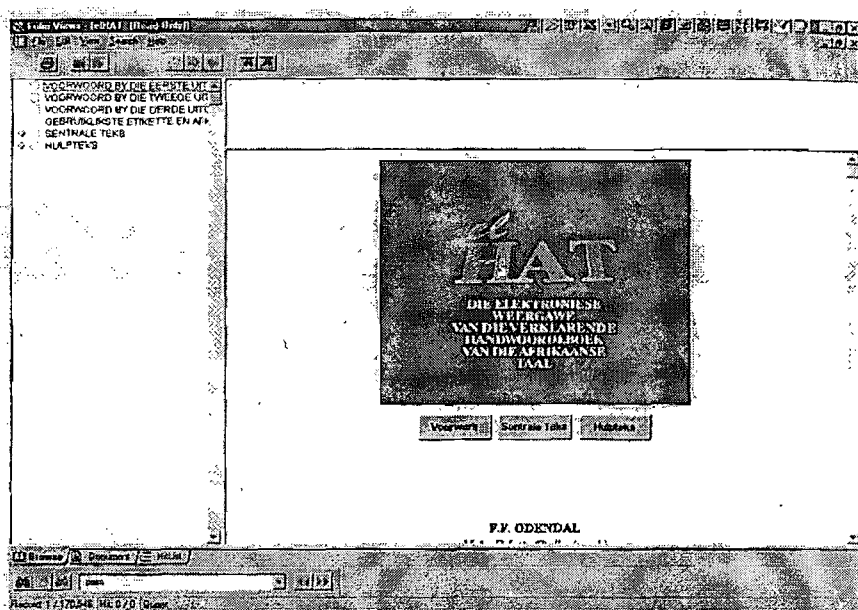


Fig. 1

An outline is given in a window on the left side of the screen. This outline provides active cross-references to transport the user to texts of the print version's front matter, the central text or an additional help text or user's guide which deals specifically with the structure, search patterns, etc. of the electronic version. The outline can provide rapid access to the macrostructure and therefore forms part of the rapid outer access structure. It shares this function with a button marked "Sentrale Teks" in the main window. A click here or in the outline will take the user to the central text, which is presented in scrollable format in the main window (the central text can be accessed separately by clicking on the icon entitled "Document" at the bottom of the screen).

At this point other elements of the rapid outer access structure take over. As a first step, the user can click on any of the 26 buttons representing the letters of the alphabet and will then be transported to the first lemma starting with that particular letter. From there the user can follow the standard outer search path by scrolling down to the lemma he/she is looking for. In order to maintain the "user addressing" (Louw and Gouws 1996: 98) of the print version, which is aimed at a very broad spectrum of Afrikaans speakers, the initial standard alphabetical outer access structure of the print version is retained unchanged. The elements of the rapid outer access structure which usually guide the user when turning the pages — the running heads — are, however, replaced. In the top right a small window shows which letter is currently consulted and displays the lemma sign of the article where the cursor is situated. This assists the user when he/she is scrolling rapidly through the central text.

By selecting a field and then clicking in the "Tik woord in"-input box, a hit list appears in the "Groep en Woordelys"-display window (see Fig. 4).

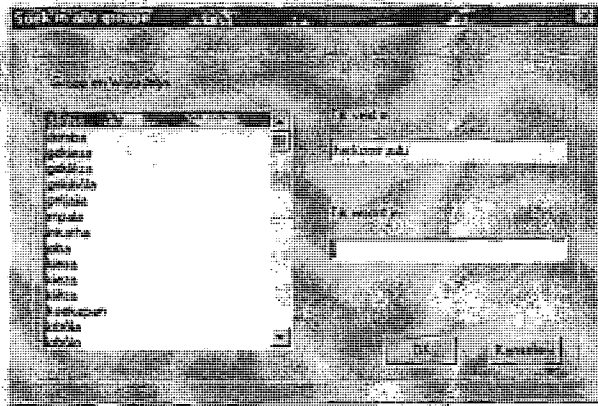


Fig. 4

Each group of lemmas represents a hit list, which can be searched by scrolling through them or by refining the search by typing in a search word. A click on "OK" will then take the user to the correct article.

2.1.2.1 EWED

The EWED-CD-ROM includes a thesaurus, a dictionary of quotations, a general monolingual dictionary and two bidirectional bilingual dictionaries. By choosing "everything" on the opening screen, searches can be conducted throughout all these dictionaries. For the purposes of this article, however, the procedures following the choice of "dictionary" on the opening screen will be focussed on. This link takes the user to the outer access structures of a general monolingual dictionary.

The outer access structure differs significantly from that of ELHAT. On the main window of both dictionaries, links are provided to various parts of the front matter, but whereas ELHAT has a link to the central text and search boxes available from menu items, EWED has a search aid that is always on the left side of the screen. This "pinpointer" is a trademark feature of the Encarta family. In the previous version of the Encarta Encyclopedia it acted as a "floating help text" (Louw 1999a: 288), but in both the Encarta Encyclopedia 2000 and EWED it is in a fixed position, which makes it easier for the target user to access.

As was the case with the search boxes in ELHAT, elements of the rapid and standard outer access structures are combined very effectively in EWED. The "pinpointer" lists the macrostructural items strictly alphabetically and scrolling down to the relevant lemma and clicking on it can access these (see

Fig. 5). The main window then displays the selected article. In this procedure the "pinpointer" forms part of the standard outer access structure and has an important function. In cases where someone is unsure of the spelling of a lemma, it affords him/her the opportunity to find the desired lemma in a search procedure that he/she would be familiar with in a print dictionary.

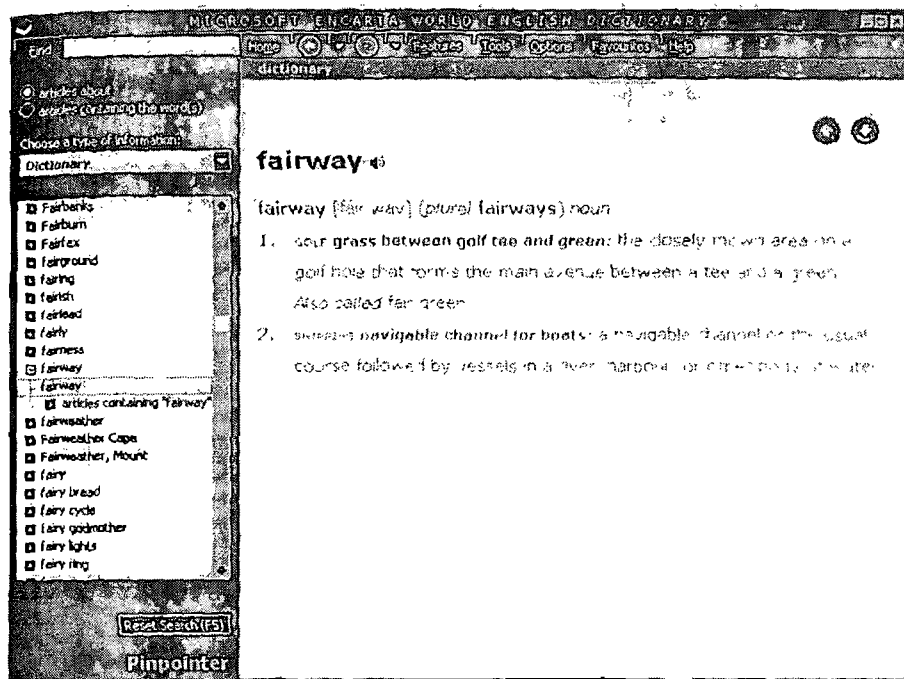


Fig. 5

This appreciation of familiarity is an important part of the user-friendliness of EWED. Louw's (1999a: 289) evaluation of the "pinpointer's" role in the Encarta Encyclopedia is therefore also valid here: "It is the preference for an initial alphabetic listing as the first step in searches using the 'pinpointer', that proves this commitment to the needs of target users, as an alphabetical listing would probably be preferred by families within the current dictionary culture".

Yet the "pinpointer" also functions as an element of the rapid outer access structure. Once the radio button labelled "Articles about" is activated a search word or phrase can be entered into the "Find" input box. The macrostructure of the dictionary is immediately scanned and the results are narrowed down as each letter is typed in. The relevant lemma is highlighted in the main display box of the "pinpointer". From here the article can be accessed directly and swiftly by a click, or the user can choose to follow a different, integrated search path.

2.2 An integrated approach to searches on the outer and inner search paths

Not all the search options in CD-ROM dictionaries necessarily maintain the boundaries of the outer and inner access structure. Whereas passive structural markers and other search aids in print dictionaries are as a rule restricted in their application to either the outer or inner access structure, active markers and search functions in electronic dictionaries can follow an integrated approach to searches on the outer and inner search paths.

2.2.1 EWED

Both dictionaries offer advanced search functions. In the EWED there is a radio button option to select all the "articles containing the word(s)". The entry of a search term provides a selection of relevant articles in which the search term can be found. This selection is displayed in the main list box of the "pinpointer", and the user can either scroll down or click on the desired article. His or her choice will then be displayed in the main window with the relevant term highlighted, be it a word, sublexical lexical item, multilexical lexical item, phrase or name.

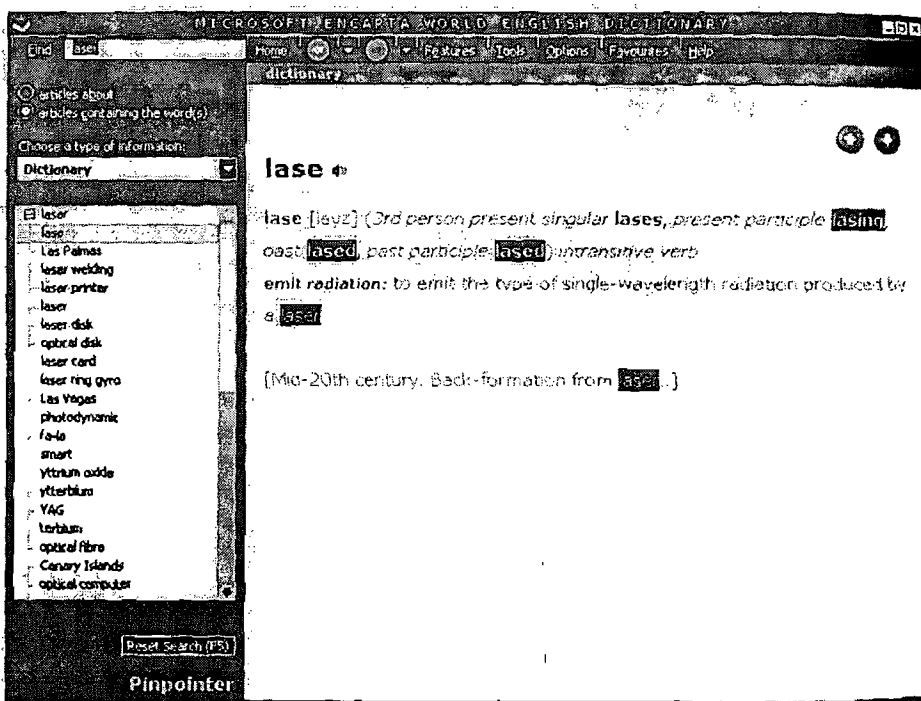


Fig. 6

This search function is especially useful if the user needs more detailed information about a search term than is contained in the article with that term as lemma, if the user needs to know if a term not lemmatised as part of the macrostructure exists, or if the user is eager to learn about the semantic relations between the search term and other lexical items. The procedure is, however, not always applied consistently enough to be of full value to the user. If one searches for "laser", for example, the first lemma on the list is *lase*.

In this article all derivatives of *lase* are highlighted in the microstructure, whereas the user searched for "laser" (see Fig. 6). The user must therefore wade through useless information to reach the requested item. The situation is even less desirable in the next article, that for *Las Palmas*. The search term "laser" is not even present, with "Las" being highlighted. This degree of inaccuracy in searches frustrates the user and should be addressed in future editions of the EWED.

2.2.2 ELHAT

The ELHAT is far more advanced in this type of procedure. It provides two valuable query options. The first is a basic query that is similar to that of the EWED. A query-form can be selected from the search menu option and a search term entered (see Fig. 7).

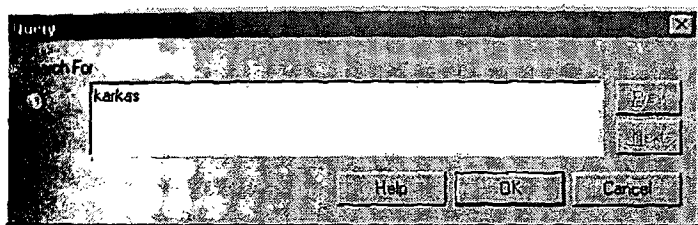


Fig. 7

The results will be compiled in a hit list, which can be accessed by clicking on the "HitList"-icon at the bottom of the main window. See for example the hit list for *karkas* (Fig. 8).

