

AFRILEX -reeks series 6:1996

LEXIKOS 6

BURO VAN DIE WAT

AFRILEX

Lexikos 6

Lexikos 6

Redakteur

Editor

D.J. van Schalkwyk

Resensieredakteur

Review Editor

D.C. Hauptfleisch



African Association for Lexicography

AFRILEX-REEKS 6:1996

AFRILEX SERIES 6:1996



BURO VAN DIE WAT
STELLENBOSCH

Uitgewer Publisher

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

**Kopiereg © 1996 deur die uitgewer
Alle regte streng voorbehou
Eerste uitgawe 1996**

**Tipografie en uitleg
deur Etienne Botha, Tanja Harteveld en
Hermien van der Westhuizen
Bandontwerp deur Piet Grobler**

Geset in 10 op 12 pt Palatino

**Gedruk en gebind deur Rotapress
Stewartstraat 59 Goodwood**

ISBN 0 9583222 9 5

**Geen gedeelte van hierdie boek mag sonder skrif-
telike verlof van die uitgewer gereproduseer of in
enige vorm of deur enige elektroniese of mega-
niese middel weergegee word nie, hetsy deur foto-
kopiëring, plaat- of bandopname, vermikrofilming
of enige ander stelsel van inligtingsbewaring**

Hierdie uitgawe van

Lexikos

in die

AFRILEX-REEKS

word opgedra aan

die nagedagtenis van

Pieter Harteveld

3 Desember 1941 - 8 Maart 1996

redakteur van die eerste vyf

uitgawes van *Lexikos*.

Mederedakteur 1966 - 1993

Bestuurder: Redaksionele Steundienste 1993 - 1996

van die

WOORDEBOEK VAN DIE AFRIKAANSE TAAL

This issue of

Lexikos

in the

AFRILEX SERIES

is dedicated to

the memory of

Pieter Harteveld

3 December 1941 - 8 March 1996

editor of the first five

issues of *Lexikos*.

Co-editor 1966 - 1993

Manager: Editorial Support Services 1993 - 1996

of the

WOORDEBOEK VAN DIE AFRIKAANSE TAAL

Adviesraad / Advisory Board

- Dr. H. Béjoint (Frankryk/France)
Prof. V. February (Nederland/The Netherlands)
Dr. R.R.K. Hartmann (Groot-Brittanje/Great Britain)
Prof. W. Martin (België en Nederland/Belgium and The Netherlands)
Prof. I.A. Mel'čuk (Kanada/Canada)
Prof. M. Schlaefer (Duitsland/Germany)
Prof. J. Taeldeman (België/Belgium)
Prof. P.G.J. van Sterkenburg (Nederland/The Netherlands)
Prof. H.E. Wiegand (Duitsland/Germany)
Prof. L. Zgusta (VSA/USA)

Redaksiekomitee / Editorial Committee

- Prof. W.R.G. Branford (RSA)
Prof. A. Carstens (RSA)
Prof. W.A.M. Carstens (RSA)
Prof. A.D. de V. Cluver (RSA)
Prof. C.J. Conradie (RSA)
Prof. L.G. de Stadler (RSA)
Prof. R.H. Gouws (RSA)
Drs. M. Haron (RSA)
Prof. M.H. Heliel (Egipte/Egypt)
Prof. E.F. Kotzé (RSA)
Prof. J. Lenake (RSA)
Dr. J.S. Mdee (Tanzanië/Tanzania)
Prof. B.M. Mini (RSA)
Mev. P. Mtintsilana (RSA)
Prof. A.C. Nkabinde (RSA)
Mnr. L. Pokpas (RSA)
Prof. F.A. Ponelis (RSA)
Prof. E.H. Raidt (RSA)
Prof. P.H. Swanepoel (RSA)
Prof. G.J. van Jaarsveld (RSA)
Prof. E.B. van Wyk (RSA)

Inhoud / Contents

Voorwoord	xi
Foreword	xiii
<i>D.J. van Schalkwyk</i>	
'n Woord van AFRILEX	xv
A Few Words from AFRILEX	xvi
<i>Rufus Gouws</i>	
Redaksionele doelstellings	xvii
Editorial Objectives	xviii
Redaktionelle Ziele	xix
Navorsingsartikels / Research Articles	
Remarks on the Lexicographical Treatment of Metaforms	1
<i>William Frawley</i>	
Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society	14
<i>Rufus H. Gouws</i>	
A Translation Dictionary of Phrasal Verbs: An Ongoing Project	32
<i>Mohamed H. Heliel</i>	
Dictionaries and Orthography in Modern Africa	44
<i>M. Lynne Murphy</i>	
Ancestor Non-worship in Mampruli	71
<i>Tony Naden</i>	
Die inkortingsneller <i>moet</i> en die woordeboek	104
<i>A.F. Prinsloo</i>	
Using Learner Corpora for L2 Lexicography: Information on Collocational Errors for EFL learners	116
<i>Yukio Tono</i>	

- Textual Condensation in Printed Dictionaries. 133
A Theoretical Draft
Herbert Ernst Wiegand

Beskouende artikels / Contemplative Articles

- Die WAT en etimologie: Word die kringloop voltooi? 159
Willem Botha
- Leksikografiese behoeftevervulling 171
Rufus H. Gouws
- Lexical Culture 184
Reinhard Rudolf Karl Hartmann
- Lexicographical Resources in a Multilingual Environment 199
Willy Martin
- Authenticating the Vocabulary: A Study in 215
Seventeenth-Century Lexicographical Practice
N.E. Osselton

Resensieartikels / Review Articles

- Die negende deel van die *Woordeboek van die Afrikaanse Taal* 233
A.E. Feinauer

Resensies / Reviews

- William Branford. *The South African Pocket Oxford Dictionary of Current English* 272
L.A. Barnes
- J. Combrink en J. Spies. *SARA: Sakboek van regte Afrikaans* 274
Tom McLachlan
- Jonathan Crowther (Editor), Kathryn Kavanagh (Assistant Editor) 280
and Michael Ashby (Phonetics Editor). *Oxford Advanced Learner's Dictionary of Current English*
Ilse Feinauer
- L.G. de Stadler en Amanda de Stadler (Medewerker). 283
Groot Tesourus van Afrikaans
C. van Schalkwyk

Margaret Doyle. <i>the A—Z of Non-Sexist Language</i> Annie Gagiano	294
I.E. du Toit. <i>Tweetalige Polisiewoordeboek. Bilingual Police Dictionary</i> B.F. Smit	297
<i>Longman Dictionary of Contemporary English</i> Geoffrey Hughes	300
N.E. Osselton. <i>Chosen Words. Past and Present Problems for Dictionary Makers</i> R.H. Gouws	303
Rajend Mesthrie. <i>A Lexicon of South African Indian English</i> Rambhajun Sitaram	307
Terminology Committee for Social Work. <i>Vaktaalkomitee vir Maatskaplike Werk. New Dictionary of Social Work. Nuwe Woordeboek vir Maatskaplike Werk</i> W.F. van Delft	311
Della Thompson (Editor). <i>The Concise Oxford Dictionary of Current English</i> Liesel Hibbert	314
G.W.R. Tobias and B.H.C. Turvey. <i>English—Kwanyama Dictionary</i> P.A. Mbenzi	318
Annelize van Rooyen. <i>Die Afrikaanse Naamboek</i> Johan Combrink	320
Leksikovaria / Lexicovaria	
Problems in Swahili Lexicography Albina R. Chuwa	323
Lexicography in Kenya: A Historical Survey M.D. Mageria	330
Publikasie-aankondigings / Publication Announcements	338

Voorwoord

Een van die belangrikste doelstellings van die Buro van die WAT met die uitgee van *Lexikos* in die AFRILEX-reeks was nog altyd 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. Met die bereiking van hierdie doelstelling is in die eerste vyf uitgawes van *Lexikos* sowel as in hierdie uitgawe reeds ver gevorder.

Leksikografiese kommunikasie in Afrika het momentum gekry met die stigting van die *African Association for Lexicography* (AFRILEX) op 14 Julie 1995. Hierdie kommunikasie word tydens AFRILEX-kongresse gestimuleer en dit sal voortaan in 'n hoër mate deur *Lexikos* gestimuleer word, want vanaf hierdie uitgawe is *Lexikos* die mondstuk van AFRILEX. Lede van AFRILEX word uitgenooi om in steeds groter mate deur *Lexikos* deel te neem aan die leksikografiese gesprek.

Die leksikografiese gesprek in Afrika moet deur die leksikograwe, metaleksikograwe, ander belanghebbendes én die leksikografiese projekte voortgesit word. Kontak moet toenemend met ander kontinente en met die leksikografiese verenigings van ander kontinente bewerkstellig word. Koöperatiewe leksikografie is die strategie waardeur die leksikografiese vakspecialiste met mekaar in kontak gebring kan word en *Lexikos* moet daarvoor benut word.

Vanaf volume 4 is *Lexikos* deur die Departement van Onderwys as inkomstegenererende joernaal geakkrediteer. Dit bevestig die standaard van *Lexikos*.

Dr. H. Béjoint van die Université Lumière Lyon 2 in Frankryk word verwelkom as nuwe lid van die Adviesraad van *Lexikos*.

Op 8 Maart vanjaar is mnr. P. Harteveld wat as redakteur van die eerste vyf uitgawes van *Lexikos* opgetree het, oorlede. Hy het 'n groot bydrae tot die totstandbrenging en uitbouing van *Lexikos* gelewer. Die wetenskaplike standaard en die toepaslikheid van die navorsingsbydraes was vir hom van wesentlike belang. As Bestuurder: Redaksionele Steundiens by die Buro van die WAT het hy toegesien dat die aanbod, uitleg en tipografie van *Lexikos* maklik toeganklik vir die leser moet wees. Sy perfeksionistiese ingesteldheid het uit die versorging van elke artikel in *Lexikos* geblyk.

Die reaksie van medewerkers om bydraes te lewer tot hierdie uitgawe ter nagedagtenis aan Pieter Harteveld, was so positief dat sommige bydraes sal moet oorstaan vir publikasie in volume 7 van *Lexikos*. U word hartlik bedank vir u medewerking. Dank word ook uitgespreek teenoor die keurders en almal

wat die redaksie bygestaan het gedurende die verskillende produksiestadiums van die publikasie.

Mnr. D.C. Hauptfleisch, resensieredakteur van *Lexikos*, word bedank vir die nougesette deeglikheid waarmee hy sy aandeel aan hierdie volume afgehandel het. Hermien van der Westhuizen, Etienne Botha en Tanja Harteveld van die afdeling Redaksionele Steundienste in die Buro van die WAT asook kollegas in die redaksionele komponent, word opreg bedank vir hulle onmisbare en toegewyde hulp met die administrasie, redigering, proeflees en elektroniese setwerk van hierdie publikasie.

Lexikos 6 is 'n konstruktiewe huldeblyk ter nagedagtenis aan 'n gewaardeerde kollega.

D.J. van Schalkwyk

Buro van die Woordeboek van die Afrikaanse Taal

Foreword

One of the most important aims of the Bureau of the WAT with the publication of *Lexikos* in the AFRILEX series has always been to create a communication channel for national and international lexicographic discussion, and particularly to serve lexicography in Africa, with its rich language diversity. The first five issues of *Lexikos* as well as the present one have already contributed greatly towards the achievement of this goal.

Lexicographic communication in Africa gained momentum with the founding of the *African Association for Lexicography* (AFRILEX) on 14 July 1995. This communication is stimulated during AFRILEX congresses and it will in future be stimulated to a greater extent by *Lexikos*, because from this issue onwards, *Lexikos* is the mouthpiece of AFRILEX. Members of AFRILEX are invited to share, through *Lexikos*, to an ever greater degree in lexicographic discussion.

Lexicographic discussion in Africa must be continued by lexicographers, metalexicographers, other interested parties, and lexicographic projects. Contact needs to be expanded to other continents and lexicographic societies on other continents. Co-operative lexicography is the strategy whereby lexicographic specialists may be put into contact with one another and *Lexikos* should be used for this purpose.

From volume 4 onwards *Lexikos* has been accredited by the Department of Education as an income generating publication. This confirms the standard of *Lexikos*.

We welcome Dr H. Béjoint of the Université Lumière Lyon 2 in France as new member of the Advisory Board of *Lexikos*.

Mr P. Hartevelde, who acted as editor of the first five issues of *Lexikos*, passed away on 8 March this year. His contribution to the establishment and expansion of *Lexikos* was remarkable. The scientific standard and relevance of the research contributions was of great importance to him. As Manager: Editorial Support Services at the Bureau of the WAT he saw to it that the presentation, layout and typography of *Lexikos* was easily accessible to the reader. His perfectionist attitude was apparent in the editorial treatment of every article in *Lexikos*.