The hits can be selected individually and viewed in the document section by double-clicking on the selected item.

There is also a more advanced query procedure that illustrates the attention to the needs of the target users of the dictionary. The compilers of the dictionary knew that the target users frequently employ the dictionary as an aid in crossword puzzles. Rather than dismissing this need as unacademic, they have chosen to accommodate the users by making a "wild card"-search available. The "Advanced Query"-form can be accessed from the "Search"-menu item and is easy to use (see Fig. 9).

The question mark was chosen as a structural marker in the search term to denote which letters of the term are not known. It is a good choice, as most users are familiar with this marker and what it denotes. The entry of a search term will lead to a selection of articles in the "Word"-list box and a corresponding selection in the hit list which can be accessed in the same way as that of the normal query.

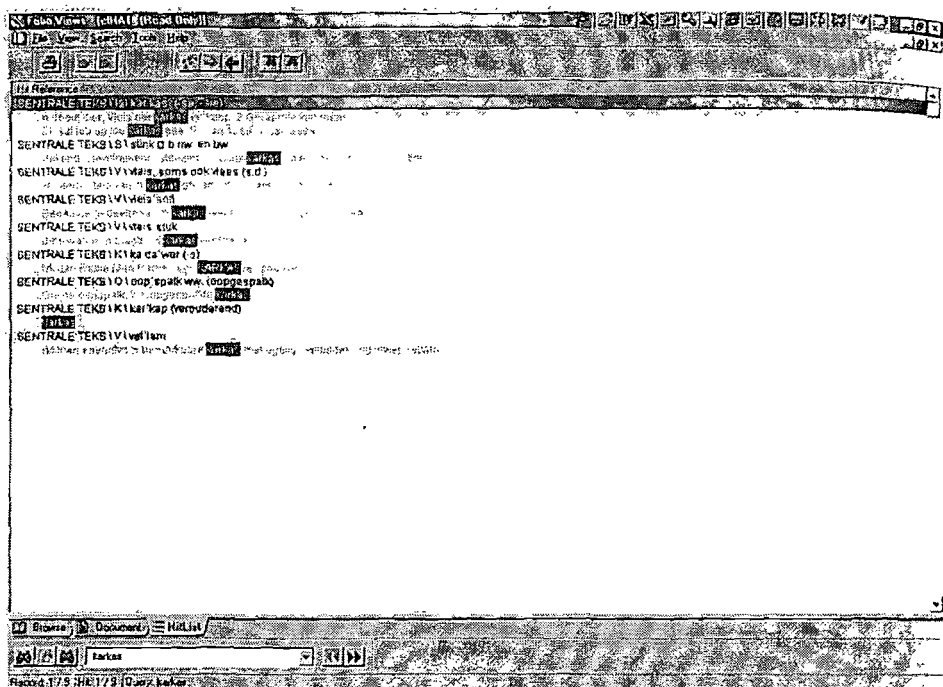


Fig. 8

2.3 The inner access structure

2.3.1 The inner search path

Most of the innovations in the CD-ROM dictionaries guide the user on the outer search path. As was shown in 2.2, however, there are some innovative search procedures that stretch to the inner search path. There are also some elements of the inner access structure which are unique to electronic dictionaries.

These unique elements will be discussed in this section. Elements of the inner access structure which are truly innovative in EWED, but are also included in the print version of this dictionary, will therefore not be discussed.

Some care has been taken in EWED to enhance the "findability" of microstructural information. Unfortunately, the same cannot be said for ELHAT.

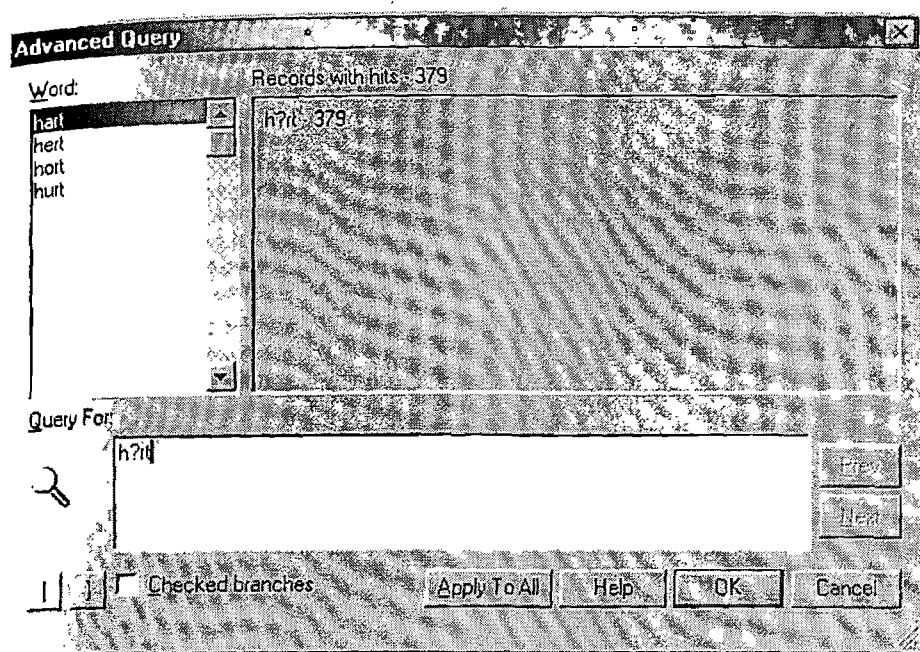


Fig. 9

2.3.2 ELHAT

ELHAT is in essence purely an electronic adaptation of a print dictionary, which has been on the market a long time and has not undergone major structural changes. This lack of structural innovation is reflected in the lack of elements of the inner access structure.

2.3.2 EWED

In the EWED, on the other hand, the use of multimedia has necessitated a fresh approach to the inner access structure. The first and most important multimedia element in EWED is the aid in pronunciation. There is an icon denoting a speaker next to some lemmas, which, if clicked on, provides the pronunciation in the form of a sound clip. The user is therefore granted rapid access to an accurate pronunciation. The icon denoting the speaker acts as a structural marker that forms part of the rapid inner access structure.

The EWED can be classified as an encyclopaedic dictionary that breaks with the traditional distinction between encyclopaedia and linguistic dictionary, opting for a hybrid format. On a macrostructural level this has, for example, led to the inclusion of place names, etc. On a microstructural level it has not only had implications for the level of encyclopaedic information contained in the items giving the meaning paraphrases, but also owing to the inclusion of multimedia items that supports the transfer of encyclopaedic information. These include maps, diagrams, pictures (mostly photographs) and sound clips (such as the anthems of countries whose names have been lemmatised). The visual items are usually included as inserted inner texts, but in many cases these small items act as active cross-references which provide rapid access to enlarged versions of the inserted inner texts. The sound clips are introduced by means of an abstract icon. In all these cases the items act not only as cross-references, but also as effective, unique structural markers of the inner access structure, which introduce the multimedia.

3. The mediostructure

CD-ROM dictionaries have not only broken new ground with the inner and outer access structures. Some of the most telling innovations are to be found in their mediostructure, which encompasses all lexicographic cross-references in a dictionary.

3.1 EWED

As was shown in 2.3.2, active cross-references have a significant role in the successful use of multimedia, where smaller versions and abstract icons are used as links to separate audio and visual elements. In the case of visual elements such as maps or pictures, these links invariably lead to an illustration, which operates as a text that does not form part of the central text. The cross-references in this case are therefore text-external cross-references.

Active text-external cross-references are employed to great effect elsewhere in EWED. The active link to audio clips giving the pronunciation was discussed in 2.3.2, but these are not the only innovations in the pronunciation information. Whereas only some lemmas have audio clips, transcriptions are provided for almost all lemmas. These transcriptions have a green font colour, which shows that they are active text-external cross-references. A click on the pronunciation takes the user to a pronunciation guide, which can also be accessed separately as part of the front matter (see Fig. 10).

These cross-references offer instant help to the user who struggles with the pronunciation of a specific lemma. They improve the user-friendliness of EWED considerably and should be adopted by other CD-ROM dictionaries.

Pronunciation Key		Click any pronunciation symbol or word to hear it spoken.	
a	at	l, ll	let, silly
aa	father	'l	apple
aw	all	m, mm	mother, hammer
ay	day	'm	rhythm
air	hair	n, nn	not, funny
b, bb	but, ribbon	'n	fashion
ch	chin	ng	song
d, dd	do, ladder	o	odd
a	about, item, edible, common, circus	ō	open
e	egg	ōō	good
ee	eel	oo	school
f, ff	fond, differ	ow	owl
g, gg	go, giggle	oy	oil
h	hat	p, pp	pen, happy
hw	when	r, rr	road, carry
i	it, happy	s, ss	say, lesson
r	ice	sh	sheep
j, jj	juice, pigeon	th	thin
k	key, thick	th	this
h	hat	t, tt	tell, butter
		u	up
		ur	urge
		v, vv	very, savvy
		w	wet
		y	yes
		z, zz	zoo, blizzard
		zh	vision
In foreign words:			
kh	as in German Bach,		
	as in Scottish loch		
aN	as in French vin		
aaN	as in French blanc		
oN	as in French bon		
ōN	as in French un		
ō	as in German schön		
ū	as in French rue		
Stress marks: ' as in <i>secret</i> [sɛk' rət] ˈ as in <i>secretary</i> [sɛk' rɪ tɛəri]			
Note: This dictionary uses double letters to show many sounds in the middle of words, because English spelling normally uses double letters in these positions. This is one of several ways that the system of spelling pronunciations here is more natural and easier to use than the systems in many other dictionaries.			

Fig. 10

EWED's more conventional text-internal cross-references are also active. The explicit remote article-external cross-references are given in blue text and have an important role to play in maintaining the textual cohesion of this dictionary. A click on the explicit reference "See also genie" in the article of *jinni* will take the user directly to *genie*, where the cross-reference is met by a corresponding active cross-reference to complete the circle. This is unfortunately not always the case. The explicit cross-references to *abalone* at *ear shell* and *sea ear* are active, but their corresponding cross-references at *abalone* are not.

Explicit article-external cross-references that are complicated by polysemy are not treated adequately. In the case of the synonymous pair *hemp* and *marijuana*, for example, the use of cross-references is confusing. A cross-reference "See also marijuana" is listed under the correct sense of *hemp*, but it is not specified whether *hemp* can be used as a synonym for both senses of *marijuana* or, if not, for which sense it can act as a synonym. Once again the circle is not completed by a corresponding active cross-reference at the address of this cross-reference (i.e. *marijuana*). This creates even more confusion in the user's mind and compounds the error made by the compilers.

Deliberately monodirectional explicit article-external cross-references are also employed in EWED, with more success. They are particularly effective

when used in the etymology article position. In the case of *genie*, for example, the etymology is related to the Latin "genius". Instead of repeating the full etymology of "genius", the user can click on the cross-reference "(see genius)" and be transported to the article for *genius* in which the full etymology is explained.

Not all the article-external cross-references are explicit though. Probably the most innovative mediostuctural elements are the implicit cross-references. A double-click on any word in the microstructure, will prompt a small screen called "Quickdefine" to appear. One can, for example, highlight "panama hats" in the the microstructure of *jipijapa* and double click with the result as shown in Fig. 11.

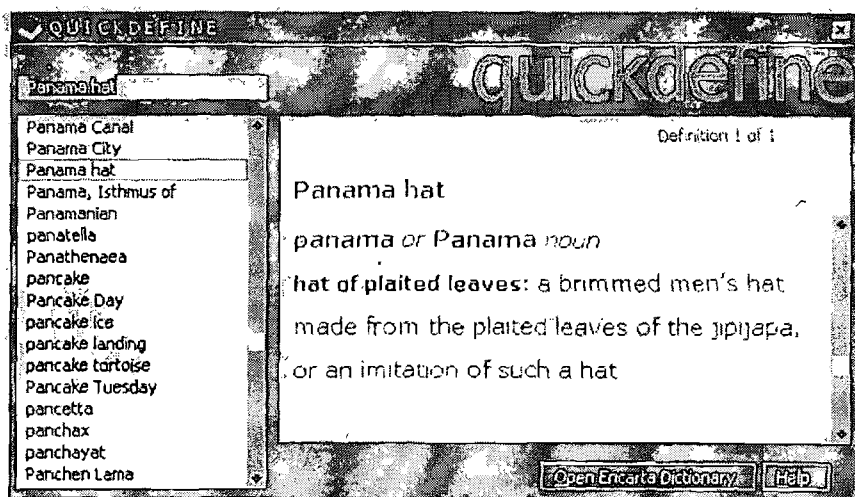


Fig. 11

This is yet another user-friendly feature of EWED that gives the user rapid access to the information he/she desires. EWED goes one step further by allowing users to obtain a "quickdefine"-window for search terms in ordinary Windows-based word processors such as Microsoft Word by highlighting the term and then clicking on the quickdefine desktop item, rather than opening the whole dictionary. The emphasis in both these procedures is on fast, effective access to the desired information.

3.2 ELHAT

ELHAT's mediostucture has unfortunately not even remotely reached the standards of innovation set by EWED, as is to be expected in such a far more modest attempt. The mediostucture is restricted to explicit article-external cross-references. These are generally well implemented and in the case of cross-

references to polysemous articles, the sense to which the cross-reference applies is mentioned, e.g. "Vgl. KAALVUIS, 2" at *kaalkop*. This level of accuracy is what was needed in EWED.

ELHAT fails in another important aspect though, by not consistently completing the bidirectionality of active cross-references. At *kaalvuis* the cross-reference is met by the entry "kaalkop (b.nw. en bw. bet. 4)", but in this case the cross-reference is passive. Cross-references such as this one should also be active so that the dictionary can be the cohesive entity it has the potential to be.

4. Conclusion

Both the dictionaries under discussion have similar target user groups, with the primary difference being the language spoken by each group. The compilers of each of these dictionaries should take note of the innovative procedures followed by the compilers of the other. In order to keep these products viable at a time when CD-ROM dictionaries are increasingly under threat of being rendered obsolete by counterparts on the Internet, lexicographers need to make full use of the superior searchability provided by the CD-ROM format at the moment. In this regard the guide structures, and in particular the access and mediostructures, have an important role to play. A cross-pollination of features between CD-ROM dictionaries will improve the accessibility of the dictionaries and ultimately benefit their target users.

Bibliography

Dictionaries

- Elektroniese Verklarende Handwoordeboek van die Afrikaanse Taal. 2.00.2000. Perskor.
Microsoft Encarta World English Dictionary. 1999. Microsoft Corporation

Other Sources

- Gouws, R.H. and D.J. Prinsloo. 1998. Cross-Referencing as a Lexicographic Device. *Lexikos* 8: 17-36.
Hausmann, F.J., et al. (Eds.). 1989-1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie Dictionaries/An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Berlin: Walter de Gruyter.
Hausmann, F.J. and H.E. Wiegand. 1989. Component Parts and Structures of Monolingual Dictionaries. Hausmann, F.J. et al. (Eds.). 1989-1991: 328-360.
Louw, P.A. 1999. Access Structures in a Standard Translation Dictionary. *Lexikos* 9: 108-118.
Louw, P.A. 1999a. Portals to Knowledge: CD-ROM Encyclopaedias, with Specific Reference to *Microsoft Encarta 99 Encyclopaedia*. *Lexikos* 9: 283-293.
Louw, P.A. and R.H. Gouws. 1996. Lemmatiese en niemlemmatiese adressering in Afrikaanse vertalende woordeboeke. *South African Journal of Linguistics* 14(3): 92-100.

Juta's Concise Dictionary of Accounting Terms with their Afrikaans Equivalents

N. Stegmann, *Departement Rekeningkunde, Randse Afrikaanse Universiteit, Johannesburg, Republiek van Suid-Afrika*

Opsomming: Die derde uitgawe van die woordeboek is in 1996 saamgestel in 'n tyd toe die standaardiseringsproses van terme in die middel van 'n internasionaliseringproses was. Die meeste van die Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk wat as riglyn vir die rekeningkunde dien, is hersien en in ooreenstemming met die Internasionale Rekeningkundige Standaardekomitee se voorskrifte gebring. Aangesien rekeningkundige terminologie bepaal word deur hierdie Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk, is 'n groot aantal van die terme in die woordeboek verouderd. Omdat enige skrywer van 'n rekenkundige handboek weerloos is teen sulke veranderinge, het die resensent probeer om terme te identifiseer wat sedertdien verander het of nie voorheen in die rekeningkunde gebruik is nie. Die woordeboek is 'n nuttige werkboek vir enigeen wat geïnteresseerd is in die terminologie wat in die rekeningkunde gebruik word, maar sal veral bruikbaar wees vir studente wat op hoogte moet kom van die vaktaal van die rekeningkunde.

Sleutelwoorde: UNIEKE POSISIE, VERKLARINGS VAN ALGEMEEN AANVAARDE REKENKUNDIGE PRAKTYK (GAAP), GEKODIFISEERDE GAAP, INTERNASIONALISERINGSPROSES, VEROUERDE TERME, ONVOLDOENDE OMSKRYWINGS, AANGEPASTE OMSKRYWINGS, NUWE TERMINOLOGIE

Abstract: *Juta's Concise Dictionary of Accounting Terms with their Afrikaans Equivalents.* The third edition of the dictionary was composed in 1996 at which time the standardization process of terms was in the middle of an internationalisation process. Most of the South African Statements of Generally Accepted Accounting Practice which serves as guideline for accounting were revised and brought into compliance with the International Accounting Standards Committee's regulations. As accounting terminology is determined by these Statements of Generally Accepted Accounting Practice, a large part of the terms in the dictionary is outdated. Because any author of an accounting text book is defenceless against such changes, the reviewer attempted to identify terms that have since been changed or not been used in accounting previously. The dictionary is a helpful tool for anybody interested in the terminology used in accounting, but will be especially useful for students who need to acquaint themselves with the technical language of accounting.

Keywords: UNIQUE POSITION, STATEMENTS OF GENERALLY ACCEPTED ACCOUNTING PRACTICE (GAAP), CODIFIED GAAP, INTERNATIONALISATION PROCESS,

ANTIQUATED TERMS, INSUFFICIENT DESCRIPTIONS, MODIFIED DESCRIPTIONS, NEW TERMINOLOGY

1. Inleidend

1.1 Unieke posisie van rekeningkunde

Rekeningkunde beklee 'n unieke posisie in die ekonomiese en bestuurswetenskappe aangesien die vakgebied deur 'n standaardliggaam gereguleer word. Hierdie liggaam staan as die Suid-Afrikaanse Instituut van Geoktrooieerde Rekenmeesters (SAIGR) bekend. Die SAIGR kodifiseer Algemeen Aanvaarde Rekeningkundige Praktyk (AARP) in Suid-Afrika deur middel van Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk na goedkeuring deur die Rekeningkundige Praktykeraad. Die Rekeningkundige Praktykeraad bestaan uit 'n wye groep verteenwoordigers van alle belanghebbende institute en liggamme wat belange by finansiële verslagdoening in Suid-Afrika het. Die terminologie wat in die Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk gebruik word, is dus aanvaarbaar vir 'n wye reeks belanghebbers. Die SAIGR is ook tans besig met 'n internasionaliseringsproses deur middel waarvan Suid-Afrikaanse Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk in ooreenstemming met internasionale standaarde gebring word.

Dit behoort dus duidelik te wees dat terminologie in rekeningkunde tot 'n groot mate gereguleer word deur Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk deurdat die meeste rekeningkundige terme in Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk omskryf word.

1.2 Gekodifiseerde AARP as uitgangspunt

Die *Juta's Concise Dictionary of Accounting Terms* word hoofsaaklik deur studente en rekenmeesters gebruik. Daar word ook aanvaar dat gebruikers van finansiële state die woordeboek benut wanneer finansiële state bestudeer word. Indien die beperkende raamwerk van terminologie soos vasgestel in Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk in ag geneem word, behoort dit duidelik te wees dat dit sinneloos is om van hierdie vasgestelde terminologie af te wyk, aangesien voorbereiders van finansiële state juis veronderstel is om binne die gegewe raamwerk te opereer.

2. Verbeteringe

Die woordeboek is volledig indien die datum van publikasie in ag geneem word, aangesien die woordeboek midde-in die omvattende internasionalisering van Standpunte geskryf is. Vanweë die snelontwikkelende aard van

gekodifiseerde Algemeen Aanvaarde Rekeningkundige Praktyk, is die woordeboek noodwendig verouderd en het die werk met die huidige inhoud nie veel gebruiksnut nie. Hierdie situasie is heeltemal buite die outeurs se beheer en daarom is die kommentaar hoofsaaklik daarop gerig om terme wat sedert die publikasie van die woordeboek omskryf of gewysig is, te identifiseer.

3. Formaat

Die formaat van die woordeboek is baie gebruikersvriendelik. Die waarde van die werk word aansienlik verhoog deur die verskaffing van voorbeelde van waardasiemetodes vir voorraad, waardeverminderingsberekennings en finansiële state. Die voorbeelde van die finansiële state is egter verouderd, aangesien 'n nuwe standpunt 'n gewysigde formaat voorstel. Die verskaffing van Afrikaanse terme met hul Engelse ekwivalente is baie nuttig vir Afrikaanse studente en voorbereiders van finansiële state.

4. Werkswyse

Die woordeboek is noukeurig bestudeer en kommentaar word oor alle aspekte gelewer wat aandag verdien. Met inagneming van die raamwerk wat hierbo uiteengesit is, word verouderde terme uitgewys, veranderinge wat deur gewysigde Standpunte van Algemeen Aanvaarde Rekeningkundige Praktyk vereis word, aanbeveel en 'n lys verskaf van terme wat sedert publikasie algemeen aanvaarde rekeningkundige terme geword het en wat in die woordeboek bygevoeg behoort te word. Die woordeboek bevat slegs enkele foute wat ook geïdentifiseer word met aanbevelings ter regstelling.

5. Foutiewe of ontoereikende beskrywings

Die omskrywing van die volgende term word bevraagteken:

Account (for) Die gebruik van die term "explain" in die beskrywing "To explain, substantiate or record the effect of a transaction" is foutief. In die rekeningkunde beteken die term "account (for)" dat 'n transaksie in die boek van eerste inskrywing aangeteken word.

Die omskrywing van die volgende term behoort eerder te lui:

Accounting *Rekeningkunde*. The process of identifying, measuring, recognising, presenting and disclosing economic information to facilitate informed judgements and decision making by users of the information.

Die volgende vertalings van terme word bevrageeteken:

Accrual (adjustments, basis of accounting) en **Accrue** *Oplopiingsaansuiwering, Oplopiingsbeginsel* en *Oploop* behoort vervang te word met *Toevallingsaansuiwering, Toevallingsaansuiwering* en *Toeval* (verwys RE 000.22).

Assessed loss *Belastingsverlies* behoort vervang te word met *Aangeslane verlies*. Die beskrywing behoort ook gewysig te word na "The excess of tax deductible expenses over taxable income as confirmed by SARS".

Cash on delivery *Kontant by aflewering* is die meer aanvaarde term. Dit word algemeen as *KBA* afgekort.

Directors' report *Direkteursverslag* is verkieslik bo *Direkteureverslag*.

Excess of expenses over revenue *Tekort*. Nie *Oorskot* soos dit tans vertaal is nie.

Expenditure *Besteding*. Nie *Uitgawes* soos dit tans vertaal is nie.

Inflation accounting *Inflasie-rekeningkunde*. Hierdie tipe rekeningkunde verantwoord die effek van algemene prysverhogings en nie bloot die effek van stygende pryse nie. Die beskrywing "rising prices" behoort vervang te word met "general increases in prices".

Purchases journal *Aankoopjoernaal*. Nie *Aankopejoernaal* nie.

Purchases returns *Aankoopterugsendings; aankoopretoere*. Nie *Aankopeterugsendings; aankoperretoere* nie.

Purchases returns journal *Aankoopterugsendingsjoernaal*. Nie *Aankopeterugsendingsjoernaal* nie. Verder hoort *Aankopejoernaal*, die vertaling van *Purchases journal*, glad nie hier nie.

6. Onvolledige omskrywing van terme

Die volgende terme se omskrywings is na die skrywer se mening onvolledig:

Liquidate *Likwiedeer*. Daar behoort 'n verwysing te wees na *Wind up*.

Fair value *Billike waarde*. Die volgende omskrywing behoort bygevoeg te word: "The amount for which an asset could be exchanged or a liability settled between knowledgeable, willing parties in an arm's length transaction" (AC 123.10).

Reserve *Reserwe*. Daar behoort bygevoeg te word dat reserwes 'n onderdeel van ekwiteit is en dat dit ook ontstaan as gevolg van winste wat nie aangewend is nie.

Rights issue *Regte-uitgifte*. Daar moet vermeld word dat sodanige uitgifte plaasvind in 'n sekere verhouding tot bestaande aandelebesit.