The reaction of contributors to this issue in memory of Pieter Hartevelde was so positive that some contributions will have to be carried over to volume 7 of *Lexikos*. You are sincerely thanked for your co-operation. We also express our gratitude to the judges and everybody who assisted the editorial staff

during the various production stages of the publication.

Mr D.C. Hauptfleisch, review editor of *Lexikos*, is thanked for the care and thoroughness with which he accomplished his task in this volume. Hermien van der Westhuizen, Etienne Botha and Tanja Harteveld of the Division: Editorial Support Services in the Bureau van die WAT, as well as colleagues in the editorial component, are thanked for their vital and dedicated assistance with the administration, editing, proofreading and electronic typesetting of this publication.

Lexikos 6 is a constructive tribute to the memory of an esteemed colleague.

D.J. van Schalkwyk

Bureau of the Woordeboek van die Afrikaans Taal

'n Woord van AFRILEX

Vir enige vakvereniging is dit belangrik om 'n eie tydskrif te hê waarin lede hulle navorsing kan publiseer. Die totstandkoming en vestiging van so 'n tydskrif asook die erkenning van die wetenskaplike gehalte daarvan verg nie net baie beplanning en harde werk nie maar ook baie tyd. AFRILEX is in die bevoorregte posisie om in *Lexikos* 'n mondstuk te kry wat reeds gevestig is as 'n vaktydskrif en wat plaaslike en internasionale erkenning geniet. Vir hulle betrokkenheid by AFRILEX se aktiwiteite en hulle bereidwilligheid om van *Lexikos* 'n gesamentlike onderneming te maak, bedank AFRILEX die Buro van die WAT van harte.

Die entoesiasme waarmee AFRILEX as vakvereniging ontvang is en die sukses van dié vereniging se eerste internasionale kongres wat in Julie 1996 gehou is, gee meer as genoeg rede om te glo dat daar in Afrika 'n wesenlike behoefte aan gestruktureerde leksikografiese samewerking bestaan. In en deur *Lexikos* sal AFRILEX poog om die leksikografiese gesprek in Afrika aktief en produktief te hou. Verskeidenheid wat 'n wesenlike kenmerk van die leksikografie is, sal erkenning kry en daarom moet die klem in die artikels wat in *Lexikos* gepubliseer word, steeds op sowel teoretiese as praktiese aspekte van die leksikografie wees.

Rufus Gouws

Voorsitter: AFRILEX

A Few Words from AFRILEX

It is important for any association to have a journal in which its members can publish the results of their research. However, it takes a lot of time and hard work to produce such a journal and to establish its status as an acknowledged scientific publication. AFRILEX is in the privileged position to acquire in *Lexikos* a well-established journal with a scientific reputation respected both locally and internationally. AFRILEX wishes to thank the Bureau of the WAT for their willingness to produce *Lexikos* as a joint venture with AFRILEX and for their commitment to the goals and activities of this association.

The enthusiasm with which the founding of AFRILEX was greeted as well as the success of its first international conference, hosted in July 1996, gave a clear sign of the need for structured lexicographic co-operation in Africa. With *Lexikos* as its mouthpiece AFRILEX will endeavour to stimulate an active and productive lexicographic discussion in Africa. Lexicography is characterised by its diversity. This will be maintained in *Lexikos* and the papers published in this journal will reflect both the theoretical and the practical aspects of lexicography.

Rufus Gouws

Chairperson: AFRILEX

Redaksionele doelstellings

Lexikos is 'n tydskrif vir die leksikografiese vakspecialis 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 interdisciplinêre 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 vakspecialiste 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 co-ordinate co-operation 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.

Remarks on the Lexicographical Treatment of Metaforms

William Frawley, *Department of Linguistics, University of Delaware, Newark, DE USA 19716*

Abstract: This paper examines the way dictionaries describe metaforms, items that refer to the linguistic system itself (e.g., pronouns and indexicals). The paper first shows the inconsistent grammatical classification of metaforms, which are usually and incorrectly categorized as adverbs, conjunctions, prepositions, or interjections. It is then argued that metaforms should be classified in their currently known grammatical categories: discourse marker (*now*), focus particle (*even, like*), quotative (*like*), and so on. Discussion then turns to the definitions of metaforms, with illustration of their incorrect or misleading semantic characterization and suggestions for capturing their denotations consistently. This examination of the lexicography of metaforms is ultimately a case study in how dictionary making might be informed by judicious use of current grammatical, semantic, and pragmatic theory.

Keywords: ADVERB, COMPLEMENTIZER, CONJUNCTION, DEFINITION, DISCOURSE MARKER, EVEN, FOCUS PARTICLE, FORM CLASS, INDEXICALS, INTERJECTION, JUST, LEXICOGRAPHY, LIKE, METAFORMS, NOW, PART OF SPEECH, PREPOSITION, QUOTATIVE, THEN, USAGE

Opsomming: **Opmerkings oor die leksikografiese hantering van meta-vorme.** In hierdie artikel word die metode bespreek waarop woordeboeke metavorme beskryf, d.w.s. items wat na die taalsisteem self verwys (bv. voornaamwoorde en verwysingswoorde). Die artikel toon ten eerste die inkonsekwente grammatikale klassifikasie van metavorme aan, wat gewoonlik en verkeerdlik gekategoriseer word as bywoorde, voegwoorde, setsels en tussenwerpsels. Daar word dan geredeneer dat metavorme in hulle tans bekende grammatikale kategorieë geklassifiseer behoort te word: diskoersmerker (*now* "nou"), fokuspartikel (*even* "selfs", *like* "soos"), kwotatief (*like* "soos"), ensovoorts. Die definisie van metavorme, met illustrasie van hulle verkeerde en misleidende semantiese karakterisering, en voorstelle om hulle denotasies konsekwent vas te vang, word daarna bespreek. Hierdie studie van die leksikografie van metavorme is eintlik 'n gevallestudie van hoe die maak van woordeboeke beïnvloed sou kon word deur die oordeelkundige gebruik van die hedendaagse grammatikale, semantiese en pragmatiese teorie.

Slutelwoorde: ADVERBIUM, DEFINISIE, DISKOERSMERKER, EVEN, FOKUSPARTIKEL, GEBRUIK, JUST, KOMPLEMENTEERDER, KONJUNKSIE, KWOTATIEF, LEKSIKOGRAFIE, LIKE, METAVORME, NOW, SETSEL, THEN, TUSSENWERPSEL, VERWYSINGSWOORDE, WOORDSOORT, WOORDVORMKATEGORIE

1. Introduction

This paper explores the lexicographical treatment of forms that refer to the

informational system of language itself — either the formal linguistic code or the conditions of exchange in which the code operates. I refer to this class of reflexive forms broadly as **metaforms**.

This investigation has two main purposes:

- A. to show the need for careful form-class judgments of metaforms (and, by implication, forms in general) and the enhancement of these judgments by the use of current grammatical theory;
- B. to show the way consistent definitions can be constructed for terms that defy definition (or seem not to merit it, in some cases) if defining practice is informed by current semantic and pragmatic theory.

2. What are metaforms and do dictionaries include them?

Metaforms denote, all or in part, some aspect of the linguistic system or context of exchange. Typically included in the list of metaforms are indexicals, focus particles, evidentials, discourse markers, certain interjections, quotatives, meta-linguistic forms, and so on (see König 1991, Schiffrin 1987, and Lucy 1993 for illustrations). For example, an essential part of the meaning of indexicals is reference to properties of the system of exchange. *There* means 'some spatial position away from the source of speech,' and so crucially involves the linguistic and contextual system itself.

Indexicals are fairly obvious and uncontroversial metaforms, but there are many other items that are similarly reflexive, and not often understood as such: focus particles, for example. *Even* gets its meaning by signaling that the hearer must add information to that currently within the scope of the particle and consider that new information more important or stronger than any presupposed in the context (König 1991: 69ff.). *Even* is thus a metapragmatic signal to the hearer to examine both the proposition asserted and the context and then explicitly add information that exceeds that encompassed by the particle itself.

While metaforms are semantically productive and determinate, their lexicographical treatment has not followed suit, perhaps most significantly in their very inclusion in the dictionary. As long as the metaform dovetails with an established form-class or can be (erroneously, I will argue) forced into one, it is likely to be included. So no dictionary excludes *even* (adverb — wrongly categorized, I will argue) or *I* (pronoun). But few include the discourse-level metaphor *you know* (*I was, you know, surprised by that*).

Furthermore, form-class recognizability frequently interacts with usage judgments. Metaforms are sometimes excluded because they are incorrectly assigned to certain registers or varieties (e.g., *really* and *huh*), or attributed unsystematic or idiosyncratic distributions (e.g., *like*: see below).

We can see this conspiracy of uncertain grammatical class and usage judgments in the uneven lexicographical treatment of a long-debated form —

like. It has a notorious history in English lexicography, the most obvious illustration of which is its substitution for the ostensibly more refined *as*. The AHD (1992: 1042) puts this amusing history nicely:

Writers since Chaucer's time have used *like* as a conjunction, but 19th-century and 20th-century critics have been so vehement in their condemnations of this usage that a writer who uses the construction in formal style risks being accused of illiteracy or worse.

In modern American English, *like* has undergone a fairly established, but only recently noticed (and "horrifying," to purists), innovation to a metaform in certain meanings and distributions. *Like* now has two meta-uses: as a focus marker and as a quotative verb.

In the former, *like* is a free particle, indicating either attenuation (1a) or significantly new information (1b) (examples from Underhill 1988: 236):

- (1a) But it's like a five, ten minute hike to the cabin
- (b) She like paid for the whole thing ...

In (1a), *like* means 'something like,' and hedges the commitment of the speaker to the information that follows. In (1b), however, *like* means 'and here's the real new information or the point of this exchange, namely ...'; in this sense, *like* is a focus particle. Interestingly enough, these two meanings are closely related both conceptually and historically. But even more important, they are formally unifiable since they are both instances of focus (though Underhill 1988 reserves focus for only the latter usage).

In the latter use, as a marker of quotation, *like* has also drawn some attention (Romaine and Lange 1991: 227):

- (2) And she's like, "Um ... Well, that's cool."

In this function, *like* signals reported speech or thought (or even behavior: "So I was like: [gesture or behavioral demonstration]."). This *like* appears to be a grammatical innovation of an alternate for *say*, *express*, or even *do* (Cf. colloquial *go*: Butters 1980). It is thus a metaform, reflexively signaling the fact of speech, thought, or behavior.

Both uses of *like* are quite widespread. Underhill (1988) reports instances of focal *like* across social class, gender, and age. A literature professor is reported as saying about E.M. Forster: "Then, all of a sudden, he like stopped producing completely" (Underhill 1988: 238). Romaine and Lange (1991) observe a similarly wide distribution of quotative *like*. Although they find it more common among young females (like many grammatical changes), they also report written instances (e.g., from the *New York Times* and *Washington Post*) and some usage in British English. Indeed, the *New York Times Magazine*

(May 5, 1996: 54) quotes Christine Whitman, the current Governor of the state of New Jersey, as follows: "I was like, 'Please don't let my Dad see me ...'"

The linguistic and social distributions of *like* as a metaform should recommend it for lexicographical treatment, but it has been picked up only sporadically. The AHD (1969, 1992) reports only the second focal use, calling it **nonstandard**. This is a curious usage label because the attenuating sense has a long history and is closer to standard. In fact, depending on how the form's semantic history is to be interpreted, it could be argued that the OED CD-ROM (1993) reports related uses and meanings back to the 13th century (Cf. the first sense in the OED CD-ROM (1993)). The BDNE (1973, 1990), which tracks new usage and hence should report these innovations, contains only the second focal use (again as nonstandard). Still, neither the AHD nor BDNE accurately classes or defines the term.

WNWCD (1996: 783) reports the focal use but gives no definition or grammatical class. The RHD (1971) reports only the first focal sense (1a) twice, defs. 18 and 20, apparently differentiating the two by register and syntactic distribution. Chambers (1983) excludes both focal and quotative altogether. The COD (1995: 789) reports the second focal sense (as 'so to speak'), but classes it as an adverb and slang. And if you look closely in the OED CD-ROM (1993), you find both focal senses well attested for over 1 000 years (*like* adv., B.7).

Quotative *like*, unfortunately, loses out completely. It rarely appears anywhere, as far as I can tell, not even in the BDNE. One exception is Spears' (1989) dictionary of American slang, which records both quotative and focal *like*. The appearance of these uses in a dictionary of slang is a striking illustration of the socio-political conditions on choice of entries (see Willinsky 1988 for some interesting discussion in this respect).

The lexicographical success and failures with *like* open a window on dictionary-making itself. Why is there such uneven treatment of metaforms?

3. The form class of metaforms

Part of the problem with metaforms is their grammar. Certainly the present variation and grammatical innovations of forms such as *like* make form-class judgments difficult. But when you look at how metaforms have been treated in dictionaries as a whole, a number of puzzles about their part of speech immediately surface, suggesting that they have never been properly classed and that such errors have been passed down out of lexicographical tradition.

Focus particles are almost uniformly classed as adverbs. The AHD (1992: 979), for instance, categorizes *just* this way, and Chambers (1983: 686), the COD (1995: 737), and the RHD (1971: 775) follow suit. The OED CD-ROM (1993) does likewise, and cites, as supporting evidence, *even*, which it also classes as an adverb, as do the other dictionaries.¹

But there is little to recommend this construal of focus particles as

adverbs. Admittedly, some derive from adjectives, and the adjectival base of some adverbs is well known. *Just* and *even* fit this pattern, but this does not mean that they currently are adverbs. Many other focus particles have non-adjectival and non-adverbial sources, like numerals and verbs, and others are unquestionably nonadverbial in their current manifestation (such as the modal particles: König 1991: 163ff.)

In fact, the relevant counter-evidence to the construal of focus particles as adverbs can be found in the dictionaries' own constructed examples. The AHD (1992: 979) exemplifies the focal meaning of *just* ('merely') with *just a scratch*. The RHD (1971: 775) has a similar exemplification: *just a clerk* and *just one of those things*. The COD (1995: 737) illustrates the focal meaning 'no more than' with *just a minute*. These cannot possibly be adverbial uses, as the failure of the substitution of other adverbs attests:

- (3a) ?? happily a scratch
 - (b) ?? quickly a clerk
 - (c) ?? fast one of those things
 - (d) ?? excitedly a minute
- (Cf. ran happily, quickly, fast, excitedly)

While adverbs are known to have a fairly unrestricted cooccurrence, there are limitations (McCawley 1988: 192). But focus particles have **unrestricted categorical distribution** because their communicative purpose is to mark the status of **pieces** of propositions (or propositions as a whole) with respect to context. In this function, they behave like all scope-bearing items (adverbs included) and have their effects left to right. *Even* or *just* can thus modify a determiner (*I saw even THAT dog*) or a preposition (*Donna went just INTO the street, not ACROSS it*), and *like* can modify a whole sentence (*Like, Bob caught a fish*) or part of one (*Bob caught a, like, FISH*).

Similar conclusions hold for discourse markers, forms that signal speaker and hearer position with respect to the discourse context (*now, well, oh, and then*, e.g. Schiffrin 1987): *Now, I was going to say that ..., Well, I think I'll go, Oh, I was going to say that ..., So you're leaving, then?* Dictionaries usually class these as adverbs, conjunctions, or interjections. The AHD (1992: 1239) reports *now* in this use as an adverb. Chambers (1983: 866) has it as both an adverb and interjection. Curiously, the OED CD-ROM (1993) has it as an adverb, but it cites the very form that Chambers classes as an interjection (*now then*, which the OED CD-ROM (1993) traces back to the year 100 as an adverbial use). WNWCD (1996: 929) has it as both an adverb and interjection, and the COD (1995: 932) has it as only an adverb.

Similar alternate categorizations can be found for the other discourse markers. The COD (1995: 1445, 3c and d) classes two uses of discourse-level *then* as an adverb. But other dictionaries exclude these uses altogether. The RHD does not cite *then* in any of its discourse forms, nor does WNWCD. Per-

haps by this practice, the RHD and WNWCD mean to indicate that the discourse-marker sense is subsumed by, or is indistinct from, the adverbial or conjunctive sense and so not necessary to report. However, this would overlook the unique surface distribution of the discourse marker: *We decided on going, then ...*

Like focus particles, **discourse markers** must be differentiated from adverbs, conjunctions, or interjections (whatever this latter category might be!). They are, to put it baldly, discourse markers. Why not call them what they are?

Perhaps most confusing and hence (ironically) illustrative is again the case of *like*, which has not only a checkered history of inclusion but an equally checkered history of categorization, both in its "standard" uses and in its metaform innovations.

Almost without divergence, dictionaries list a standard use of *like* as a preposition, as in the following expressions (AHD 1992: 1042): *It's not like him to take offense, lived like royalty, looks like a bad year for farmers*. The RHD, WNWCD, and COD echo this classification, as do Romaine and Lange (1991: 244ff.) in their study of the quotative. But I have my doubts about this (as does the OED CD-ROM (1993), which, as far as I can determine, never classifies it as a preposition).

In all its ostensible prepositional uses, *like* fails the standard tests for constituency as the head of a prepositional phrase (PP). A PP can take a pro-form:

- (4) Bob sat *on the chair*, and Bill sat *there*, too

This holds even for apparent confluents of prepositions with conjunctions when prepositions take sentential objects:

- (5) Bill ran *after Bob shot the gun*, and Sam ran *then*, too

But no such substitution is possible for the *like*-phrases that supposedly exemplify prepositional uses:

- (6a) ?? They lived *like royalty*, and we lived *such*, too (AHD)
 (b) ?? He works *like a beaver*, and she works *such*, too (RHD)
 (c) ?? She acted *like an idiot*, and he acted *such*, too (COD)

Another test is movement. PP's allow object extraction and fronting for topicalization, sometimes leaving behind a resumptive pronoun. *Like*-phrases do not:

- (7a) He looks like a good prospect for the job (RHD)
 (b) ?? A good prospect — he looks like it for the job.
 (8a) He looked over the fence for his father
 (b) The fence — he looked over it for his father

Like cannot be a preposition in these standard uses. It behaves much more like a **complementizer**, signaling a full or reduced clause. (Quirk et al. 1978: 727-28, 754-55 call it simply a **subordinator** and note its peculiar behavior, though they also give it prepositional uses (323).) But no dictionary I know of uses the category **complementizer**. Indeed, *that*, uncontroversially a complementizer, is called a **conjunction** in WNWCD (1996: 1386), COD (1995: 1444), and RHD (1971: 1470). But complementizers have quite restricted distributions because they mark certain types of clauses (McCawley 1988) while conjunctions link categories of **any type**. Note that the only possible pro-form for the *like*-phrases in (6) is *so* (9a), which is also an acceptable substitute for a noun clause marked by *that* (9b), but not acceptable as a pro-form for either a clause with a conjunction (9c) or a preposition (9d):

- (9a) They lived *like royalty*, and we lived *so*, too.
- (b) Tom believes *that Tom was here*, and Donna believes *so*, too,
- (c) ?? Tom cried *because Bob was here*, and Donna cried *so*, too.
- (d) ?? Bob sat *on the chair*, and Bill sat *so*, too.

These clear problems with the categorization of standard *like* make me worry about its classification in the focal and quotative — i.e., metaform — uses. The BDNE (1973: 260) classes focal *like* as a conjunction, citing a number of written instances. The OED CD-ROM (1993) does likewise, giving citations that reach back two centuries. The AHD (1969, 1992) calls the focal use an adverb; WNWCD (1996: 782) puts the focal use under the verbal use (*like to died*, which it astutely calls **intransitive**), but it is said to be syntactically empty even though it precedes words phrases and clauses. The COD (1995: 789) has focal *like* as an adverb.

But all these categorizations fail the standard constituency tests, and do so in the very illustrations the dictionaries use to exemplify them. A conjunction must conjoin grammatical objects of like category: *boys and girls*, not ?? *boys and with*. The BDNE cites the following as illustrative of conjunctive *like*: *And I thought like wow, this is for me*. While one might think that this exemplifies the conjunction of two clauses — *and I thought* + CONJ + *this is for me* — note that no other conjunction can appear in the CONJ slot: ?? *And I thought since wow / and wow / if wow, this is for me*. The OED CD-ROM's (1993) examples are equally misleading: *As we say pragmatically in Huddersfield: "C'est la vie, like!"* Here, nothing is "conjoined" by *like*; in any case, no conjunction appears in sentence-final position in English. *Like* is not a conjunction here, but a focus particle.

These examples underscore the kinds of form-class choices made by dictionaries and the way lexicographical tradition reproduces outdated and misleading grammatical theory. Perhaps the most radical lesson of the uneven grammatical treatment of metaforms is the suggestion that form-class judgments should be eliminated from dictionaries altogether. After all, users infrequently consult the dictionary for grammatical class anyway, so why not ter-

minate the practice and thus avoid mischaracterization? This would surely evoke protests from the public, who want to see grammar in the dictionary even if the public does not use the dictionary for such a purpose. It would also evoke protests from the publisher's marketing division, which wants the public to see exactly what the public wants to see (Willinsky 1988 has some revealing comments about the interaction of marketing and editorial decisions in dictionaries).

Still, if form-class categorizations are to be included in the dictionary, they should at least be accurate. Modern linguistics has added to the list of form-class categories, so why not class things as they are? At least four new parts of speech should be included in dictionaries as a consequence of an examination of the lexicographical treatment of metaforms: complementizer, discourse marker, focus particle, and quotative. There is no need to use the term **metaform**, which would be too inclusive and too abstruse to the user. The prefatory material in the dictionary could be easily emended to include these categories — without fancy abstract explanations but with clear illustrations. Some suggestions to this effect are in the last section of this paper.

4. The definition of metaforms

Just as curious as the grammatical treatment of metaforms in dictionaries are their definitions. Sometimes metaforms are defined as if they uncontroversially denote, just like nouns, verbs, and prepositions. At other times, they are defined as if they defectively denote or, worse, are meaningless. These discrepancies can be found both within and across dictionaries, suggesting an uneven treatment of metaforms both in-house and across the profession as a whole.

There appear to be three strategies to define metaforms:

- A. define them in the regular defining formulas of the dictionary and so treat them like all other (normally) denoting items;
- B. define them via a usage or distribution statement, not a formulaic definition, and so call attention to them as a special class;
- C. avoid defining them or define them as **empty**.

The first strategy, normal form, typically applies only to those metaforms that are also easily (and wrongly, see above) put into traditional form-classes. Thus, the AHD (1992: 979) defines the focus particle (*n=née* adverb) *just* as "precisely; exactly ... barely ... merely ... simply ..." The OED CD-ROM (1993) does likewise, subdividing senses by distributions: "exactly, precisely ... used freely before a demonstrative ... as an emphatic expletive, strengthening an assertion ..." WNWCD (1996: 734) follows suit and elegantly captures, in a single definition ("neither more nor less than"), the way the particle focuses both sides of a gradient: 'no less than,' *just great*, and 'no more than,' *just five*.

Similar results emerge from an examination of the definitions of *even*, though again the WNWCD has the most semantically sophisticated treatment. WNWCD (1996: 470) initially defines *even* as "though it may seem improbable; moreover; indeed," and goes on to give clear variations of this essential meaning, all in a strikingly clear definitional vocabulary. From both a linguistic and lexicographical standpoint, this is an admirable characterization of *even* because it is semantically accurate in casting the form as a focus marker, and hence a comparative expression, yet it remains simple and faithful to the dictionary's entire denotational style. (This practice is one of the features that makes the WNWCD an excellent book.)

Discourse markers and other metaforms fare less well in terms of normal form and often evoke strategy #2. Chambers (1983: 866) defines the discourse marker *now* as "used ... in remonstrance ... or taking up a new point ... expressing admonition, warning ..." The AHD (1992: 1239) reports: "Used to indicate a change of subject or to preface a remark ..." The OED CD-ROM (1993) notes: "in sentences expressing a command or request, with the purely temporal sense effaced ... Used to introduce an important or noteworthy point ... Inserted parenthetically, or at the end of a clause, with similar force." Even the WNWCD cannot resist distributional statements, which it uses to precede both the excellent definition of *even* and discourse markers like *well* (1996: 1516): "an exclamation used to express ..."

While this defining style suggests an awareness of these forms as metaforms since they are defined by their conditions of use, the practice has two problematic consequences. First, despite (or because of?) the gesture to pragmatics, this sort of definition often obscures the meaning of the form. What does it mean to say that *now* means 'used in taking up a new point'? All points? Whose points? Only points? Why are the following disallowed?

- (10a) ?? Now, hello. (but Cf. Now, what's your name?)
- (b) ?? She said that, now, she wants to take up a new point, in remonstrance.

Second, this style unnecessarily brackets and calls attention to the meta-use. Indeed, in some dictionaries, metaforms are defined literally in bracketed form. WNWCD (1996: 470) begins the definition of *even* with "*used as an intensive or emphatic particle meaning*," with the usage statement italicized, presumably because all usage labels in WNWCD are in italics. But, in fact, this is a statement of the form-class, not of usage.

The RHD's (1971: 987) definition of *now* also illustrates the practice of bracketing: "(used to introduce a statement or question)."² The discourse marker *now* means, in fact, 'upcoming discourse focused on the speaker with the current moment of speech as an initiation point and so this is a break in the flow of information' (Schiffrin 1987: 228-46). This definition resolves both problems in a single sweep — it is in the spirit of normal definition and incor-

porates the meta-use into the definition proper. Furthermore, this definition can be constructed only with the help of current pragmatic and discourse theory.