7. Onvoldoende vertaling van terme in Afrikaans

Die volgende Engelse terme kan aangevul word met alternatiewe vertalings:

- Acceptor** *Aannemer*. Dit word ook algemeen in die praktyk as *akseptant* vertaal.
- Accounting date** *Rekeningkundige datum*. Dit word meer algemeen as *rekenpligtige datum* vertaal.
- Accounting period** *Rekeningkundige tydperk*. Dit word tans as *rekenpligtige tydperk* vertaal (verwys RE 101.50).
- Ageing** *Debiteureveroudering*. Die term word meer algemeen met *debiteure-ontleding* vertaal.
- Allocate** *Toewys*. Die term word meer algemeen met *toedeel* vertaal.
- Application and allotment account** *Aansoek- en toekenningsrekening*. Dit kan ook vertaal word met *aansoek- en toewysingsrekening*.
- Closing balance** *Sluitingsaldo*. Dit word meer algemeen as *eindsaldo* vertaal. Die byvoeging van die sin "Stock is usually valued at the lower of cost and net realisable value" as deel van die verklaring van hierdie term word bevestig.
- Credit By (1)** moet *Krediet* vervang word met *Krediteer*. Die volgende moet bygevoeg word: *Crediting Kreditering*. The recording of a credit entry in the ledger account.
- Current ratio** *Bedryfsverhouding*. Dit word meer algemeen met die term *bedryfskapitaalverhouding* vertaal.
- Debit** *Debiet* moet vervang word met *Debiteer*. Die volgende moet bygevoeg word: *Debiting Debitering*. The recording of a debit entry in a ledger account.
- Favourable balance** *Batige saldo*. *Voordelige saldo* word ook algemeen gebruik.
- Fixed overhead** *Vaste bokoste*. *Vaste drakoste* is die verkieslike term ingevolge RE 108.
- Flat rate of interest** *Uniforme tarief; rentekoers*. *Bytelkoers* word die meeste in die praktyk gebruik.
- Gearing** *Hefboomwerking*. Die omskrywing is eerder dié van *hefboomwerking* as van *hefboomfinansiering*. (Verwys *Leverage*).
- Participating preference shares** *Deelnemende voorkeuraandeel*. *Winsdelende voorkeuraandeel* kan as alternatief bygevoeg word.
- Pre-incorporation profits/losses** *Voorinkorporasiewinste/-verliese*. *Voorinlywingswinste/-verliese* kan as alternatiewe bygevoeg word.
- Profit-sharing ratio** *Winsverdelingsverhouding*. Hiernaas kan *winsdelingsverhouding* bygevoeg word.
- Subsidiary books** *By-boeke*. Dié term word selde gebruik. *Hulpboeke* is meer gebruiklik.
- Subsidiary ledger** *Memorandumgrootboek*. Die term *hulpgrootboek* word meer algemeen gebruik.

Work in progress *Onvoltooide werk; werk-aan-die-gang*. Die volgende kan ook bygevoeg word: *werk-in-proses; werk-in-voortgang*.

8. Verouderde terme

Die volgende terme is verouderd ingevolge Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk en behoort nie meer algemeen gebruik te word nie:

Abnormal item *Abnormale item*.

Associated company *Geassosieerde maatskappy*. Die terme word vervang met **Associate** en *Geassosieerde*.

Book value *Boekwaarde; drawaarde*. In RE 123.10 word die terme **Carrying amount** *Drabedrag* gebruik.

Contingency *Gebeurlikheid*.

Contingent gain/loss *Voorwaardelike wins/verlies*.

Discontinued operation *Beëindigde bedrywigheid*. Die terme **Discontinuing** en *Eindigende* word tans in AC 117 en RE 117 gebruik.

Fixed asset *Vaste bate*. Die terme word vervang deur **Non-current asset** en *Niebedryfsbate* in AC 101 en RE 101.

Holding company *Houermaatskappy*. Hierdie terme is vervang met **Parent company** *Moedermaatskappy*, maar word nog baie in die praktyk gebruik en oorweging behoort geskenk te word aan die behoud van die ou terme in die woordeboek.

Net current assets *Netto bedryfsbates*. Ondernemings mag nie meer bates en laste teen mekaar verreken nie (RE 101.34). Dieselfde geld vir **Net monetary assets** *Netto monetêre bates*.

Net income *Netto inkomste*. Die term word nou vervang met *Netto wins*. (Verwys ook na *Net income before/after tax*).

Outside shareholders *Buite-aandeelhouders*. Hierdie terme is vervang met **Minority interest** *Minderheidsbelang* in RE 131 en AC 131, maar word nog baie in die praktyk gebruik en oorweging behoort geskenk te word aan die behoud van die ou terme in die woordeboek.

Permanent difference *Permanente verskil/Blywende verskil*. Hierdie terme word nie meer gebruik in die nuwe AC 102 en RE 102 nie.

Source and application of funds statement *Staat van bron en aanwending van fondse*. Dit is 'n staat wat nie meer voorberei word nie.

Stock Dit is in AC 108 algemeen met **Inventories** vervang.

Timing differences *Tydsberekeningsverskille*. Hierdie terme is vervang met die terme **Temporary differences** *Tydlike verskille* in AC 102 en RE 102.

9. Wysigings vereis deur Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk

Die volgende terme se omskrywings of vertalings moet gewysig word soos aangedui, omdat die toepaslike Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk verander is:

- Accounting assumptions** *Rekeningkundige aannames*. Dit behoort as *Rekeningkundige veronderstellings* vertaal te word (verwys RE 000.22).
- Accounting policies** *Rekeningkundige beleid*. The specific principles, bases, conventions, rules and practices adopted by an enterprise in preparing and presenting financial statements. (AC 103.05)
- Amortisation** *Amortiseer*. Systematic allocation of the depreciable amount of an intangible asset over its useful life. (AC 129.09)
- Annual report** *Jaarverslag*. A report for shareholders and other interested parties prepared once a year. Includes a balance sheet, an income statement, a statement showing either all changes in equity or changes in equity other than those arising from capital transactions with owners and distributions to owners, cash flow statements, accounting policies and explanatory notes. (AC 101.08)
- Associate** *Geassosieerde*. An enterprise in which the investor has significant influence and which is neither a subsidiary nor a joint venture of the investor. (AC 110.02)
- Cash flow statement** *Kontantvloeiistaat*. A statement that reflects the sources and uses of cash and cash equivalents for a period.
- Contingent asset** *Voorwaardelike bate*. A possible asset that arises from past events and whose existence will be confirmed only by the occurrence or nonoccurrence of one or more uncertain future events not wholly within the control of the enterprise. (AC 130.11)
- Contingent liability** *Voorwaardelike aanspreeklikheid*. A possible obligation that arises from past events and whose existence will be confirmed only by the occurrence or nonoccurrence of one or more uncertain future events not wholly within the control of the enterprise, or a present obligation that arises from past events but is not recognised because it is not probable that an outflow of resources embodying economic benefits will be required to settle the obligation, or the amount of the obligation cannot be measured with sufficient reliability. (AC 130.11)
- Cost** *Koste*. The amount of cash or cash equivalents paid or the fair value of the other consideration given to acquire an asset at the time of its acquisition or production. (AC 129.08)
- Current asset** *Bedryfsbate*. An asset that is expected to be realised in, or is held for sale or consumption in the normal course of the enterprise's operating cycle, or is held primarily for trading purposes or

for the short term, and expected to be realised within twelve months of the balance sheet date, or is cash or a cash equivalent asset that is not restricted in its use. (AC 101.58)

Current liability *Bedryfslas*. A liability that is expected to be settled in the normal course of the enterprise's operating cycle, or is due to be settled within twelve months of the balance sheet date. (AC 101.61)

Current tax *Lopende belasting*. The amount of income tax payable (recoverable) in respect of the taxable profit (tax loss) for a period (AC 102.09)

Debtor *Debiteur*. Person who has an obligation to pay another (creditor). Also referred to as *Account receivable* or *receivables*. (AC 101 Appendix)

Deferred tax *Uitgestelde belasting*. The amount of taxation expense, the amount of which has been postponed or prepaid as a result of temporary differences. If the deferred tax has a cumulative credit balance, income tax will be payable in future periods. Where a cumulative debit balance exists income tax will be recoverable in future. (AC 102)

Depreciable amount *Afskryfbare bedrag*. The cost of an asset, or other amount substituted for cost in the financial statements, less its residual amount which is to be depreciated over its useful life. (AC 123.10)

Depreciation *Depresiasie; waardevermindering*. The systematic allocation of the depreciable amount of an asset over its useful life. (AC 123.10)

Earnings per share *Verdienste per aandeel*. Die omskrywing van *P* moet soos volg gewysig word:

P = Profit, including equity profit from associate and extraordinary item, before tax.

Equity *Ekwiteit*. Die omskrywing wat gegee word, verskil van die omskrywing verskaf in RE 000, daarom moet die verwysing na AC 000 verkieslik uitgelaat word, of die omskrywing moet aangevul word met die volgende:

"the residual interest in the assets of the enterprise after deducting all its liabilities". (AC 000.49)

Equity basis/method of accounting *Ekwiteitsbasis/-metode van verantwoording*. Investment is initially recorded at cost and adjusted thereafter for the postacquisition change in the investor's share of net assets of the investee. The income statement reflects the investor's share of the results of operations of the investee. (AC 110.02)

Financial statements *Finansiële state*. Die gekursiveerde gedeelte moet bygevoeg word: "Financial statements of a company comprise a balance sheet, an income statement, a statement showing changes in equity, a cash flow statement, ..."

- Income statement** *Inkomstestaat*. Hierdie staat bevat nie meer aanwendings van netto wins indien die vereistes van RE 101 nagevolg word nie. Die beskrywing is egter steeds korrek indien toegepas op eenmansake en vennootskappe.
- Intangible asset** *Ontasbare bate*. An identifiable nonmonetary asset without physical substance held for use in the production or supply of goods or services, for rental to others, or for administrative purposes. (AC 129.08)
- Joint ventures** *Gesamentlike ondernemings*. A contractual arrangement whereby two or more parties undertake an economic activity that is subject to joint control. (AC 119.02)
- Lease** *Huur*. An agreement whereby the lessor conveys to the lessee in return for a payment or series of payments the right to use an asset for an agreed period of time. (ED 128.04)
- Minority interest** *Minderheidsbelang*. That part of the net results of operations and of net assets of a subsidiary attributable to interests that are not owned, directly or indirectly through subsidiaries, by the parent. (AC 131.09)
- Monetary item** *Monetêre item*. Money held, and assets and liabilities to be received or paid, in fixed or determinable amounts of money. (AC 112.08)
- Net realisable value** *Netto realiseerbare waarde*. The estimated selling price in the ordinary course of business less the estimated costs of completion and the estimated costs necessary to make the sale. (AC 108.04)
- Offset** *Verrekening*. To deduct one amount from another where a Statement of Generally Accepted Accounting Practice requires it or where gains, losses and related expenses arising from the same or similar transactions and events are not material. (AC 101.35)
- Overhead costs** *Bo-koste*. Die term *Drakoste* word vereis deur RE 108.
- Parent** *Moedermaatskappy*. An enterprise that has one or more subsidiaries. (AC 131.09)
- Provision** *Voorsiening*. A liability of uncertain timing or amount. (AC 130.11)
- Reducing balance method** *Verminderdesaldo-metode*. A method of calculating depreciation by applying the rate to the carrying amount (nie "book value" nie).
- Residual value** *Reswaarde*. The net amount that the enterprise expects to obtain for an asset at the end of its useful life after deducting the expected costs of disposal. (AC 123:10)
- Revenue** *Inkomste*. The gross inflow of economic benefits during the period arising in the course of the ordinary activities of an enterprise when those inflows result in increases in equity, other than

increases relating to contributions from equity participants. (AC 111.09)

Taxable profit/loss *Belasbare wins/verlies*. The profit/loss for a period, determined in accordance with the rules established by the taxation authorities, upon which income taxes are payable (recoverable). (AC 102.09)

Revenue *Inkomste*. The gross inflows of economic benefits during the period arising in the course of the ordinary activities of an enterprise when those inflows result in increases in equity, other than increases relating to contributions from equity participants. (AC 111.09).

10. Nuwe terminologie vereis deur Standpunte van Algemeen Aanvaarde Rekeningkundige Praktyk

Die volgende terme wat algemeen gebruik word of deur nuwe Standpunte oor Algemeen Aanvaarde Rekeningkundige Praktyk gebruik word, is uitgelaat en oorweging behoort geskenk te word aan hul insluiting in die werk:

Absorption *Absorpsie*.

Accounting profit *Rekeningkundige wins*. The net profit or loss for a period before deducting tax expenses. (AC 102.09)

Borrowing costs *Leenkoste*. Interest and other costs incurred by an enterprise in connection with the borrowing of funds. (AC 114.04)

Carrying amount *Drabedrag*. The amount at which an asset is recognised in the balance sheet after deducting any accumulated depreciation (amortisation) and accumulated impairment losses thereon. (AC 128.06)

Cash *Kontant*. Comprises cash on hand and demand deposits. (AC 118.07)

Cash equivalents *Kontantekwivalente*. Short-term, highly liquid investments that are readily convertible to known amounts of cash and are subject to insignificant risk of changes in value. (AC 118.07)

Charge *Debiteer*.