Bringing the definitions of metaforms in line with existing defining formulas and removing all trace of special bracketing might improve definition in other ways. Consider the first two senses of the RHD definition of *I* (1971: 704):

-pron. 1. the nominative singular pronoun used by a speaker in referring to himself ... -n. 2. (used to denote the narrator of a literary work written in the first person singular.).

These senses seem unnecessarily confusing to me, mixing grammatical, semantic, and pragmatic facts and bracketing special literary uses. For one thing, *I* is classed as a pronoun elsewhere in the entry before the first definition, so why repeat this grammatical classification in the first sense? (And why change the classification to a noun in the second sense?) For another, the second sense is really a type of the first.

A simple definition for *I* could be given that solves all these problems in a single sense: 'speaker self-reference, whatever the ultimate source of the speech itself, say an author (*I*) writing the speech of a first person narrator (*I*) speaking it.' (See Levinson 1989: 68). Narrator self-reference is a type of recursive speaker self-reference, and both can be accommodated in a straightforward denotation, without mixing criteria and without bracketing metafunction. (Cf. WNWCD 1996: 666, which has a remarkably similar definition.)

The third defining strategy is to avoid definition or to call metaforms **meaningless**. Chambers (1983: 866) defines the metaform *now* as "used meaninglessly ..." (sic) and then goes on to characterize its meaningful uses! The OED CD-ROM (1993) is also guilty of empty definition for *now*: "used elliptically in various ways ..." This practice, quite obviously, is not much help, though it is really not much different in spirit, if not letter, from a distribution statement, which fails to pinpoint the meaning of the term and defers the definition to its collocation. Imagine defining *dog* as if it had no content: "used to identify, describe, or solicit certain animals."

It turns out that all three definitional options can be found in lexicographical treatments of focal *like*. In the BDNE's (1973: 260) initial treatment, *like* is defined in bracketed and empty fashion — "used without a definite meaning" — but then given a determinate characterization of its meaning — "to understate or de-emphasize the word or phrase that follows or precedes it." The later BDNE (1990: 289) defines it in normal fashion. The 1969 AHD brackets it as a pausal or emphatic and assigns it expletive usage (1969: 757); the later AHD (1992: 1042) eliminates the expletive characterization but keeps the same bracketing.³ Even the WNWCD defines focal *like* as "used without meaning or syntactic function," but the same dictionary takes on equally non-standard metaforms and straightforwardly defines them: the discourse phrase *you know*

is defined as "you understand" (748).

These disparities both within and across dictionaries return us to our earlier lessons on form class. Why not treat metaforms as they are? Focal *like* denotes 'attenuated, intensified, or very new or very salient information.' The COD (1995: 789), in fact, captures the first part of this denotation by defining focal *like* as 'so to speak,' a definition also found in BDNE (1990: 289). If focal *even* can be defined simply, accurately, and consistently as a focus particle, why should other metaforms be given different treatment?

5. Conclusion

The foregoing arguments lead to two recommendations:

- A. Increase the number of part-of-speech labels to be more sensitive to the variety of grammatical categories in a language.
- B. Write definitions that define, even if they are to characterize discourse-level or pragmatic forms. Empty definitions — or defining an item as **meaningless** — can be misleading.

With respect to the first, dictionaries might include the following categories and descriptions in the prefatory material:

Complementizer: a form that precedes and signals only the presence of a clause. Unlike a coordinating conjunction (*and*) or subordinating conjunction (*because*), which may optionally mark clauses, a complementizer can signal the presence of a subject clause, and, when marking an object, can sometimes be deleted. Examples of complementizers are (*the fact*) *that*, *if*, and *whether*: *That John bought a new car surprised me* (subject); *I know that/if/whether/0 John bought a new car* (object and deletion). Cf. ?? *Since John bought a new car surprised me* and ?? *I know since/because/and John bought a new car*.

Discourse Marker: a form that indicates how the information is exchanged in a conversation or the speaker's or hearer's position in the exchange. Examples of discourse markers are *now*, *then*, *so*, and *well*: *Now, what were you going to say?*

Focus Particle: a form that modifies any part of speech and signals new information or contrast between what is said and what is assumed to be known or have been said. Focused information is generally stressed or somehow made salient. Examples of focus particles are *even* and *just*: *Even the DOG ate that casserole*, *The dog even ATE that casserole*, *The dog ate even THAT casserole*.

Quotative: a form that indicates the presence of direct or imitated speech, thought, or behavior. Examples of quotatives are *say*, *do*, and, colloquially, *go* and *like*: *She said / went / was like, "What are you doing?"*

These form classes require modern syntactic and discourse theory for proper identification and description. Importing current linguistics and its categories into lexicography is a difficult, uphill battle against the medieval, part-of-speech tradition. But editorial decisions should at least be sensitive to the negative tradeoffs in accuracy of adherence to such a tradition. It is entirely possible to continue to please the public without overdoing the prefatory material by adding abstruse form-classes.

The second recommendation has a motivation similar to the first's. Current semantic and pragmatic theory can help dictionaries be consistent and complete in definitions. It is one thing to exclude metaforms from the list of entries. But it is another thing to include them and then inaccurately capture the denotation. I know of no "meaningless forms" in any language. Even those forms that appear to be entirely grammatical in their meaning and hence "empty" (e.g., *of* in *a picture of Bob*) have determinate senses that are captured by the traditional defining formulas of dictionaries: Cf. WNWCD on *of* (1996: 940). Why not define all metaforms instead of focusing on a subset and characterizing those in terms of distribution statements?

COBUILD has shown that it is possible to have a simple and consistent defining style. It also is possible to have a defining vocabulary that accepts form classes beyond what lexicography has traditionally propagated without overcomplicating the attendant definitions. Lexicography is very up-to-date in its method — witness the prevalence of machine-readable databases. Why, then, perpetuate an outdated semantic metatheory?

By excluding certain metaforms or, more importantly, by including them and treating them as they do, dictionaries make choices that sometimes quietly reproduce social and political sympathies contrary to the otherwise expressed leanings of the dictionaries themselves. Thus a dictionary that includes focal *like*, calls it **nonstandard** or **colloquial**, and then defines it as **meaningless** sends very mixed signals. The dictionary satisfies its expressed obligation to record actual speech, but then it simultaneously devalues that speech by calling it **empty**.

Notes

1. This judgment is not restricted to lexicography. The most recent theoretical semantics text lists some focus particles as adverbs (Larson and Segal 1995: 301).
2. I have been unable to find an explanation of this practice in the RHD's guide to the dictionary.

3. The AHD's definition is incorrect even as a distribution statement. Focal *like* obligatorily collocates with focal stress, not a pause, and must always precede the stress (Underhill 1988, Romaine and Lange 1991).

References

- American Heritage Dictionary of the English Language*. 1969. Boston: Houghton Mifflin. (AHD)
- American Heritage Dictionary of the English Language*. 1992. Third Edition. Boston: Houghton Mifflin. (AHD)
- Barnhart Dictionary of New English since 1963*. 1973. New York: Barnhart/Harper & Row. (BDNE)
- Barnhart Dictionary of New English (Third)*. 1990. New York: H.H. Wilson Co. (BDNE)
- Butters, Ronald**. 1980. Narrative Go 'Say.' *American Speech* 55: 304-07.
- Chambers 20th Century Dictionary*. 1983. New Edition. Cambridge: Cambridge University Press.
- Concise Oxford Dictionary of Current English*. 1995. Ninth Edition. Oxford: Oxford University Press. (COD)
- König, Ekkehard**. 1991. *The Meaning of Focus Particles*. London: Croom Helm.
- Larson, Richard and Gabriel Segal**. 1995. *Knowledge of Meaning*. Cambridge: MIT Press.
- Levinson, Stephen**. 1989. *Pragmatics*. Cambridge: Cambridge University Press.
- Lucy, John (Ed.)**. 1993. *Reflexive Language*. Cambridge: Cambridge University Press.
- McCawley, James**. 1988. *The Syntactic Phenomena of English*. Vol. 1. Chicago: University of Chicago Press.
- Oxford English Dictionary On Compact Disc*. 1993. Second Edition. Oxford: Oxford University Press. (OED CD-ROM)
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik**. 1978. *A Grammar of Contemporary English*. London: Longman.
- Random House Dictionary of the English Language*. 1971. Unabridged. New York: Random House. (RHD)
- Romaine, Suzanne and Deborah Lange**. 1991. The Use of *Like* as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech* 66: 227-279.
- Schiffrin, Deborah**. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Spears, Richard**. 1989. *NTC's Dictionary of American Slang and Colloquial Expressions*. Lincolnwood, Ill.: National Textbook Co.
- Underhill, Robert**. 1988. *Like is, Like, Focus*. *American Speech* 63: 234-46.
- Webster's New World College Dictionary*. 1996. Third Edition. New York: Macmillan. (WNWCD)
- Willinsky, John**. 1988. Cutting English on the Bias: Five Lexicographers in Pursuit of the New. *American Speech* 63: 44-66.

Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society

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

Abstract: In a multilingual society dictionaries play an important role in assisting to achieve communicative success between the speakers of the different languages. Speakers in a multilingual society often employ a bilingual dictionary as the only instrument to meet their lexicographic needs. This implies that a bilingual dictionary becomes a polyfunctional instrument, presenting more information than just translation equivalents. This article focuses on the contents and the presentation of bilingual dictionaries. To achieve the optimal transfer of information for South African bilingual dictionaries, some general problems are identified and discussed. With the emphasis on the user perspective, metalexicographical criteria are used to investigate problems regarding the access structure and the addressing procedures in Afrikaans dictionaries. Suggestions are made to expand the outer access structure and to employ innovative methods, including the use of inserted inner texts, to improve the inner access structure. Changes in the addressing procedures to make provision for the more frequent use of nonlemmatic addressing procedures are also suggested.

Keywords: ACCESS STRUCTURE, ADDRESSING PROCEDURES, BILINGUAL DICTIONARY, COMMUNICATIVE EQUIVALENCE, DICTIONARY, INNER ACCESS STRUCTURE, INSERTED TEXT, LEMMATIC ADDRESSING, LEXICOGRAPHY, NON-LEMMATIC ADDRESSING, SEARCH PATH, OUTER ACCESS STRUCTURE, USER PERSPECTIVE

Opsomming: *Vertalende woordeboeke en kommunikatiewe ekwivalensie vir 'n meertalige gemeenskap.* In 'n meertalige gemeenskap speel vertalende woordeboeke 'n belangrike rol om kommunikatiewe sukses tussen die sprekers van die verskillende tale te help bewerk. Vir taalgebruikers wat in 'n meertalige gemeenskap woon, is 'n vertalende woordeboek dikwels die enigste woordeboek waaroor hulle beskik. So 'n woordeboek moet gevolglik 'n meerdoelige gebruiksinstrument wees met 'n hoër inligtingslading as net die aanbieding van vertalekwivalente. In hierdie artikel word die inhoud sowel as die aanbieding van vertalende woordeboeke aan die orde gestel. Om 'n optimale inligtingsoordrag vir Suid-Afrikaanse vertalende woordeboeke te probeer bereik, word sekere tekortkominge aangedui en bespreek. Met die klem op die gebruikersperspektief, word daar aan die hand van metaleksikografiese bevindinge gekyk na probleme met die toegangstruktuur asook die adreeringsprosedures in Afrikaanse vertalende woordeboeke. Die uitbreiding van die eksterne toegangstruktuur en innoverende aanpassings in die interne toegangstruktuur, onder meer die benutting van toegevoegde tekste, word voorgestel. 'n Verskuiwing in die adreeringsprosedures om vir meer gevalle van nielemmatiese adreering voorsiening te maak, word ook voorgestel.

Sleutelwoorde: ADRESSERINGSPROSEDURES, EKSTERNE TOEGANGSTRUKTUUR, GEBRUIKERSPERSPEKTIEF, INTERNE TOEGANGSTRUKTUUR, KOMMUNIKATIEWE EKWIVALENSIE, LEKSIKOGRAFIE, LEMMATIESE ADRESSERING, NIELEMMATIESE ADRESSERING, SOEKROETE, TOEGANGSTRUKTUUR, TOEGEVOEGDE TEKS, VERTALENDE WOORDEBOEK, WOORDEBOEK

Introduction

Dictionaries are regarded as containers of knowledge and more specifically as comprehensive sources of linguistic information. This has definite consequences for the article structure of general monolingual and bilingual dictionaries, and implies and necessitates the inclusion of a representative variety of information types as microstructural entries in the treatment of each lemma. In spite of all the different linguistic categories accommodated in each dictionary article, the average user still perceives dictionaries primarily as reference works aimed at the transfer of information on the meaning of words. This attitude of the user often influences the lexicographer when making decisions on the inclusion and presentation of information.

The comprehensive transfer of linguistic information in general monolingual and bilingual dictionaries is often impeded by the presence of a semantic bias that dominates the microstructural presentation. In monolingual dictionaries the focus on the definition and the lack of an extensive treatment of other categories such as grammatical information often illustrates this bias. A less explicit but often more aggressive application of the semantic bias can be identified in bilingual dictionaries. This procedure is most probably motivated by the typical usage patterns identified in the average dictionary user's utilisation of a bilingual dictionary. Users focus their lexicographical inquiries on semantic aspects in their search for "the target language meaning of a source language item". Having found a target language item, no further attention is paid to additional information which might be of paramount importance for the correct comprehension or use of the specific item. Users too often eschew the other entries which form an integral part of the overall linguistic treatment of a lemma.

The status of a bilingual dictionary as a source of semantic information is especially evident in a multilingual society. In a monolingual society, bilingual dictionaries are not used as reference works in the day to day linguistic needs of the average member of a speech community. Monolingual descriptive dictionaries are employed for this function. In a multilingual society the use of bilingual dictionaries forms an integral part of the daily communication process. Where the majority of dictionary owners in a monolingual society will possess a monolingual dictionary, a multilingual society gives evidence of the flip side of the coin with bilingual dictionaries being possessed by the majority of one-dictionary owners. They employ their bilingual dictionaries not only as

encoding or decoding aids during translation, but the dictionary functions as the only source of semantic information for both the native and the foreign language treated in the dictionary. If you need to know the meaning of a word from the lexicon of your native language you consult a bilingual dictionary and interpret the meaning from the given translation equivalents. A bilingual speaker can often use a bilingual dictionary with success to obtain semantic information regarding his native language. For the skilful dictionary user who is competent in both languages the listing of translation equivalents without the interference of other microstructural entries can be used for instant semanticising. However, this is only possible where the dictionary user is very skilful. Lexicographers may not rely on the presumed intuition of their users but have to be as explicit as possible to ensure an optimal transfer of information. This implies the inclusion of additional entries not aimed directly at transferring the meaning of the lemma but rather at supporting and supplementing the given translation equivalents.

Although bilingual dictionaries are employed as polyfunctional sources of semantic information, their main function is not a transfer of meaning. Bilingual dictionaries are aids in interlingual translations and have to focus on a treatment that enables the user to render a good and sound translation. The main aim of the dictionary should not only be the establishment of a relation of semantic equivalence between source and target language. Instead, a lexicographer has to endeavour to reach communicative equivalence.

Translation equivalents and translation equivalence

When people consult a bilingual dictionary they seldom realise that the information given is not essentially a statement about meaning but a list of translation equivalents, cf. Louw (1985: 53, 54). The functional status of these translation equivalents is that they may be used in certain contexts to substitute the source language item. Where the specific contexts in which translation equivalents can be used to substitute the lemma are not given as part of the lexicographical treatment, it is hardly possible that the creation of semantic equivalence can lead to the establishment of communicative equivalence.

A relation of equivalence exists between a lemma and the translation equivalent paradigm, that is the selection of translation equivalents presented in the article of any one lemma. Translation equivalence, the first aim of the lexicographer of a bilingual dictionary, implies a semantic co-ordination between a lemma and its translation equivalent paradigm. Unfortunately, lexicographers often restrict their attempts to ensure translation equivalence to the mere listing of a number of target language items. Although these items represent the semantic value of the lemma and create a relation of semantic equivalence, the dictionary user receives no assistance in choosing the correct equivalent for a specific context. This lack of additional information impedes the

possibility to reach communicative equivalence; the form of equivalence that should be the lexicographer's first priority.

The existence of different types of equivalent relations adds to the problems of the dictionary user in his attempts to utilise a bilingual dictionary as a functional linguistic instrument. Where congruence holds as equivalent relation, there is a one-to-one relation between source and target language. In such an article semantic equivalence generally implies communicative equivalence. In the case of a relation of divergence there is a one to more than one relation between source and target language. In these articles the lexicographer has to distinguish between lexical and semantic divergence, and this distinction must have an influence on the microstructural representation, cf. Gouws (1989: 167). Lexical divergence is the equivalent relation between a lemma and a translation equivalent paradigm or a subsection of such a paradigm, where all the members of the paradigm are synonyms. Semantic divergence comes into play when the members of the paradigm are not synonyms because the lemma is a polysemous lexical item and different translation equivalents are needed to represent the different polysemous senses of the lemma. Both types of divergence often occur in one article. In the Afrikaans-English / English-Afrikaans dictionary *Tweetalige woordeboek / Bilingual Dictionary*, the following translation equivalent paradigm is given in the article of the polysemous lexical item **betrokke**:

betrokke "overcast, cloudy, dull (*sky*); clouded (*face*); concerned; relevant; committed, involved ..."

Although the first three translation equivalents are only partial synonyms they can be used to substitute the lexical item **betrokke** in the same context because they represent one polysemous sense of this lexical item. Between the lemma and these target language forms exists a relation of lexical divergence. The other target language items, e.g. **concerned**, **relevant** and **committed**, represent other polysemous senses of the word **betrokke**. This illustrates a relation of semantic divergence.

According to Zöfgen (1991: 2889) the use of an alphabetically arranged dictionary presupposes a specific communication situation, and requires skill in acknowledging lexical shortcomings as well as the ability to find a reference form for them. One of the major problems of the user of a bilingual dictionary arises in articles with divergence as equivalent relation. To ensure communicative equivalence the dictionary article has to accommodate additional entries to support the given semantic information. This will be the focus of the rest of this paper.

Criteria for modern bilingual dictionaries

In utilising the results of the metalexicographical research done in the past two

decades, lexicographers can apply a number of well-formulated criteria to improve the quality of their dictionaries. Kromann et al. (1991: 2713) argue convincingly in favour of three criteria applicable in the evaluation and compilation of bilingual dictionaries. They are the user aspect, the linguistic aspect and the empirical aspect. The user aspect focuses on ascertaining a specific target user group and on the distinctive lexicographical needs of this target group. The empirical aspect includes the establishment of relevant databases and the collection of material, whereas the linguistic aspect demands a systematic approach from the lexicographer in the treatment of equivalent relations between lexical fields of the source and the target language but also in the treatment of other syntagmatic and paradigmatic relations that exist between these fields. Both the empirical and the linguistic aspects may not be isolated from the framework set by the user aspect.

The importance that Kromann et al. (1991) attach to the user aspect is in line with one of the mainstream research areas in present-day metalexigraphy, i.e. dictionary use and the focus on the dictionary user, his needs and reference skills, cf. Hartmann (1989; 1989a), Kühn (1989) and Zöfgen (1994). The position of the user is so important that Hartmann (1989: 103) hypothesises that an analysis of users' needs should precede dictionary design. This implies that the design of a dictionary in terms of both its macro- and its microstructure has to be co-determined by the needs and reference skills of the intended target user. The user perspective can have such a comprehensive influence on the design of a dictionary that it could alter its proposed typological status. In meeting the needs of a well-defined target user, the lexicographer might find it necessary to deviate from the traditional criteria for certain dictionary types. Adhering to the demands of the user perspective may lead to a reassessment of existing typological classifications and to a more precise functional differentiation within these categories. The use of bilingual dictionaries in a multilingual society necessitates a comprehensive profile of the target user, his needs and reference skills. Divergence compels the lexicographer to present the target language information in such a way that the intended user will be able to make an unambiguous choice of the correct equivalent for any specific context.

A reassessment of the category of bilingual dictionaries gives evidence of the need for a detailed subtypological categorisation. This topic will not be discussed in this paper, although certain remarks, suggestions and proposals will underline the need for such a refined typological classification.

Bilingual dictionaries: a typological perspective

Considering the different needs and reference skills of users when consulting a bilingual dictionary, one has to argue in favour of a differentiation within this typological category. One bilingual dictionary cannot be everything to everyone. The collection of Afrikaans bilingual dictionaries, that is bilingual dictionaries with Afrikaans as one of the members of the language pair to be treated

in the dictionary, displays a lack of subtypological differentiation. This is a symptom of a more general problem, which influences many dictionaries in many ways, caused by an insufficient linguistic and metalexicographical basis.

Although a distinction is made between different types of Afrikaans bilingual dictionaries, the metalexicographical motivation for this distinction is insignificant. Too often an Afrikaans bilingual dictionary professes to be compiled for use at school level. However, an evaluation of the dictionary gives evidence that the microstructural treatment is not directed at the needs of the target user. School dictionaries are often produced as spin-offs from general dictionaries. In many of these cases, cf. Gouws (1995: 19; 23), the lexicographer diminishes the macrostructure of the existing general dictionary and omits some of the microstructural entries to produce a new dictionary without any real attention to the needs of the new target user. Each typological category has to be compiled with the specific target user in mind, adhering to specific macro- and microstructural demands and criteria.

According to Zöfgen (1994: 16) the conception of future dictionaries will have to follow a lexicographical-pragmatic approach orientated at at least three parameters, i.e. the user, the situation in which the dictionary will be used and the nature and extent of the information to be treated in the dictionary. This approach deviates from a lexicographical tradition which only focuses on the transfer of information, and it emphasises the importance of the user in modern-day lexicography. There often exists a gap between the claims of publishers and lexicographers and the real efficiency of a dictionary for its specific target user. Zöfgen (1994: 245) argues that the usage value of a dictionary is often inversely proportional to the promises of the publisher. Within a multilingual society, the optimization of the added value which a bilingual dictionary can have for its target user depends on the lexicographer's knowledge of the real needs of that user.

The access structure of bilingual dictionaries

Dictionaries display three major structural components, i.e. the *macrostructure*, *microstructure* and *access structure*, cf. Hausmann and Wiegand (1989). The macrostructure can be regarded as the collection of lemma signs being the basic treatment units of the dictionary, whereas the microstructure is the total set of data categories following the lemma sign. The dictionary user has to get to the information he needs, and the structure of the dictionary should help him to reach that information. This component of the dictionary is known as the *access structure*. A distinction is made between the *outer access structure* and the *inner access structure*. The outer access structure guides the user up to the lemma, whereas the inner access structure helps the user to reach the specific information category within the article. The access structure can also be described as the *search path* for the dictionary user.

One of the possibilities a lexicographer should consider to enhance the user-friendliness of his dictionary, is to improve the quality of the access structure in such a way that the search path can lead to an instant retrieval of information. Besides the distinction between the outer and the inner access structure, a lexicographer should pay attention to a third possible component in the access structure, i.e. a component providing a search path from the inner access structure to the communicative context in which the specific lexical item is to be employed. This component can be called the communicative access structure or the exit structure.

Where a bilingual dictionary is compiled for users who are L1 speakers of one of the languages of the dictionary, with a certain level of efficiency in the second language and who share the same cultural background, the traditional outer access structure of a monoaccessible bilingual dictionary would be quite sufficient. This would apply in a bilingual country where both languages are compulsory school subjects and where every speaker has a relative good command of the second language. This situation motivated the present state of bilingual Afrikaans-English / English-Afrikaans dictionaries. In the previous political dispensation Afrikaans and English were the only two official languages of South Africa. Afrikaans- and English-speaking South Africans had a fairly good command of both languages and the Afrikaans-English / English-Afrikaans bilingual dictionaries relied, admittedly often too much, on the knowledge of the intended target user. This approach to bilingual dictionaries has not only had a detrimental influence on Afrikaans-English / English-Afrikaans dictionaries, but it impeded the development of other South African bilingual dictionaries due to the role-model influence of the leading Afrikaans-English dictionaries, cf. Gouws and Ponelis (1992) and Gouws (1993). A lexicographical result of this situation is that a single search path in both the outer and inner access structure was a typical characteristic of Afrikaans-English / English-Afrikaans as well as other bilingual South African dictionaries.

In the new political dispensation South Africa has eleven official languages, i.e. Afrikaans and English; the Nguni languages isiXhosa, isiZulu, siSwati and isiNdebele; the Sotho languages Sesotho sa Leboa and Sesotho, and Setswana as well as Tshivenda and Xitsonga. This has definite implications for the nature and extent of bilingual dictionaries. One of these implications is the fact that lexicographers compiling bilingual dictionaries may no longer assume their target users to be familiar with both the languages treated in the dictionary. This leads to a situation where dictionary users need more guidance to ensure a successful retrieval of information. One way of achieving this is to enhance the accessibility of the dictionary by improving the quality of the access structure.

Improving the outer access structure

Changes to the outer access structure could be done in various ways. On the

one hand the lexicographer could add special entries or alter the presentation of macrostructural elements and, on the other hand, the lexicographer could include additional access structures to create a poly-accessible dictionary. Existing dictionaries can be used to illustrate both these options.