Control *Beheer*. The power to govern the financial and operating policies of an economic activity so as to obtain benefits from it. (AC 119.02)

Defined benefit plan *Omskrewevoordeel-plan*. Retirement benefit plan under which amounts to be paid as retirement benefits are determined by reference to a formula, usually based on employees' remuneration and/or years of services. (AC 116.06)

Defined contribution plan *Omskrewedydrae-plan*. Retirement plans under which amounts to be paid as retirement benefits are determined by reference to contributions to a fund, together with the investment earnings thereon. (AC 116.06)

- Equity instrument** *Ekwiteitsinstrument*. Any contract that evidences a residual interest in the assets of an enterprise after deducting all of its liabilities. (AC 125.07)
- Financial instrument** *Finansiële instrument*. Any contract that gives rise to both a financial asset of one enterprise and a financial liability or equity instrument of another enterprise. (AC 125.07)
- Financial asset** *Finansiële bate*. Any asset that is cash, a contractual right to receive cash or another financial instrument, a contractual right to exchange financial instruments with another enterprise under conditions that are potentially favourable, or any equity instrument of another enterprise. (AC 125.07)
- Financial liability** *Finansiële las*. Any liability that is a contractual obligation to deliver cash or another financial asset to another enterprise, or to exchange financial instruments with another enterprise under conditions that are potentially unfavourable. (AC 125.07)
- Financing activities** *Finansieringsaktiwiteit*. Activities that result in changes in the size and composition of the equity capital and borrowings of the enterprise. (AC 118.07)
- Forward cover** *Termyndekking*.
- Fundamental error** *Fundamentele fout*. Errors discovered in the current period that are of such significance that the financial statements of one or more prior periods can no longer be considered to have been reliable at the date of their issue. (AC 103.05)
- Impairment loss** *Waardedalingsverlies*. The amount by which the carrying amount of an asset exceeds its recoverable amount. (AC 128.06)
- Interim period** *Tussentydperk*. A financial reporting period shorter than a full financial year. (AC 127.05)
- Investing activities** *Beleggingsaktiwiteite*. The acquisition and disposal of long-term assets and other investments, not included in cash equivalents. (AC 118.07)
- Joint control** *Medebeheer*. The contractually agreed sharing of control over an economic activity. (AC 119.02)
- Onerous contract** *Beswarende kontrak*. A contract in which the unavoidable costs of meeting the obligation under the contract exceeds the economic benefits expected to be received under it. (AC 130.11)
- Ordinary activities** *Gewone bedrywighede*. Any activities that are undertaken by an enterprise as part of its business and such related activities in which the enterprise engages in furtherance of, incidental to, or arising from these activities. (AC 103.05)
- Ordinary shares** *Gewone aandele*. An equity instrument that is subordinate to all other classes of equity instruments. (AC 104.07)
- Payables** Die term word in AC 101 gebruik. Daar bestaan nie 'n algemeen aanvaarde Afrikaanse term daarvoor nie.

- Potential ordinary shares** *Potensiële gewone aandele*. A financial instrument or other contract that may entitle its holder to ordinary shares. (AC 104.07)
- Proportionate consolidation** *Eweredige konsolidasie*. A method of accounting and reporting whereby a venturer's share of each of the assets, liabilities, income and expenses of a jointly controlled entity is combined on a line-by-line basis with similar items in the venturer's financial statements, or reported as separate line items in the venturer's consolidated financial statements. (AC 119.02)
- Receivables** Die term word in AC 101 gebruik. Daar bestaan nie 'n algemeen aanvaarde Afrikaanse term daarvoor nie.
- Record** *Verantwoord*.
- Recoverable amount** *Verhaalbare bedrag*. The higher of an asset's set selling price and its value in use. (AC 128.06)
- Redemption premium** *Aflossingspremie*.
- Related parties** *Verwante partye*. Parties are considered to be related if one party has the ability to control the other party or exercise significant influence over the other party in making financial and operating decisions. (AC 126.05)
- Reorganisation** *Reorganisasie*.
- Reporting period** *Verslagdoeningstydperk*.
- Restructuring** *Herstrukturering*. A programme that is planned and controlled by management, and materially changes either the scope of a business undertaken by an enterprise, or the manner in which that business is conducted. (AC 130.11)
- Retirement benefit plans** *Aftreevoordeelplanne*. Arrangements whereby an enterprise provides benefits for its employees on or after termination of service (either in the form of an annual income or as a lump sum) when such benefits, or the employer's contributions towards them, can be determined or estimated in advance of retirement from the provisions of a document or from the enterprise's practices. (AC 116.06)
- Significant control** *Beduidende invloed*. The power to participate in the financial and operating decisions of an economic activity, but without control or joint control over those policies. (AC 119.02)
- Temporary differences** *Tydlike verskille*. Differences between the carrying amount of an asset or liability in the balance sheet and its tax base. (AC 102.09)
- Tax base** *Belastingbasis*. The amount attributable to an asset or liability for tax purposes. (AC 102.09)
- Trade mark** *Handelsmerk*. The book value of a trade mark is regarded as a noncurrent asset.
- Useful life** *Nutsduur*. Either the period of time over which an asset is expected to be used by an enterprise, or the number of production

or similar units expected to be obtained from the asset by an enterprise. (AC 128.06)

Value in use *Gebruikswaarde*. The present value of estimated future cash flows expected to arise from the continuing use of an asset and from its disposal at the end of its useful life. (AC 128.06)

Warrants/options *Bewyse/opsies*. Financial instruments that give the holder the right to purchase ordinary shares. (AC 104.07)

11. Gevolgtrekking

Die woordeboek is 'n baie bruikbare hulpmiddel vir alle gebruikers van finansiële state, maar veral vir studente wat die eerste keer met rekeningkundige terme te doen kry. Die outeurs van enige tegniese navorsings- of handboek oor rekeningkunde is egter uitgelewer aan die snelle verandering as gevolg van die internasionaliseringsproses wat tans steeds aan die gang is. Die proses het egter alreeds ver gevorder en enige veranderinge wat aan die woordeboek aangebring word, behoort 'n redelike lewensduur te hê.

Bronnelys

(Al die dokumente is deur die Suid-Afrikaanse Instituut van Geotrooieerde Rekenmeesters, Kengray, Suid-Afrika gepubliseer.)

1. Afrikaanse dokumente

Jaar	Verwysingsnr.	Naam van dokument
1990	RE 000	Raamwerk vir die opstel en aanbieding van finansiële state
1999	RE 101	Aanbieding van finansiële state
1999	RE 102	Inkomstebelasting
1995	RE 108	Voorraad
1999	RE 117	Eindigende bedrywighede
1999	RE 123	Eiendom, aanleg en toerusting
1999	RE 131	Besigheidsamevoegings

2. Engelse dokumente

Jaar	Verwysingsnr.	Naam van dokument
1990	AC 000	Framework for the preparation and presentation of financial statements
1998	AC 101	Presentation of financial statements

1999	AC 102	Income taxes
1995	AC 103	Net profit or loss for the period, fundamental errors and changes in accounting policies
1998	AC 104	Earnings per share
1995	AC 108	Inventories
1999	AC 110	Accounting for investments in associates
1994	AC 111	Revenue
1998	AC 112	Accounting for the effects of changes in foreign exchange rates
1995	AC 114	Borrowing costs
1996	AC 116	Retirement benefit costs
1999	AC 117	Discontinuing operations
1996	AC 118	Cash flow statements
1999	AC 119	Financial reporting of interests in joint ventures
1999	AC 123	Property, plant and equipment
1997	AC 125	Financial instruments: disclosure and presentation
1997	AC 126	Related party disclosure
1997	AC 126	Interim financial reporting
1999	AC 128	Impairment of assets
1999	AC 129	Intangible assets
1999	AC 130	Provisions, contingent liabilities and contingent assets
1999	AC 131	Business combinations
1999	AC 132	Consolidated financial statements and accounting for investments in subsidiaries
1998	ED 128	Leases

Igor Burkhanov. *Lexicography: A Dictionary of Basic Terminology.* 1998, 285 pp. ISBN 83-87288-56-X. Rzeszów: Wyższa Szkoła Pedagogiczna. Price 16 PLN.

A long-awaited dictionary has at last been published, going beyond a glossary of lexicographic terms (cf. Robinson 1984) or a dictionary of a particular aspect of lexicography (cf. Cluver 1989). It is the first dictionary of basic lexicography terms. In 265 pages Burkhanov gives us a balanced view of the state of the art in dictionary form. It comprises terms typical of lexicographic work, including terms originally related to different linguistic disciplines.

The selection of the data is founded, as can be inferred from the bibliography (p. 267-285), on a solid basis. It covers publications over the last twenty years, deals with various aspects of the field and includes the most recent trends from the late eighties and the nineties, e.g. Bergenholtz and Tarp (1995), Bright (1992), Bussmann (1996), Cowie (1987), Cluver (1989), Diab (1990), Fillmore and Atkins (1992), Ilson (1991), Benson, Benson and Ilson (1986), Landau (1989), Hüllen (1990), Snell-Hornby (1990).

It is not difficult to find one's way in the dictionary: few symbols are used, most important of which are the ones indicating cross-references within the word-list. In this dictionary, though the arrangement of terms is alphabetical, conceptual relatedness between terms is not lost. Related terms are listed at the end of the explanatory text with the guide words "see also" and "compare". "See also" refers to superordinate and subordinate terms, e.g. in the lemma *lemmatization*:

In *lexicography*, the term "lemmatization" is used to refer to the reduction of *inflectional word forms* to their *lemmata*, i.e. *basic forms*, and the elimination of *homography*. It is one of the major procedures of computational lexicography and computational linguistics in general (see *corpora analysis*) which attributes word forms to their *lexemes*. In practice, lemmatization involves the assignment of a uniform heading under which elements of the *corpora* containing the word forms of same lexeme are represented. Lemmatization procedures are a constituent of *concordancing programmes* intended to produce *concordances*, *reverse dictionaries*, and so on. [see also *corpora analysis*; *corpus*; *lemma*; *word form*].

A "compare" reference leads us to co-hyponyms, antonyms or otherwise associated forms of the same rank, e.g. in the lemma *complementarity*:

The term "complementarity" is used in reference to a semantic relationship of opposition, particularly a lexical-semantic relation between pairs of *lexical items* called *complementaries* or *nongradable antonyms*. If two lexical items are complementaries, the negation of one implies the assertion of the other; for instance, 'John is not *married*' implies 'John is *single*'. [compare *nongradable antonym*; *converseness*; see also *antonymy*].

Therefore, what has been sacrificed in the alphabetical arrangement is almost regained through the cross-references, which in a way represent a conceptual system.

Within the entries there are references to:

- (a) bibliographical sources to which the reader could refer, e.g. Denisov (1977), Kipfer (1984), Wiegand (1989) in the lemma of *entry*,
- (b) dictionaries, e.g. *The Concise Oxford Synonym Dictionary* (1995) in the entry for *synonym*, and
- (c) the person who introduced the term, e.g. Landau (1989) in the entry for *frequency label*, Catford (1965) in the entry for *co-text*, and Trier (1931) in the entry for *lexical field*.

The dictionaries referred to are for English, French, German and Russian, and many of them have been published recently (see Bibliography: 281-285), e.g. *The American College Dictionary of the English Language* (1995), *The Facts on File: Visual Dictionary* (1995), *The New Lexicon Webster's Dictionary of the English Language* (1993), *NTC's Thesaurus of Everyday American English* (1995), and *Thesaurus Larousse: des idées aux mots, des mots aux idées* (1992).

On the whole, the cross-reference system is manageable and easy to use. However, an alphabetical index of all terms contained in the dictionary, including "see", "see also" and "compare" cross-references with page numbers, should be included, e.g. in the lemma *lexicographic description* (p. 132), see *macrostructure of a dictionary* (p. 146), see also *lexicographic investigation* (p. 132); *lexicographic presentation* (p. 134), or in the lemma *lexicographic investigation* (p. 132), see also *lexicographic description* (p. 132); compare *lexicographic presentation* (p. 134).

One of the advantages of the dictionary is its reference to Russian scholars and their contributions, e.g. Morkovkin, Akhmanova, Apresyan, Mel'žuk, Ufimceva, Smirnicky, Ščerba, Cyvin and Karaulov. More information on these scholars is needed for non-Russian-speaking users of the dictionary and short notes on their contributions could be included in a new edition.

In his definition of essential terms such as *lexicography* and *lexicology*, Burkhanov gives a balanced, unbiased viewpoint. In *lexicography*, for example — which is a controversial term — he has highlighted different approaches: Richards, Platt and Weber (1992), Ilson (1991), Denisov (1977), Dubois et al. (1973), Akhmanova (1969), Henne (1973) and Wierzbicka (1985). Throughout the dictionary, Burkhanov sticks to his firm belief stated in the preface "not to impose the author's ideas or only one, even predominant, viewpoint on the intended audience, but to introduce the dictionary user into the heart of present-day controversy".