Special entries in the form of structural markers as well as a typographical innovation to change the face of certain lemma signs have been employed by the bilingual learner's dictionary *Kenkyusha's Lighthouse English-Japanese Dictionary*. This dictionary employs asterisks on a systematic basis as structural markers to indicate lemmas with a high usage frequency and to give the user access to the core vocabulary. This system grades certain lemma signs with between one and four asterisks. One asterisk indicates that the lemma belongs to the 3 000 basic words of general usage; two asterisks indicate membership of the first 2 000 basic words of general usage; three asterisks mark a lemma as part of the first 1 000 basic words for high school usage whereas four asterisks signal a word as belonging to the first 1 000 basic words for junior high school. In addition to this system the *Lighthouse Dictionary* gives rapid access to the last two categories of words by presenting the lemma signs in a larger letter type and coloured red. By merely following the outer search path a user looking for this type of information can easily distinguish these lemmas from the rest of the macrostructure without having to resort to an article internal search.

The *Reader's Digest Afrikaans-Engelse Woordeboek / English-Afrikaans Dictionary* (hereafter abbreviated as RD) owes its poly-accessible character to the inclusion of additional outer access structures. The central word list is the same text used in another bilingual Afrikaans-English / English-Afrikaans dictionary, i.e. *Tweetalige woordeboek / Bilingual Dictionary* (hereafter abbreviated as TW). One of the major problems of TW is that it has traditionally been directed at a bilingual target user. This dictionary, albeit one of the best members of its typological category, demands too much linguistic and pragmatic insight from a user who is unfamiliar with one of the treated languages. The editor of RD has detected this problem, but a contractual stipulation forbids changes to the central word list. This restriction has proved to be a lexicographical bonus because it has led to the introduction of an innovative alternative to enhance the attempt at communicative equivalence.

The original macrostructure of RD is presented in two columns occupying the centre of each page. An additional column is included on either side of the central word list. These columns contain two kinds of entries — clearly distinguished from each other. On the one hand, and of less importance for the present discussion, there are boxes called "Words in action" in which a word, taken from the macrostructure, is repeated and supplied with additional information on pragmatics, semantics, grammar, etc. The information in these boxes is given against either a blue background ("Words in action", in the Afrikaans-English component where the information, given in English, refers to Afrikaans words) or against a pink background ("Woorde in aksie", in the English-Afrikaans component where the information, given in Afrikaans, refers to

English words). On the other hand these columns contain an additional access structure which is a second, much reduced macrostructure. This supplementary text contains lemmas also treated on the same page in the main text of the central word list.

The treatment in the articles of this supplementary text is directed at the user who cannot choose the correct translation equivalent from the presentation in the article of the main text. One of the problems of the translation equivalent paradigm in the main text, which is actually a problem of TW, is a lack of contextual information to support the translation equivalents. Ever so often an uncommented list of translation equivalents gives no assistance to the user who does not know which equivalent will be the correct choice for the specific context in which he needs to use the word. The specific contents of these supplementary text articles will be commented on in another section of this paper. Here it will suffice to mention that the focus in these articles is on a retranslation of the translation equivalents.

Although the front matter of RD does not give criteria for the selection of lemmas to be included in this secondary macrostructure, it can be argued that semantic divergence is one of the dominating reasons for the selection. What is important here is the information conveyed to the user on the search path of the outer access structure. RD allows access to the main text as well as to the supplementary text. Being familiar with the system employed in this dictionary, the user who consults the dictionary for a translation equivalent of a specific word, will follow the primary search path to the main text lemma he is looking for. Knowing that RD offers a retranslation of certain lemmas where an insufficient treatment of semantic divergence as equivalent relation impedes the retrieval of a communicative equivalent, the user should turn to the supplementary text to see if the lemma of the main text is also presented in the supplementary text. If there is no such lemma the user can proceed to the inner access structure of the main text article. The occurrence of the wanted lemma as part of the supplementary macrostructure will lead the user to the inner search path of that article where a more comprehensive treatment is given. The inclusion of this additional access structure increases the functionality of the dictionary because it helps the user to a quicker and unambiguous retrieval of the desired information.

Improving the inner access structure

A variety of information categories have to be provided for by the microstructure of both bilingual and monolingual dictionaries. A microstructure typically displays a linearly ordered set of information categories and the search path of the inner access structure leads a dictionary user from the top of this ordering down to the bottom. One of the many problems experienced by dictionary users is the unclear division between information categories. Especially in a

dictionary with a high density of information it takes time to identify a specific microstructural category before being in the position to retrieve the necessary information.

A user-orientated dictionary should try to empower its users by constructing the article in such a way that a quick and unambiguous retrieval of information, preceded by a successful identification of the wanted microstructural category, is always possible. This means that the user has to be familiar with the inner structure of the articles in the dictionary and with the order of the different article positions. The inner access structure can be improved by providing structural markers to indicate the borders between information categories as well as the position and scope of different search areas.

In the Afrikaans-English / English-Afrikaans *Tweetalige Aanleerderswoordeboek / Bilingual Learner's Dictionary* (hereafter abbreviated as TAW), the inner access structure contains structural markers to indicate different types of microstructural categories. Following the system used in monolingual dictionaries to differentiate between polysemous senses, TAW gives a number for each translation equivalent representing a different polysemous sense of the lemma. The occurrence of collocations is always preceded by a diamond, and the search area containing article positions which accommodate the part of speech and information about forming plurals or other derived forms, is marked by a small square. In the descriptive monolingual Afrikaans dictionary *Verklarende Handwoordeboek van die Afrikaanse Taal* the inclusion of cited examples is marked by an upside down triangle. These structural markers are part of the access structure and their function is to guide the user to a specific information category.

In his analysis of the lexicographical text formation Wiegand (1995) makes provision for the inclusion of *inserted inner texts* as part of the article, and he regards this type of text as especially appropriate for learner's dictionaries. These inserted inner texts can easily be distinguished from the rest of the article because they are presented in a different manner, e.g. in boxes or tables. If the lexicographer wants to draw the attention of his intended target user to a specific aspect of a lemma, the inclusion of an inserted inner text will be a clear milestone on the search path of the inner access structure. This system is employed in TAW for the presentation of additional information, mostly of a pragmatic nature, not covered by the defined categories in the ordered article positions. The articles of the lemmas **meat** and **scissors** contain the following inserted inner texts respectively:

The meat from some animals has a different name from the animal itself: the meat from a cow is called **beef** and that from a pig **pork**, but the meat from a lamb is called **lamb**. For fish and for birds such as chicken or duck the same word is used for both meat and animal.

scissors take a plural verb, but a **pair of scissors** is singular. "Where are the scissors?" This pair of scissors is mine.

The inserted texts in TAW are not only characterised by their occurrence on microstructural level, but also by their macrostructural functioning. Inserted texts help to constitute the macrostructure and form part of the outer access structure. The lemma **examination** in TAW is preceded by an inserted text as an additional article:

exam is an abbreviated, informal word for **examination** (its plural is **exams**)

Between the articles of the lemmas **hanger** and **happen** the boxed article of the lemma **hanky** is included as an inserted text:

hanky is an informal word for **handkerchief** (its plural is **hankies**)

These inserted texts add some diversity to the macrostructure. Their boxed presentation within the outer access structure enhances the procedure of identifying a certain category of macrostructural elements.

The kind of article formation where inserted texts are used, illustrates an approach that is aimed at a more comprehensive transfer of information with a strong communicative focus. In a multilingual society lexicographers will do well to include inserted inner texts in their attempt to enhance the communication skills of the intended target user of their dictionary. The use of inserted inner texts is not restricted to bilingual dictionaries. In monolingual dictionaries, especially learner's dictionaries, these constituents can be functional text components. In the monolingual Afrikaans learner's dictionary *Basiswoordeboek van Afrikaans* text boxes add valuable pragmatic information, cf. the use of one of these inserted inner texts in the article of the lemma **afbrand**:

Afbrand word meestal ten opsigte van geboue gebruik. 'n Motor wat deur vuur vernietig word, sal uitbrand.

(**Afbrand** (to burn down) is generally used to refer to buildings. If a car is destroyed by fire it will burn out.)

The inner access structure should also contain information to guide the user to related articles or to the practical application of microstructural information in the real language context. To enable users following the search path of the

inner access structure to achieve rapid access to these microstructural entries, the lexicographer should also employ structural markers to indicate the relevant article positions. TAW uses arrows to refer a user to another lemma. In the article of the lemma *cow* the user is directed by means of an arrow to the note (the inserted inner text) in the article of the lemma *meat*, whereas the only article position in the article of the lemma *till* is the arrowed reference to *until*. In the *Kenkyusha's Lighthouse Dictionary* an even more explicit structural marker is used for intra-article reference. A pointed hand precedes the cross-reference and forms a clearly noticeable signpost on the inner search path.

Intra-article structural markers place the lemma and its translation equivalents into communicative contexts while inter-article structural markers help the user to see further than the isolated lemma by displaying the lexicon as a network of semantic relations and the dictionary as a text with a definite cohesion.

The addressing procedure in bilingual dictionaries

According to Hausmann and Wiegand (1989: 328) the way in which a form and information relating to that form are brought together is the *addressing procedure*. Each information item is addressed to a form called the *address*. In a dictionary article the addressing procedure occurs when information is related to the lemma (the form). The lemma is the most typical lexicographical address and is always the target of *lemmatic addressing*. However, information can also be directed at other components of the article and display a non-lemmatic addressing procedure. In bilingual dictionaries lexicographers should give much more consideration to the use of non-lemmatic addressing procedures. The implementation of these procedures can add to the enhancement of communicative equivalence. This aspect will be discussed in a following section.

Lematic addressing

Linguists working in the field of lexical semantics have drawn attention to the existence of lexical gaps, cf. Lehrer (1974) and Lyons (1977). Lexical gaps are language-specific and this poses problems to the lexicographer compiling a bilingual dictionary. Dagut (1981) identifies different types of semantic gaps in the transfer of a text from one language to another. He distinguishes between gaps due to linguistic factors and gaps due to extra-linguistic factors. These two categories are called *linguistic* and *referential* gaps respectively.

A linguistic gap exists where the speakers of both languages are familiar with a certain object but where the one language does not have a word to refer to the object, whereas the other language does have such a word. A referential gap can be postulated where a lexical item from language A has no translation

equivalent in language B because the speakers of language B do not know the referent of the lexical item from language A.

To achieve communicative equivalence a lexicographer has to identify potential lexical gaps and distinguish between linguistic and referential gaps. This distinction determines the extent of the additional information needed for an optimal transfer of semantic and pragmatic information. Lexical gaps in the target language of a bilingual dictionary compel the lexicographer to employ a more detailed lemmatic address. Linguistic gaps offer less serious problems to a lexicographer because the speakers of the language with these gaps are familiar with the object to which the non-existing word would have referred. Quite often a loanword is used as a surrogate equivalent. To ensure a proper transfer of the meaning of the lemma, lexicographers complement the surrogate equivalent by including a brief description.

The game of rugby is known to speakers of all the different South African languages. However, all these languages do not have translation equivalents for this word and the loanword is often used as a surrogate equivalent. The monodirectional Afrikaans-Zulu dictionary *Afrikaans / Zoeloe-woordeboek*, compiled for the South African market, has the lemma *rugby* with the surrogate equivalent in Zulu *iragbhi*. This equivalent is followed by the brief explanation *ibhola lombhoxo* which literally means "the ball of an oval object." This entry has a lemmatic address because it explains something of the meaning of the lemma. The Nguni languages have a word *ubuntu* which has no one-to-one translation equivalent in English. The best attempt would be *human-heartedness*. In the *Dictionary of South African English* the equivalent *human-heartedness* is given, but this entry is followed by the brief explanation "quality embodying all the traditional virtues and values of *isintu* ..." The word *isintu* refers to African characteristics, style, values and traditions. The inclusion of the explanation enhances the communicative equivalence because the user will know that the translation equivalent *human-heartedness* is not an exhaustive equivalent of the lemma. Once again an additional lemmatic addressing procedure is employed to ensure a more comprehensive treatment.

Even though the speakers of a certain language may be well familiar with a specific referent where their language has a linguistic gap, the lexicographer should, once again, not rely on the intuition of his users, but by means of additional lemmatic addressing procedures he has to be as explicit as possible in the transfer of information to ensure a sound understanding of the lemma as well as the communicatively correct application of a possible surrogate equivalent. The failure to do so can cause real problems. Afrikaans-speaking South Africans are well familiar with the referential value of the lexical item *fat cat*. However, Afrikaans has a linguistic gap with no real translation equivalent for this item. Where no satisfactory surrogate equivalent can be found, the lexicographer of a bilingual English-Afrikaans dictionary should restrict the treatment of the lemma *fat cat* to a brief description addressed at the lemma. TW treats this lemma by giving the surrogate equivalent *rykaard* (person of wealth).

Besides the semantic differences between **fat cat** and **rykaard** there is also a difference in register. No additional information is given to convey the real meaning and register of the lemma. The lack of either an additional lemmatic addressing procedure or a non-lemmatic addressing procedure to give more information on the surrogate equivalent impedes communicative success.

Because a referential gap implies that the speakers of the language are not familiar with a certain referent, the lexicographer is compelled to treat the lemma in a more comprehensive manner. The multilingual South African society is also a multicultural society. Cultural traditions amongst the speakers of a language group are often unknown to speakers of the other languages. This leads to referential gaps and poses many a problem for the lexicographer. The lexicographer should still endeavour to find a surrogate equivalent but the article has to contain an additional microstructural treatment of the lemma to ensure the retrieval of the correct information.

The Nguni word **hlonipa** "refers to a system of reverence and taboos observed by the Nguni woman towards her male relatives-in-law, involving a whole substitute vocabulary of (hlonipa) words to avoid speaking the radical syllable of any one of their names ..." (*Dictionary of South African English*). **Hlonipa** is used as a loanword in South African English and Afrikaans. Yet the treatment of this lemma in a bilingual Zulu / Xhosa-Afrikaans / English dictionary needs additional lemmatic addressing in the article of this lemma. By merely giving the loanword or an unsatisfactory surrogate equivalent, e.g. *taboo*, the user is not equipped with sufficient information.

Non-lemmatic addressing

The addressing procedure in both monolingual and bilingual dictionaries is dominated by lemmatic addressing. Although the lemma remains the main treatment unit of a dictionary with the majority of microstructural entries directed at the lemma, users also need information regarding other microstructural items, e.g. translation equivalents. In a bilingual dictionary lemmatic addressing keeps the lexicographical focus only on the source language item, and ignores one of the most fundamental differences between bilingual and monolingual dictionaries in terms of the treatment of semantic information. Semantic information in a monolingual dictionary focuses necessarily on the lemma. Hansen (1990: 97) argues that the presentation and ordering of translation equivalents should be determined by the equivalence relations between the lemma and the translation equivalents and not by the semantic structure of the lemma, as is the case with the ordering of polysemous senses in a monolingual dictionary. The co-ordination of two languages in one reference work often necessitates the elucidation of some of the target language items. Hansen's formulation already indicates a focus shift in the presentation of semantic information in a bilingual dictionary. This is ascertained by Hausmann and Werner (1991: 2730), according to whom the principle of meaning

discrimination applied in a monolingual dictionary should be replaced in bilingual dictionaries by a principle of equivalent discrimination. Such an approach does not eschew the importance of the presentation of semantic information in bilingual dictionaries. It rather represents a shift in the semantic focus from the lemma to the equivalent. While the mere listing of translation equivalents without any complementary or supporting entries adheres to a lemmatic addressing procedure, the focus shift towards the translation equivalent paradigm implies an article-internal non-lemmatic addressing procedure. Non-lemmatic addressing procedures in bilingual dictionaries are mostly directed at the translation equivalent. Equivalent discrimination still implies the transfer of semantic information and even the discrimination of meaning. This discrimination is directed at the need of the dictionary user to get familiar with the target language items.

The habit of focusing on the meaning of the lemma has led to the presentation of semantic equivalence as the prime objective of the lexicographical treatment in bilingual dictionaries. This lemmatic addressing eschews the need for sufficient contextual evidence which is a type of information primarily directed at the translation equivalent and therefore constitutes a non-lemmatic addressing procedure.

For a bilingual dictionary to function as a practical linguistic instrument in a multilingual and multicultural society the information transfer has to be focused on pragmatic aspects. There is a fundamental difference between the demands of foreign language users and those of native speakers, and therefore lexicographers have to employ innovative approaches to make the information in their dictionaries accessible to the intended target user. The demands of the users of bilingual dictionaries in a multilingual society must be dealt with from a functional and pragmatic perspective which gives first priority to the usefulness of the treatment to the target user. A reassessment is needed of the traditional treatment of semantic information in bilingual dictionaries. The way in which illustrative examples, with a translation equivalent as address, can equip a user with semantic information, can exceed the transfer of semantic information achieved by traditional methods focusing on the lemma.

In a lemmatic addressing procedure the lemma is the topic and the information related to the address is the comment. Non-lemmatic addresses imply a constant topic switching because every new address is a new topic within the article, cf. Hausmann and Wiegand (1989: 328, 329). Although topic switching increases the density of information, the new comment is directed at a topic with which the user is unfamiliar, being a translation equivalent in the target language. The additional information adds communicative value and empowers the user with better comprehension.

There are various types of non-lemmatic addressing aimed at an enhancement of communicative equivalence in bilingual dictionaries. This includes the use of glosses, collocations, illustrative phrases and sentences, the explanation of meaning as well as ostensive addressing, i.e. the use of pictorial

illustrations. In equivalent relations of semantic divergence microstructural entries like these form a significant part of the dictionary article.

Hansen (1990: 3) distinguishes between the translation profile and the example profile in a bilingual dictionary. The translation profile contains the translation equivalents and all the information directly relevant to the translation equivalents. The example profile contains the illustrative examples and all the information directly relevant to the examples. According to Hansen the lemma and the translation profile constitute the compulsory components of the article of a bilingual dictionary while the example profile is in principle optional. For a sophisticated dictionary user who is familiar with both the languages treated in a bilingual dictionary, Hansen's argument might hold, because this user does not need the examples to contextualize the translation equivalent. However, in a multilingual society the lexicographer may neither presuppose this kind of expertise from the target user nor may he rely on the intuition of the user to interpret the translation equivalent correctly for a specific context. The use of examples represents a pragmatic non-lemmatic address and offers an interaction of semantics and grammar in order to equip the user with encoding and decoding skills. It can also help to differentiate between a stylistic adequate text presentation and the mere understanding of a foreign text, cf. Kühn (1989: 117, 118).

In RD the alternative outer access structure leads the user to the supplementary text in which the lexicographer has employed an innovative method of retranslating the equivalents in order to ensure a correct choice and a pragmatic and linguistically sound application. The main text of RD, which is also that of TW, treats semantic divergence unsatisfactorily. For many of these uncommented translation equivalent paradigms, RD has created new articles in which the translation equivalents are the topics of non-lemmatic addressing procedures. In the English-Afrikaans section the lemma **pinnacle** (noun) gets the following treatment in the translation equivalent paradigm:

torinkie, toringspits; spits torinkie; siertorinkie, pinakel (*bouk.*); top, spits, piek; hoogste punt, toppunt.

In the supplementary text the lemma **pinnacle** gets a lemmatic address with the inclusion of all the translation equivalents given in the main text. A non-lemmatic address provides each one of these translation equivalents with information on pronunciation as well as a retranslation:

torinkie turret; **toringspits** spire, steeple; **spits torinkie** pointed turret; **siertorinkie** and **pinakel** slender turret or spire on a roof; **top** top, peak, summit; **spits** peak, summit, spire; **piek** peak; **hoogste punt** acme, culmination, peak, zenith, highest point; **toppunt** top, summit, highest point, acme.

Semantic divergence implies either a polysemous lemma or the translation equivalent paradigm of a monosemous lemma which includes at least one polysemous translation equivalent. Although RD's retranslation does not provide contextual entries, glosses or explanations, the non-lemmatic addressing gives an indication of the different relations of equivalence between the lemma and the translation equivalents of the main text article. This procedure enhances communicative skills. A different kind of access structure can be identified in these articles. The retranslation gives access from the target language back to the source language. This is an access structure to promote the correct use of the translation equivalent in real communicative situations. It can be regarded as a communicative access structure or an exit structure from the article to the pragmatic functioning of the translation equivalent.

Conclusion

In a multilingual and multicultural society bilingual dictionaries have to provide translation equivalents. What is, however, of prime importance, is an additional treatment aimed at the differentiation of these equivalents. By skillfully adjusting the access structure and by an innovative application of addressing procedures the dictionary article can offer a presentation of communicative equivalents instead of merely semantic equivalents.

References

- Bosman, D.B. et al. 1984⁸. *Tweetalige woordeboek / Bilingual Dictionary*. Cape Town: Tafelberg Publishers.
- Branford, J. and W. Branford. 1991⁴. *Dictionary of South African English*. Cape Town: Oxford University Press.
- Dagut, M. 1981. Semantic Voids as a Problem in the Translation Process. *Poetics Today* 2(4): 61-71.
- Du Plessis, M. 1993. *Tweetalige Aanleerderswoordeboek / Bilingual Learner's Dictionary*. Cape Town: Tafelberg Publishers.
- Gouws, R.H. 1989. *Leksikografie*. Cape Town: Academica.
- Gouws, R.H. 1993. Afrikaans Learner's Dictionaries for a Multilingual South Africa. *Lexikos* 3: 29-48
- Gouws, R.H. 1995. Woordeboeke met Afrikaans en Duits as taalpaar: 'n metaleksikografiese situering binne die Afrikaanse vertalende leksikografie. *South African Journal of Linguistics*, Supplement 25: 1-87.
- Gouws, R.H. and F.A. Ponielis. 1992. The Development of Afrikaans and the Lexicographical Tradition. Zgusta, L. (Ed.). 1992. *History, Languages and Lexicographers*: 77-104. Tübingen: Max Niemeyer.
- Gouws, R., I. Feinauer and F. Ponielis. 1994. *Basiswoordeboek van Afrikaans*. Pretoria: J.L. van Schaik.
- Grobbelaar, P. et al. (Eds.). 1987. *Reader's Digest Afrikaans-Engelse Woordeboek / English-Afrikaans Dictionary*. Cape Town: The Reader's Digest Association, South Africa (Pty) Ltd.

- Hansen, G.B. 1990. *Artikelstruktur im zweisprachigen Wörterbuch*. Tübingen: Max Niemeyer.
- Hartmann, R.R.K. 1989. *Sociology of the Dictionary User: Hypotheses and Empirical Studies*. Hausmann, F.J. et al. (Eds.). 1989-1991: 102-111.
- Hartmann, R.R.K. 1989a. *The Dictionary as an Aid to Foreign-Language Teaching*. Hausmann, F.J. et al. (Eds.). 1989-1991: 181-189.
- 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.
- Hausmann, F.J. and R.O. Werner. 1991. *Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher: eine Übersicht*. Hausmann, F.J. et al. (Eds.). 1989-1991: 2729-2769.
- Hausmann, F.J. et al. (Eds.). 1989-1991. *Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography*. Berlin: Walter de Gruyter.
- Kotzé, E. and P. Wela. 1991. *Afrikaans / Zoeloe-woordeboek*. Cape Town: Tafelberg Publishers.
- Kromann, H.P. et al. 1991. *Principles of Bilingual Lexicography*. Hausmann, F.J. et al. (Eds.). 1989-1991: 2711-2728.
- Kühn, P. 1989. *Typologie der Wörterbücher nach Benutzungsmöglichkeiten*. Hausmann, F.J. et al. (Eds.). 1989-1991: 111-127.
- Lehrer, A. 1974. *Semantic Fields and Lexical Structures*. Amsterdam: North-Holland.
- Louw, J.P. (Ed.). 1985. *Lexicography and Translation*. Cape Town: Bible Society of South Africa.
- Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Odendal, F.F. 1994³. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Midrand: Perskor.
- Takebayashi, S. and Y. Kojima (Eds.). 1990. *Kenkyusha's Lighthouse English-Japanese Dictionary*. Tokyo: Kenkyusha.
- Wiegand, H.E. 1995. *Lexikographische Texte in einsprachigen Lernerwörterbüchern. Kritische Überlegungen anlässlich des Erscheinens von Langenscheidts Großwörterbuch Deutsch als Fremdsprache*. Popp, H. (Ed.). 1995. *Deutsch als Fremdsprache. An den Quellen eines Faches*: 463-499. München: Iudicium.
- Zöfgen, E. 1991. *Bilingual Learner's Dictionaries*. Hausmann, F.J. et al. (Eds.). 1989-1991: 2888-2903.
- Zöfgen, E. 1994. *Lernerwörterbücher in Theorie und Praxis*. Tübingen: Max Niemeyer.

A Translation Dictionary of Phrasal Verbs: An Ongoing Project

Mohamed Helmy Heliel, *Alexandria University, Alexandria, Egypt*

Abstract: The paper centres on a plan for an English-Arabic phrasal verb dictionary for Arab trainee translators. Such a dictionary answers a pressing need. It is not a bilingual dictionary but a translation one. This requires special lexicographical treatment as regards selection of headwords, structure of phrasal verbs, literal and figurative usage, definitions, synonyms, illustrative examples, relatedness of meaning, order of senses, collocations and discriminated equivalents.

Keywords: PHRASAL VERBS, TRANSLATION DIFFICULTIES, PHRASAL VERB TRANSLATION DICTIONARY, HEADWORDS, DERIVATIVES, FIGURATIVE SENSE, STRUCTURE, DEFINITIONS, SYNONYMS, MEANING RELATEDNESS, SENSES, COLLOCATIONS

Opsomming: 'n Vertalende woordeboek van frasewerkwoorde: 'n voortgesette projek. Hierdie artikel handel oor die beplanning van 'n Engels-Arabiese woordeboek van frasewerkwoorde vir Arabiese leerlingvertalers. So 'n woordeboek voorsien in 'n dringende behoefte. Dit is nie 'n tweektalige woordeboek nie, maar 'n vertalende woordeboek. Dit vereis spesiale leksikografiese hantering betreffende die seleksie van trefwoorde, struktuur van frasewerkwoorde, letterlike en figuurlike gebruik, definisies, sinonieme, voorbeeldmateriaal, betekenisverwantskap, volgorde van betekenisonderskeidings, kollokasies en onderskeidende ekwivalente.