All linguistic terms included are viewed from the standpoint of lexicographic theory and practice, e.g. *descriptive approach*: "Nowadays this term is

also implemented in reference to *lexicographic* products particularly *linguistic dictionaries* and *usage guides*" (p. 61). In his definition of the term *target language*, which is borrowed from applied linguistics, he refers to its use in bilingual lexicography and the meaning it acquired: "In *bilingual lexicography*, the target language is the language whose *translation equivalents* are provided as *definiens* in the entries" (p. 239).

The terms he has chosen come from both the theoretical and practical aspects of lexicology: *synonym*, *antonymy*, *lexis*, *compound*, *thesaurus*, *idiom*, *collocation*, *collect*, *concordance*, *correction file*, *citation slip*, *back matter*, *frame*, *frequency label*, *front matter*, etc.

Typical of his lexicographic treatment is the term *monosemy*:

The term "monosemy" was coined by Bréal (1897) to designate a concept which is opposed to *polysemy*. Monosemy was originally thought to be solely a property of *lexemes*. Nowadays it is usually defined as follows. A *linguistic sign*, i.e. a *lexical item*, a *morpheme*, or a *syntactic structure*, is characterized by monosemy if it has only one *meaning*. [see also *polysemous*; *homonymous*; compare *polysemy*; *homonymy*].

Here he gives the origin of the term, a definition, its opposite, related terms, its present-day usage — all in a well-knit whole. As Picht and Draskau (1985: 62) advocate: "A term may not be viewed as an isolated unit in terminology. Its evaluation and elaboration should always be carried out within the conceptual system, which is in turn closely related to a special subject field or a discipline." This has been Burkhanov's approach throughout the dictionary and this is the strongest point in his dictionary, following sound terminological principles.

One of the assets of the dictionary is the comprehensive treatment of the meanings of the terms provided. Burkhanov provides us with sufficient information about the term, which is a great help to lexicography students, translators and terminologists. A good example is the term *paraphrase* (p. 169-170), whose meanings are given in Hartmann and Stork (1972) and Crystal (1991), but only from a linguistic point of view:

(a) Hartmann and Stork (1972)

PARAPHRASE The process or result of rewording an utterance from one level or variety of a language into another without altering the meaning, as opposed to → *metaphrase* or *translation* which converts the utterance into a different language. Some types of → *language teaching* make use of both *paraphrasing* and *translation* to develop the skills of expressing the same meaning in a number of different ways.

(b) Crystal (1991)

paraphrase A term used in LINGUISTICS for the result or process of producing alternative versions of a SENTENCE or TEXT without changing the

MEANING. One SENTENCE may have several paraphrases, e.g. *The dog is eating a bone*, *A bone is being eaten by the dog*, *It's the dog who is eating a bone*, and so on. Most SEMANTIC theories would treat all these sentences as having a single semantic REPRESENTATION (though variations of FOCUS and PRESUPPOSITION could differentiate them.) Linguists use syntactic paraphrase as a major procedure for establishing certain types of TRANSFORMATIONAL relations. See Clark and Clark 1977: Ch. 1.

Burkhanov deals with the different concepts which are associated with the term and relevant to lexicographic description. He gives the meaning of the term in (a) its broad sense, (b) logic, (c) transformational generative grammar, and (d) linguistic semantics, as well as the influence of all these concepts, whether explicitly or implicitly, on the development of lexicography: "To give but one example, all of the aforementioned ways of paraphrasing the lexical meaning of a lexical item are implemented in *monolingual lexicography*: derivation is made use of in *run-ins* and *run-ons*; replacing a one-word lexical item with a semantically-equivalent word combination is applied in *analytical definitions*, etc. In *metalexicography*, the requirements for paraphrases in logic have been applied to the evaluation of the correlation between *definiendum* and *definiens*."

He also focuses on the present usage of terms and gives the main approaches to the description of lexicographic concepts, e.g. *synonym dictionary*. The term, according to Burkhanov, is used "to designate a large group of linguistic dictionaries". Therefore a synonym dictionary is "any reference work (whatever its title may be) of a wider range of philological dictionaries featuring semantic affinities between quasi-synonyms and other lexical items closely related in meaning". He includes *analogical dictionaries* and *thesauri* amongst these reference works and gives *The Concise Oxford Thesaurus* (1995) as an example.

In some of his definitions, Burkhanov refers to the relevant sense of the term in computational lexicography, e.g. senses 3 and 4 given on p. 14 of *The Concise Oxford Thesaurus*:

3. the meaning in computational lexicography as a reference work containing an alphabetized index of descriptors
4. a stored list of usually one-word lexical items intended to be consulted in the course of computer-writing

Though Burkhanov's definitions meet the requirements of *precision*, *comprehensiveness* and *clarity* (see Heliel 1987), some of them need to be reinforced by visual illustrations. Examples are:

- (a) *analogical dictionary* where examples from Maquet's *Dictionnaire analogique* could have been supplied,

- (b) *explanatory dictionary* where examples from Apresyan et al.'s *English-Russian Dictionary of Synonyms* could have been given, and
- (c) *collocational dictionary* where examples from Benson's *BBI Dictionary of English Word Combinations* could have been provided.

Other terms that need to be supplemented by examples from different dictionaries to serve as visual aids are: *subentry, run-on, subsense, user's guide, article*, etc.

There is no doubt that the goal of the author as stated in the preface and the functions he suggests for a terminological dictionary in a burgeoning field such as that of lexicography, has been achieved. Burkhanov has:

- (a) formulated the basic concepts of the field and their systems, and
- (b) specified the concepts which have been borrowed from linguistic and nonlinguistic disciplines and the modifications which those concepts have undergone before becoming part of lexicographical terminology.

His dictionary may be a positive step towards clarifying and perhaps standardizing many basic lexicographic terms. The dictionary will be an indispensable tool in the hands of its intended audience. Both undergraduate and post-graduate students doing a course in lexicography will treasure it as a guide presenting the basic concepts in the field in simple language and clear design.

References

- Akhmanova, O.S. 1969². *Slovar' lingvističeskikh terminov*. Moscow: Sovietskaja Encyklopedija.
- Benson, M., E. Benson and R.F. Ilson. 1986. *Lexicographic Description of English*. Amsterdam/Philadelphia: John Benjamins.
- Bergenholtz, H. and S. Tarp (Eds.). 1995. *Manual of Specialized Lexicography*. Amsterdam/Philadelphia: John Benjamins.
- Bréal, M. 1897. *Essai de sémantique*. Paris: Hachette.
- Bright, William (Ed.). 1992. *International Encyclopedia of Linguistics*. New York/Oxford: Oxford University Press.
- Bussmann, H. (Ed.). 1996. *Dictionary of Language and Linguistics*. London/New York: Routledge.
- Catford, J.C. 1965. *A Linguistic Theory of Translation: An Essay in Applied Linguistics*. Oxford: University Press.
- Clark, H.H. and E.V. Clark. 1977. *Psychology and Language*. Port Worth/Philadelphia: Harcourt Brace Jovanovich.
- Cluver, A.D. de V. 1989. *Dictionary of Terminography*. Pretoria: Human Sciences Research Council.
- Cowie, A.P. (Ed.). 1987. *The Dictionary and the Language Learner*. Tübingen: Niemeyer.
- Crystal, D. 1991³. *A Dictionary of Language and Linguistics*. London: Blackwell.
- Denisov, P.N. 1977. Ob universal'noj strukture slovarnoj stat'i'. Red'kin, V.A. (Ed.). 1977. *Aktual'nyje problemy u čebnoj leksikografii*. Moscow: Russkij Jazyk.

- Dlab, T.** 1990. *Pedagogical Lexicography: A Case Study of Arab Nurses as Dictionary Users*. Tübingen: Niemeyer.
- Dubois, J. et al.** 1973. *Dictionnaire de linguistique*. Paris: Larousse.
- Fillmore, C.J. and B.T. Atkins.** 1992. Toward a Frame-Based Lexicon: The Semantics of RISK and its Neighbors. Lehrer, A. and E.F. Kittay (Eds.). 1992. *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale: Lawrence Erlbaum.
- Hartmann, R.K.K. and F.C. Stork.** 1972. *Dictionary of Language and Linguistics*. London: Applied Science Publishers.
- Helie, M.H.** 1987. Definitions of Linguistic Terms in an English/Arabic Dictionary. *Dictionaries* 9: 133-148.
- Henne, H.** 1973. Lexikographie. Althaus, H.P., H. Henne and H.E. Wiegand (Eds.). 1973. *Lexikon der Germanistischen Linguistik*, 590-601. Tübingen: Niemeyer.
- Hüllen, W.** 1990. Rudolf Hallig and Walter von Wartburg's Begriffssystem and its Non-/Acceptance in German Linguistics. Schmitter, P. (Ed.). *Essays towards a History of Semantics*, 129-168. Münster: Nodus Publikationen.
- Ijson, R.F.** 1991. Lexicography. Malmkjaer, K. (Ed.). 1991. *The Linguistics Encyclopedia*, 291-298. London/New York: Routledge.
- Kipfer, B.A.** 1984. *Workbook on Lexicography*. Exeter: University of Exeter.
- Landau, S.I.** 1989. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Picht, H. and J. Draskau.** 1985. *Terminology: An Introduction*. Guildford: The University of Surrey.
- Richards, J., J. Platt and H. Weber.** 1992. *Longman Dictionary of Language Teaching and Applied Linguistics*. Harlow (Essex): Longman.
- Robinson, J.** 1984. *Glossary of English Lexicographic Terms*. Kipfer, B.A. 1984: 173-188.
- Snell-Hornby, M.** 1990. Dynamics in Meaning as a Problem for Bilingual Lexicography. Lewandowska-Tomaszczyk, B. and J. Tomaszczyk (Ed.). 1990. *Meaning and Lexicography*. Amsterdam/Philadelphia: John Benjamins.
- Trier, J.** 1931. *Der deutsche Wortschatz im Sinnbezirk des Verstandes: Von den Anfängen bis zum Beginn des 13. Jahrhunderts*. Heidelberg: Winter.
- Wiegand, H.E.** 1989. Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. Hausmann, F.-J. et al. (Eds.). 1989-1991. *Wörterbücher: Ein internationales Handbuch zur Lexikographie / Dictionaries: An International Encyclopedia of Lexicography / Dictionnaires: Encyclopédie internationale de lexicographie*. Berlin / New York: Walter de Gruyter.
- Wierzbicka, A.** 1985. *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.

Mohammed H. Helie
 Department of English
 University of Kuwait
 Kuwait

Latha Menon (Executive Editor). *Microsoft Encarta Encyclopedia Deluxe 2000*. Microsoft Corporation. Price R399.

In the highly competitive world of CD-ROM encyclopaedias, Microsoft has consistently outperformed its competitors with products for the whole family. *Microsoft Encarta Encyclopedia Deluxe 2000* (henceforth Encarta 2000) is the latest flagship of this line of products. The millennium edition is said to be "a modern reference resource that provides a richness of learning experience for people of all ages that extends far beyond what was possible with traditional reference works" (Letter from the editor). This claim reflects the feature which sets CD-ROM encyclopaedias apart from their printed equivalents — their accessibility. The executive editor emphasizes accessibility as one of the central values of the editorial staff of Encarta 2000.

Louw (1999: 283-293) presented a detailed critical analysis of the accessibility of *Microsoft Encarta Encyclopedia Deluxe 1999* (henceforth Encarta 99). In this review the focus will be on changes the editors have made to the outer and inner access structures in order to improve the encyclopaedia's accessibility for the millenium edition. It will be shown that most of these changes have succeeded in this goal, but that there is still some room for improvement.

The first change is visible immediately. In Encarta 99, the opening screen was largely aesthetic in design. It provided links to the main macrostructure, the dictionary, online updating facilities, help files and the inserted media texts. In Encarta 2000, some of the aesthetic design has been sacrificed for added functionality. As the user's primary need would be to extract information from the articles, the editors have included a search aid called the *pinpointer* in a fixed position on the left side on the opening screen.

The pinpointer can search both the rapid and standard outer access structures of the encyclopedia. It is normally used for a type of search that encompasses both these substructures by providing rapid access to the desired lemma when it is entered as a search term in the "Find"-input box. In this regard it can be used to search only the articles, any of the multimedia features, or all the features. The editors have retained the pinpointer's form and have also opted for the same initial-alphabetical listing of lemmas in the main display window of the pinpointer. This maintains the continuity in this line of products and addresses the needs of the encyclopaedia's target users. The position of the pinpointer has, however, been changed. It is no longer presented as "a floating help text" (Louw 1999: 288), but has moved to a fixed position on the left side of the screen. This makes for a less cluttered display on the article screen than was the case in Encarta 99 and once again underlines the editors' commitment to user-friendliness.

Encarta 2000 has also retained the capacity for advanced searches. The search filters have been modified somewhat and have evolved into an effective

and easy-to-use element of the outer access structure, while managing to be stylistically pleasing at the same time.

In the outer access structure there is a definite move towards a more direct and immediate presentation of information, but in the inner access structure the process seems to be in the opposite direction regarding verbally presented encyclopaedic information in longer, more complex articles. Both Encarta 99 and 2000 have *outlines* for subdivided articles. Encarta 99 offered great freedom to the user to customise the "view" functions. The point of departure on the inner search path would always be the article itself. One could choose whether to read the article by scrolling down or to engage the outline and click on the relevant subheading to go to that point in the main text. Encarta 2000, on the other hand, makes the decision for the user. Only the outline is present when the article is activated and not the main text. See for example the article on the Second World War, where the user has to select a subheading from the outline, e.g. "Military Operations", before embarking on the main text. The outline then moves to the left side where it still allows rapid access to subsections of the article by means of active cross-references.

The multimedia elements are initially included in a neat, easily accessible cluster next to the outline. Should the user wish to access these elements first, a click on the appropriate icon will transport the user to that text or to a further grouping of relevant audio or visual clips. If the main text is accessed, icons that act as active cross-references to these multimedia elements are inserted at the appropriate points.

In theory this search path should be sufficient for the needs of any user, but once one has become accustomed to the option of customisation, it is difficult to adapt to such a rigid search pattern. In this case Encarta 99 is certainly a more flexible reference tool than its successor.

The use of the multimedia elements also bears witness to a lack of customisation. Whereas videos and pictures could be resized according to the preference of the user in Encarta 99, this option is not available in Encarta 2000. There is some improvement in the accessibility of multimedia items, though. One of the most important points of criticism levelled against Encarta 99 was that the changing of CDs was an obstacle to the smooth transfer of information. In Encarta 2000 the information is still spread over two CDs, but in general the organisation of information on each CD is much better. One can soon deduce that all the video files are stored on CD 1. The result is that the changing of CDs, if unavoidable, is at least predictable.

Smaller versions of pictures also allow the user to decide before changing the CD. These smaller pictures act as active cross-references to the comprehensive items and maintain the cohesion of the encyclopaedia as a whole.

In shorter articles such as *panpipe*, not much has changed. The colouring and layout is slightly different and the option of customisation is not given. Yet the combination of text plus multimedia is still presented as an effective, cohe-

sive, uncluttered whole which provides optimal access to the target users of the dictionary.

Encarta 2000 is a smooth, cohesive entity, which promises and generally delivers an easy and pleasurable transfer of encyclopaedic information. Hopefully it will continue evolving, with some older elements brought back and some exciting innovations brought in. It will have to do so not only to stay ahead of the pack, but also to stave off the challenges posed by its greatest rival: the Internet. As search engines become more advanced, the Internet is growing closer to fulfilling its potential of becoming a huge encyclopaedia. Until it reaches that point, users of electronic encyclopaedias will have to recognise the Encarta family as state of the art CD-ROM encyclopedias.

Bibliography

- Louw, P.A. 1999. Portals to Knowledge: CD-ROM Encyclopaedias, with Specific Reference to *Microsoft Encarta 99 Encyclopedia*. *Lexikos* 9: 283-293.
- Microsoft Encarta Encyclopedia Deluxe 1999*. 1998. Microsoft Corporation.
- Microsoft Encarta Encyclopedia Deluxe 2000*. 1999. Microsoft Corporation.

Phillip Louw
Bureau of the Woordeboek van die Afrikaanse Taal
Stellenbosch
Republic of South Africa

T.J. Kriel (original compiler), D.J. Prinsloo and B.P. Sathekge (compilers — revised edition). *Popular Northern Sotho Dictionary*. 1997. 335 pp. ISBN 1 86890 001 0. Cape Town: Pharos. Price R37,95.

Preface

The compilers of this new edition have successfully highlighted the important additions to the last edition of the dictionary. It is important to inform prospective users about new information. It is also a marketing strategy to announce the contents of a new product in both the preface and at the back of the cover page, as is the case with this dictionary. This information includes:

- 3 000 new words added
- extensively revised new edition
- look up words directly — no knowledge of grammatical rules presupposed
- revised spelling

With the inclusion of new commonly used words, a few less commonly used ones have been dropped, for example kaffir-corn (lebele, mabele(thoro)), kaffir-tree (mmale, mokhungwane), and Afrika-Gare (Central Africa).

Another important addition is the explanatory notes in the preface on the use of the verbs *thuše*, *thušê* with the negative formatives *ga/sa/se*. This information provides a good guide to second language learners on the usage of the given words.

The use of the circumflex on the low vowels is another major innovation by the compilers. The circumflex serves a dual purpose. Firstly, it assists the dictionary user with the correct pronunciation of words, and secondly, it reduces ambiguity in homonyms, for example: *seroto* (basket, measure, yield, price), and *serôtô* (industrious person, diligent person). The compilers should have highlighted this in the preface.

Structure

The dictionary has a simple structure. The headword appears in bold print. Entries are stem-based and are arranged in alphabetic order, which allows direct look-up. The stem-base arrangement is extremely user-friendly, since no knowledge of grammatical rules is presupposed.

Omissions

The compilers have omitted one important convention, *viz.* that of specifying the target user of the dictionary. This information normally appears in the

preface. It immediately helps the prospective user to decide whether or not the dictionary will provide the information he/she is looking for.

Another omission is the translation equivalent of the word *ramelamo* on page 134.

Inconsistencies

It is important for compilers to maintain consistency throughout in the treatment of lemmas and the presentation of information. This helps to build confidence in the dictionary user concerning the reliability of the information provided in the dictionary. Any deviations which may compromise consistency should be explained in the front matter.

In this dictionary the use of the locative suffix *-ng* and the verbal relative suffix *-go* is inconsistent. The suffix *-ng* has been used with nouns like *monyakong* (in the doorway), *mminong* (in the song/dance), *pelong* (in the heart), but not with *letamo* (dam), *leswafo* (lung), *kgara* (chest, breast), etc.

Likewise, the suffix *-go* has been used with verbs like *mphilego* (who gave me), *ngwadilego* (who is writing/drawing), *kgethago* (who chooses/ elects/picks), but not with *biditše* (has called, named), *ngwala* (write/draw), *bêrêka* (to work), etc.

In the above instances no explanation is given for choosing certain words and excluding similar others that can be used with these suffixes.

The inclusion of *who* in the translation equivalents of words with the suffix *-go* above is not clear. There exists no relationship between *who* and *-go*.

Conclusion

On the whole, this is a good, user-friendly dictionary which contains current words. It could be used by everyone — young and adult, learners and translators — who have an interest in the understanding and use of both Northern Sotho and English.

It is portable and will fit into one's pocket, which is particularly useful for those travelling and for quick reference.

The spelling used conforms to the rules as laid down in the last edition of the *Northern Sotho Terminology and Orthography* of 1988.

Kwena J. Mashamaite
Department of Northern Sotho
University of the North
Private Bag X1106
Sovenga
Republic of South Africa

Wilfrid H.G. Haacke and Eliphas Eiseb. *Khoekhoegowab-English / English-Khoekhoegowab Glossary / Midi Saogub*, 1999, vi + 278 pp. ISBN: 99916-0-172-X. Windhoek: Gamsberg Macmillan. Price N\$86,90 + N\$25,00 (air mail).

The publication of the two-way *Khoekhoegowab-English / English-Khoekhoegowab Glossary* by Prof. Wilfrid H.G. Haacke and Pastor Eliphas Eiseb is a most welcome contribution to the lexicographical literature of Namibia. This glossary, which is an extract from the more extensive *Khoekhoegowab Dictionary* (forthcoming) by the same authors, contains more than 4 300 Khoekhoegowab entries. Apart from reviewing the glossary, the following comments also comply with Prof. Haacke's request (p. v) to supply him with "anything useful for a future edition". The glossary has as its predecessors J. Olpp's *Nama-Deutsches Wörterbuch* (Eberfeld, 1888), J.G. Krönlein's *Wortschatz der Khoi-khoi* (Berlin, 1889) and F. Rust's *Deutsch-Nama Wörterbuch* (Windhoek, 1960). Although the glossary is not solely intended for educational purposes, it is bound to have a significant impact on the teaching and development of Khoekhoegowab in Namibia.

In order to simplify the glossary, the authors excluded regular derivations such as adverbs and abstract nouns, as well as certain compounds. Likewise, not all derivations of synonyms were given, as the users were expected to derive them from the basic entries. The problem of phrases were handled by making separate entries accounting for the most important words of the phrase. Thus an entry such as:

container for cosmetic fat/ointment *n.* †khîna.s; †khêna.s (p. 153)

would also feature under:

fat-container *n.* !gollnâ.s, !gau.s, †khina.s, †khêna.s (p. 172).

As a publication intended mainly for secondary education, the complicated indication of tone was omitted (on public request). This however poses a disadvantage, because mother-tongue as well as non-mother-tongue users would have difficulty in, for example, establishing the points of semantic, i.e. tonal, division, between heterotonal entries such as the one commencing in [gai- (p. 57). Numerous tonal pitfalls are created by the "public request" to omit the marking of tone, e.g.

lû *v.t* charge (p. 95),

lû *v.t* step over (p. 95), and

lû *v.t* graze (p. 95).

The authors supply (p. i) an orthographic list of symbols reflecting the alphabetical order of the Roman as well as the click digraphs and trigraphs. The characters used to represent the clicks in this list and in the Khoekhoegowab-English section of the glossary always appear in medium print, while the Roman characters are in bold. If possible, this incongruous printing style should be avoided in the forthcoming dictionary by using bold fonts for the click characters.

The authors also account (p. ii) for the use of the macron to symbolise "long" vowels and the circumflex to symbolise nasalisation. Explanations of "other conventions" are also supplied (pp. iii-iv). The inclusion of new terms, coined by various committees and marked with an asterisk, is to be commended. The current users of Khoekhoegowab are thereby given the opportunity to consider marked terms, such as the following, for future use in their language:

‡oamâsi.b *n.* weather chart, weather map,
 !hana‡ga.s *n.* agriculture,
 !khurilhuru!â.s *n.* (*Lit.*) one-act play, one-acter,
 !uri !uri.s *n.* pollution, and
 uri‡gâlhôa.s *n.* (*math.*) reflexangle.

Numerous editorial inconsistencies are found in the practical application of the list of abbreviations (p. iv). Although instances of this problem occur in both sections of the glossary, only a few examples, mainly from the English-Khoekhoegowab section, are quoted:

(a) The abbreviation (*med.*) is not consistently applied to mark lexical entries of this nature, e.g. see *amputate* (p. 131), *consumption* (p. 153), *locum tenens* (p. 201) and *polio* (p. 222).

(b) The abbreviation (*jur.*) is not listed on p. iv, but is used in the glossary, e.g. see *impound* (p. 190) and *seize* (p. 240).

Without comparing all the abbreviations used in the text of the glossary with those listed on p. iv, it was found that (*obs.*) and (*anat.*) do not feature in the list on p. iv, e.g. see ‡khani *vi* (*obs.*) (p. 121) and *fold n* (*anat.*) (p. 175).

Incidentally, in the inversion of the first example, the abbreviation (*obs.*) was omitted under *reappear* (p. 229) in the English-Khoekhoegowab section. This is contrary to the intention of the authors as explained in the "Acknowledgements" on p. vi.

(c) The abbreviation (*zool.*) is listed on p. iv, but does not seem to be used in the glossary. One would have expected it to be used in addition to the scientific names for fauna. Whether it would have served a general pur-

pose if it had been used in this glossary is a moot point. There is however one entry on animal names that might have been confusing without the use of this abbreviation, viz. *Damara dik dik* (p. 157).

- (d) The abbreviation (*B.*) which is listed on p. iv, is used haphazardly in the glossary. One would have expected it to accompany entries such as the following: *Calvary* (p. 144) and *Nativity* (p. 211).

Instead of using the abbreviation (*B.*), explanatory notes were added to some entries, e.g. see *cup (for Communion)* (p. 156) and *service — (divine/church)* (p. 241).

Bearing in mind that the glossary will be used for educational purposes, the authors should have accounted for the loanwords from German, English and especially Afrikaans, which feature as the origin of a large number of Khoekhoegowab words, e.g. *bridle* (p. 141), *button* (p. 143), *colt* (p. 150), *debt* (p. 158), *gun (big)* (p. 183), *Tuesday* (p. 266), *uncle — maternal* (p. 267), *watering can* (p. 273).

One would have expected to find a larger number of loanwords from German. The only example which was observed, was *car* (p. 144). There are bound to be more examples of loanwords from English and German and even examples of loanwords from the Namibian Bantu languages.

Quite a number of loanwords from the European languages will contain the letter *l*, e.g. *knitting needle* (p. 196), *laboratory* (p. 197) and *lily* (p. 200). When juxtaposed with the clicks *ǀ* and *ǁ*, the visual contrast with *l* becomes problematic, as can be seen in *school hostel* (p. 239), *school regulations* (p. 239), *school bag* (p. 239), *school bell* (p. 239) and *school boy* (p. 239, etc.). If someone were to use the NewsGothic font, it would be well-nigh impossible to distinguish *l* from the clicks mentioned above. The only solution to this problem is to devise click fonts that would extend below the line on which *l* stands.

This glossary will be extremely useful for mother tongue users as well as for scholars doing comparative studies between the Khoe and Khoekhoegowab languages. The authors deserve the commendation and appreciation of all speakers of Khoekhoegowab for this enormous contribution to a language which is already overshadowed by English in Namibia.

Professor J.W. Snyman
P.O. Box 38653
Faerie Glen
Republic of South Africa

Publikasieaankondigings / Publication Announcements

Igor Burkhanov. *Lexicography: A Dictionary of Basic Terminology*, 1998, 285 pp. ISBN 83-87288-56-X. Rzeszów: Wyższa Szkoła Pedagogiczna. URL: <http://www.univ.rzeszow.pl>. E-mail: wydaw@univ.rzeszow.pl. Fax: 0-48-17-852 20 44. Price: 16 PLN. (Review in this issue)

Igor Burkhanov. *Linguistic Foundations of Ideography: Semantic Analysis and Ideographic Dictionaries*, 1999, 288 pp. ISBN 83-87288-98-5. Rzeszów: Wyższa Szkoła Pedagogiczna. URL: <http://www.univ.rzeszow.pl>. E-mail: wydaw@univ.rzeszow.pl. Fax: 0-48-17-852 20 44. Price: 16 PLN.

Johan Dorfling (Samesteller). *Brugwoordelys (Engels-Afrikaans) en omskrywings van Afrikaanse begrippe*, 2000. Mnr. Johan Dorfling, Jamesrylaan 218, Silverton, 0184.

Madaleine du Plessis. *New Words and previously overlooked ones / Nuwe Woorde en oues wat in die slag gebly het*, 1999, 299 pp. ISBN: 1 86890 016 9. Cape Town/Kaapstad: Pharos Dictionaries/Pharos Woordeboeke. URL: <http://www.pharos.co.za>. Prys: R79,95. (Review article in this issue)

Elektroniese Verklarende Handwoordeboek van die Afrikaanse Taal, 2.00.2000. Perskor. Prys: R399,00. (Resensieartikel in hierdie uitgawe)

Graham Elliott and Jeffrey Rowlands. *Juta's Concise Dictionary of Accounting Terms with their Afrikaans Equivalents*, 3rd edition 1996, ix + 134 pp. ISBN 0 7021 3588 7. Kenwyn: Juta. URL: <http://www.juta.co.za>. E-mail: ctbooks@juta.co.za. Price: R65,00. (Review article in this issue)

Wilfrid H.G. Haacke and Eliphaz Eiseb. *Khoekhoegowab-English English-Khoekhoegowab Glossary/ Mídi Saogub*, 1999, vi + 278 pp. ISBN: 99916-0-172-X. URL: <http://www.macmillan-africa.com/Contacts/Namibia.htm>. E-mail: gmp@iafrica.com.na. Windhoek: Gamsberg Macmillan. Price: N\$86,90 + N\$25,00 (air mail). (Review in this issue)

F. Heyvaert, A. Moerdijk, M. Mooijaart, M. Smits en R. Tempelaars (Red.) *Het grootste woordenboek ter wereld: Een kijkje achter de kolommen van het Woordenboek der Nederlandsche Taal (WNT)*, s.j., 434 pp. ISBN: 90 75566 96 4.

- URL: <http://www.sdu.nl>. Den Haag: SDU Uitgevers / Antwerpen: Standaard Uitgeverij. Prijs: f49.90.
- T.J. Kriel (original compiler), D.J. Prinsloo and B.P. Sathekge (compilers — revised edition). *Popular Northern Sotho Dictionary*, 1997, 335 pp. ISBN 1 86890 001 0. Cape Town: Pharos. Price: R37,95. (Review in this issue)
- Latha Menon (Executive Editor). *Microsoft Encarta Encyclopedia Deluxe 2000*. Microsoft Corporation. URL: <http://microsoft.com/southafrica/contact/email.asp>. Price: R399,00. (Review in this issue)
- Mukumbuta Lisimba. *Kongo Proverbs and the Origins of Bantu Wisdom*, 1999. Gabon: CICIBA (International Centre for Bantu Civilisations).
- A. Moerdijk en R. Tempelaars (Red.). *Van A tot Z en verder ...: Lezingen bij het voltooiing van het WNT*, 1999, 112 pp. ISBN: 90 75566 79 4. URL: <http://www.sdu.nl>. Den Haag: SDU Uitgevers / Antwerpen: Standaard Uitgeverij. Prijs: f24.90.
- Microsoft Encarta 99 Encyclopedia — World English Edition*, 1998. Microsoft Corporation. URL: <http://microsoft.com/southafrica/contact/email.asp>. Price: ±R280 (Standard edition); ±R510 (Deluxe edition). (Review article in this issue)
- F.F. Odendal. *Kernwoordeboek*, 2000, 769 pp. ISBN 0 628 03764 3. Halfway House: Perskor. E-pos: risdon@mm.co.za. Prys: R124,95.
- K.H. van Dalen-Oskam, K.A.C. Depuydt, W.J.J. Pijnenburg and T.H. Schoonheim (Editors). *Dictionaries of Medieval Germanic Languages: A Survey of Current Lexicographical Projects. Selected Proceedings of the International Medieval Congress, University of Leeds, 4-7 July 1994*, 1997, viii + 197 pp. ISBN 2-503-50601-1. Turnhout: Brepols. E-mail: publishers@brepols.com. Price: Belgian f1613.
- Van Schalkwyk, D.J. (Hoofredakteur). *Woordeboek van die Afrikaanse Taal, Deel XI*, 2000, xxii + 935 pp. ISBN 0 9584120 6 4 (plastiekband); ISBN 0 9584120 7 3 (leerband). Stellenbosch: Buro van die WAT. E-pos: wat@wat.sun.ac.za. Prys: R130,32 (plastiekband); R220,38 (leerband).

VOORSKRIFTE AAN SKRYWERS

(Tree asseblief met die Buro van die WAT in verbinding vir 'n uitvoeriger weergawe van hierdie instruksies, of besoek ons webblad: <http://www.sun.ac.za/wat/index.htm>)

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 eenige van die volgende rubrieke geïnklassifiseer word:

(1) **Navorsingsartikels:** Grondige oorspronklike wetenskaplike navorsing wat gedoen en die resultate wat verkry is.

(2) **Beskouende artikels:** Bestaande navorsingsresultate en ander feite wat op 'n oorspronklike wyse oorsigtelik, interpreterend, vergelykend of krities evaluerend aangebied word.

(3) **Resensieartikels:** Navorsingsartikels wat in die vorm van 'n kritiese resensie van een of meer gepubliseerde wetenskaplike bronne aangebied word.

Bydraes in kategorieë (1)-(3) word aan streng anonieme keuring deur onafhanklike akademiese vakgenote onderwerp ten einde die internasionale navorsingsgehalte daarvan te verseker.

(4) **Resensies:** 'n Ontleding en kritiese evaluering van gepubliseerde wetenskaplike bronne en produkte, soos boeke en rekenaarprogramme.

(5) **Projekte:** Besprekings van leksikografiese projekte.

(6) **Leksikonotas:** Enige artikel wat praktykgerigte inligting, voorstelle, probleme, vrae, kommentaar en oplossings betreffende die leksikografie bevat.

(7) **Leksikovaria:** Eengeen van 'n groot verskeidenheid artikels, aankondigings en nuusvystellings van leksikografiese verenigings wat veral vir die praktiserende leksikograaf van waarde sal wees.

(8) **Verslae:** Verslae van konferensies en werksessies.

Bydraes in kategorieë (4)-(8) moet almal aan die eise van akademiese geskifte 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* gepubliseer 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 navorsings-, beskouende 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 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
Postbus 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 kantlyn (ongeveer 2,5 cm) oorskry nie. Manuskrip moet verkieslik in elektroniese formaat as ASCII-tekst, 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 or refer to our web page: <http://www.sun.ac.za/wat/index.htm>)

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) **Research articles:** Fundamentally original scientific research that has been done and the results that have been obtained.
- (2) **Contemplative articles:** Reflecting existing research results and other facts in an original, synoptic, interpretative, comparative or critically evaluative manner.
- (3) **Review articles:** Research articles presented in the form of a critical review of one or more published scientific sources.

Contributions in categories (1)-(3) are subjected to strict anonymous evaluation by independent academic peers in order to ensure the international research quality thereof.

- (4) **Reviews:** An analysis and critical evaluation of published scientific sources and products, such as books and computer software.
- (5) **Projects:** Discussions of lexicographical projects.
- (6) **Lexiconotes:** Any article containing practice-oriented information, suggestions, problems, questions, commentary and solutions regarding lexicography.
- (7) **Lexicovaria:** Any of a large variety of articles, announcements and press releases by lexicographic societies which are of particular value to the practising lexicographer.
- (8) **Reports:** Reports on conferences and workshops.

Contributions in categories (4)-(8) 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 Woordboek 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 research, contemplative 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 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 für eine ausführlichere Wiedergabe dieser Hinweise, oder besuchen Sie unsere Webseite: <http://www.sun.ac.za/wat/index.htm>)

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) **Forschungsartikel**, die grundlegend über neue Forschungsansätze und deren Ergebnisse berichten.
- (2) **Kontemplative Artikel**, die bestehende Forschungsergebnisse und andere Informationen selbständig, interpretativ, vergleichend oder kritisch bewertend wiedergeben.
- (3) **Rezensionsartikel**, die in der Form eines Forschungsartikels eine oder mehrere veröffentlichten wissenschaftlichen Quellen kritisch rezensieren.

Beiträge in Kategorien (1)-(3) werden streng anonym von unabhängigen wissenschaftlichen Experten begutachtet, um ein internationales fachliches Niveau in *Lexikos* zu gewährleisten.

- (4) **Rezensionen**, die veröffentlichte wissenschaftliche Quellen und Produkte, wie z.B. Bücher und Software, analysieren und kritisch bewerten.
- (5) **Lexikographische Projekte**, die vorgestellt werden.
- (6) **Notizen zum Lexikon**, die praxisbezogene Informationen, Vorschläge, Probleme, Fragen, Kommentare und Lösungen hinsichtlich der Lexikographie enthalten.
- (7) **Lexikovarlä**, die unterschiedliche Beiträge, Ankündigungen und Pressemitteilungen lexikographischer Vereinigungen, die dem praktischen Lexikographen wichtig sein können, einschließen.
- (8) **Berichte über Konferenzen und Workshops**.

Beiträge in Kategorien (4)-(8) 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 Freixemplare

Autoren erhalten fünf Sonderdrucke ihrer veröffentlichten Forschungsartikel, kontemplativen Artikel oder Rezensionen-artikel gratis sowie ein Freixemplar der betreffenden Ausgabe. Rezensenten und Autoren von Beiträgen zu den Kategorien Lexikovarlä, Projekte und Berichte erhalten fünf Sonderdrucke ihrer Beiträge gratis. Die Redaktion kann sich jedoch, abhängig von der Art und dem Umfang der Beiträge der letztgenannten vier Kategorien, vorbehalten, dem Autor ein Freixemplar der Ausgabe zu überlassen.

7. Einladung und redaktionelle Adresse

Alle Autoren, die interessiert sind, Beiträge für *Lexikos* zu liefern, sind herzlich willkommen. Sie werden gebeten, ihre Artikel an die folgende Adresse zu schicken:

Der Redakteur
LEXIKOS
Buro van die WAT
Postfach 245
7599 STELLENBOSCH
Republik Südafrika

B. VORBEREITUNG DES MANUSKRIPTS

Ein Artikelmanuskript muß den folgenden redaktionellen Anforderungen entsprechen:

1. Umfang und Format

Beiträge sollen nicht länger als 20 getippte A4-Seiten in zweizeiligem Abstand und mit Randabständen von ca. 2,5 cm sein. Das Manuskript sollte möglichst als elektronischer Text auf einer (360 KB bis 1.44 MB) Diskette vorgelegt werden, entweder im ASCII-Format, oder in formatiertem Microsoft Word (DOS oder Windows) bzw. WordPerfect (DOS oder Windows). Ein Ausdruck des vollständig formatierten Artikels soll mit der Diskette eingereicht werden. Jedem Artikel ist eine Zusammenfassung im Umfang von 150-250 Wörtern beizufügen. Ferner sollen etwa 10-30 inhaltskennzeichnende Stichwörter zu jedem Artikel angegebene werden.

2. Abbildungen

Ein reproduktionsfähiger Satz der originalen Abbildungen, Illustrationen, Tabellen, Graphiken und Diagramme oder Qualitätsdrucke 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.