Sleutelwoorde: FRASEWERKWOORDE, VERTALINGSPROBLEME, VERTALENDE WOORDEBOEK VAN FRASEWERKWOORDE, TREFWOORDE, AFLEIDINGS, FIGUURLIKE BETEKENIS, STRUKTUUR, DEFINISIES, SINONIEME, BETEKENISVERWANTSKAP, BETEKENISONDSKEIDINGS, KOLLOKASIES

1. Introduction

Under current teaching conditions of English as a foreign language, Egyptian students in the Department of English Language and Translation at Alexandria University have difficulties with English phrasal verbs. Though phrasal verbs are to some extent consistent in their behaviour and are now more teachable, as reflected in some excellent textbooks such as Britten and Dellar (1991), Murphy (1983) and Goodale (1993), our students who are native speakers of Arabic find them difficult to translate.

Phrasal verbs are by no means confined, as is claimed, solely to spoken English. Through our translation teaching experience of mass media texts, phrasal verbs proved to be of one of the major problems. To illustrate, we quote

the following ten sentences that the students were asked to translate from an English newspaper:

- (a) Further gains would *ward off* any quick recognition by the outside world of Aden regime.
- (b) Heavy rebel guns *opened up* 15 minutes after UN envoy left Kigali.
- (c) Britain yesterday *brushed off* positive signals from the IRA political leadership.
- (d) Mortar battles *picked up* despite a 36-hour cease-fire.
- (e) Residents had been *called up* to help regular troops retain control of provinces.
- (f) He was pressing for further quick military gains to *head off* foreign recognition of the *breakaway* state.
- (g) Poland to *speed up* privatisation.
- (h) The suicide bomber *blew himself up*.
- (i) The Cosmos 2281 satellite *blasted off* aboard a Cyclon-3 booster rocket on a military mission ...
- (j) The fall-off in demand was also evident at Banbury cattle market in central England.

It can safely be said that many of the phrasal verbs are perfectly acceptable in formal writing. As a matter of fact, they supply meanings which are not otherwise concisely expressed.

From translation teaching experience in the Department of English Language and Translation we realised that the learning of phrasal verbs at school is generally not satisfactory. The problems encountered in learning and translating phrasal verbs could be summed up in the following points:

- (a) The difficulty of translating word combinations whose meaning bears little or no relation to the individual words of which they are composed.
- (b) Confusion between *forms* of phrasal verbs and their *meanings*.
- (c) The intricate network of relations that a phrasal verb forms with other words whether collocationally or colligationally.
- (d) Difficulty of finding the appropriate Arabic equivalents to phrasal verbs due to lack of accurate understanding of their meaning.
- (e) The phrasal verb may not only be polysemic in having both an idiomatic and a non-idiomatic use, it may well be polysemic in having more than one idiomatic or figurative use e.g. (a) *put up* (for the night) (b) *put up* (a fight).
- (f) Phrasal verbs are by no means *static*, they move from colloquial to more formal contexts: mass media, radio news broadcasts, etc. and are always on the increase.

- (g) The phrasal verb is a feature typical of English and not of Arabic. In Arabic the verb may govern (a) the accusative of a noun, or (b) a preposition with the genitive of a noun, e.g.:

laḥīqa	bi-	'abīhi
caught up-	with-	his father

which takes the place of the accusative and gives greater precision and accuracy to the expression (Wright 1967: 44).

In Arabic, the verb retains most of its meaning and the preposition retains some degree of its literal or 'physical' meaning, although some combinations may be used figuratively. In English the range is far more complex: the meaning may be totally unpredictable and the verb and the particle fused.

- (h) The freedom with which derivatives form phrasal verbs appear and the semantic specialisation that goes with them (e.g., *falling off* in theatre attendances).
- (i) The unsatisfactory lexicographical treatment of phrasal verbs in English-Arabic general and specialized dictionaries (Heliel 1995).

Hence there is a pressing need for an English-Arabic dictionary of phrasal verbs which would serve as a *translatory* dictionary.

2. Dictionaries of English-Arabic phrasal verbs

The first attempt is Mazhar's (1949) *Dictionary of Sentences and Idioms* which contains a mixture of phrasal verbs and idiomatic expressions including verbs, nouns and adjectives as headwords. It is limited in scope. The target group is that of primary and secondary school levels of education. The second attempt is *A Dictionary of English Idioms* (1985) where some phrasal verbs are included together with their idiomatic expressions.

The only dictionaries devoted solely to phrasal verbs are:

- (a) Khalaili's (1979) *English Phrasal Verbs in Arabic* which covers 162 pages and treats only a limited number of 44 phrasal verbs. It is by no means "comprehensive or academically rigorous" (p.2). The targeted audience are intermediate learners.
- (b) Saadat and Naser's (1993) *Dictionary of Phrasal Verbs* is the only medium-sized specialized dictionary of English-Arabic phrasal verbs available. The target group is not specified. On the whole, the dictionary is far from being a help to the translator. In the main, it provides us with literal, non-discriminated decontextualised and thus meaningless Arabic equivalents. In many cases, the Arabic equivalents are rather odd para-

phrases or literal translations of the definitions given in an English monolingual dictionary. It is not clear on what basis it has incorporated some phrasal verbs and excluded others. The Arabic used in the dictionary is far from being accepted.

With all the difficulties encountered by the Arab translators and the present state of dictionaries of phrasal verbs in mind, we started planning for a new English-Arabic dictionary of phrasal verbs.

3. A translation dictionary

The proposed dictionary is by no means a bilingual one that confines itself to Arabic equivalents of English phrasal verbs. It is a dictionary which aims at helping the translator understand the phrasal verbs and stimulate him to find the Arabic equivalents that will fit in with the text he is translating. To achieve this goal the English and Arabic divisions should be carefully planned and worded and this is what distinguishes it from other English-Arabic dictionaries. The following are the lexicographical principles that will be adopted in the dictionary.

4. Headwords

Frequency of use should be the overriding criterion for selection. In this respect, it has to be noted that:

- (a) *The Collins COBUILD Dictionary* (1989) lists over 3000 combinations of phrasal verbs explaining over five and a half thousand different meanings. It claims that these combinations are the ones in common use in everyday modern English. What lends support to this statement is the fact that the dictionary is based on a detailed computational study of the Birmingham collection of English texts. In the absence of reliable statistical studies the *Collins COBUILD* will be used as a guide.
- (b) *The Particles Index* in *Collins COBUILD* (1989) is an extensive guide to the way in which particles are used in English phrasal verbs. It also gives the actual number of occurrences of each particle which will be of interest in the selection of the headwords for our dictionary. It will help us determine which phrasal verbs to choose and focus on. According to *Collins COBUILD* (1989: 449-491) the commonest particles are *up*, *out*, *off*, *in*, *on*, and *down*, in descending order of frequency. *Up* (482) and *out* (410) are extremely common comprising (28%) of the phrasal verbs listed in the *COBUILD* dictionary. They are followed by *off*,

in, on, down, away, into, over, around as is clear from the COBUILD computer data study:

<i>aback</i>	001	<i>forth</i>	011
<i>about</i>	088	<i>forward</i>	010
<i>above</i>	005	<i>from</i>	123
<i>across</i>	006	<i>in</i>	209
<i>after</i>	015	<i>into</i>	112
<i>against</i>	030	<i>off</i>	233
<i>ahead</i>	014	<i>on</i>	198
<i>along</i>	023	<i>onto</i>	012
<i>among</i>	004	<i>out</i>	410
<i>apart</i>	010	<i>over</i>	111
<i>around</i>	104	<i>round</i>	058
<i>as</i>	008	<i>through</i>	059
<i>aside</i>	015	<i>to</i>	064
<i>at</i>	045	<i>together</i>	022
<i>away</i>	134	<i>towards</i>	011
<i>back</i>	083	<i>under</i>	009
<i>before</i>	005	<i>up</i>	482
<i>behind</i>	010	<i>upon</i>	083
<i>by</i>	022	<i>with</i>	056
<i>down</i>	191		
<i>for</i>	063		

- (c) Emphasis should also be placed on certain verbs (i.e. *be, come, get, look, put, take, turn, work*) which seem "to combine more frequently with particles than other verbs. The more particles with which a verb combined, the wider the range of meanings for each form." (Ryan 1984:20).
- (d) Usually, these function words (*out, off, on, ...etc.*) do not have consistent meanings in the phrasal verbs of which they are a part. There are, however, a few 'families', so to speak, of phrasal verbs in which the function words carry a fairly consistent meaning. These families require special lexicographical treatment in formulating their definitions, indicating the concepts to which they refer and hence helping the translator find the appropriate way of expressing them in Arabic. The proposed dictionary should benefit from:
- (a) The Particles Index in Collins COBUILD (1989: 450-491)
- (b) Harrap's English Verbs (1989: 83-89)

4.1 Derived forms

Any noun, adjective or modifier formed from a phrasal verb should be entered under the phrasal verb headword. British English, it has to be noted, absorbs many nominalised forms:

- (a) nouns indicating activities: e.g. *blast-off, sit-in*
- (b) gerunds: *beating up, dressing down*
- (c) past participles used as nouns, e.g. *leftover, dugout*

Some of these forms may enter a compound, e.g. *breakdown truck, breakaway movement*. The freedom with which these noun forms appear and the semantic load they carry justifies incorporating them. Such forms, though quite common in newspaper English, for example, are missing from our English-Arabic dictionaries. Mention should be made of the fact that the derived nouns from phrasal verbs do not have all the senses of the phrasal verb. Only nouns derived with the same sense of the verb should be included together with illustrative examples and Arabic equivalents. Some of these nouns, however, may have a particular sense which is not there in the verb, e.g.:

- | | |
|---------------------|--|
| <i>blow out</i> (v) | <ol style="list-style-type: none"> 1. (a flame or a fire) it stops burning 2. (a tyre) burst 3. (gas) escaping |
| <i>blow out</i> (n) | <ol style="list-style-type: none"> 1. tyre 2. escape of gas 3. <i>a big expensive meal or large social occasion</i> |

Such a sense (no. 3) will not be included.

5. Structure

Stating that the structure of a phrasal verb is either *v + prep* or *v + adv*, as is common in some dictionaries (Saadat and Naser 1993), does not help the translator. However, since meaning varies according to the different types of structures in which the phrasal verb occurs, it should be clearly stated whether the phrasal verb is *transitive* or *intransitive*. In the first case both the subject and the object collocate should be indicated and in the second only the subject collocate. The translator should correlate difference in form with difference in meaning.

6. Meaning

The literal meaning of phrasal verbs should be given to relate it to the idiomatic or figurative one. The proposed dictionary will comprise:

- (a) Idiomatic phrasal verbs, i.e., idiomatic combinations where the meanings of the separate parts tell us little or nothing about the meaning of the whole.
- (b) If the verbal combination has a normal or literal meaning as well as an idiomatic or figurative one, then both are listed. The compilers of *Collins Dictionary of Phrasal Verbs and their Idioms* (1974) argue that "differences of meaning occur within a continuum of meaning beginning with the *simple verb + particle* and ranging through to the more opaque idioms" (p.6).

The following types of phrasal verbs are not included:

- (a) Idioms which are connected with phrasal verbs, e.g. *to pull in one's horns*. These should be entered in a dictionary of idioms.
- (b) Phrasal verbs such as *depend on*, *depart from* where the verb, though it always has the same meaning, can only be used with a certain preposition when used in a combination. In such cases the information value of the preposition is either reduced or nil. It is no more than a dummy element (Rastall, 1994). These non-idiomatic or literal prepositional verbs, which simply combine the meanings of their parts, are excluded.

However, there are many phrasal verbs which are neither fully idiomatic nor fully literal. The guide for selection in these cases will be *comprehensibility*.

From experience we realised that trainee translators have difficulties with:

- (a) Verbs of movement with prepositions of direction, e.g.
 - (i) The crowd *pressed in* through the cordon of guards.
 - (ii) I *need her over* to help me with the children.

Jowett's (1951) words about the English language ring quite true in the translation classroom:

"It is indeed a language in which we call upon simple verbs concerned with acts, motions and attitudes of the body to express a vast range of human activities. We tend to think in terms of muscular and tactile sensation, to try to express sensuous reality, and so to employ kinaesthetic rather than visual imagery. Thus we *go for* an opponent when we attack

him, we feel *run down* when we are nervously exhausted, we *fall out* when we quarrel ..." (Jowett 1951: 155).

- (b) Phrasal verbs which are used figuratively. Such phrasal verbs, though by far the most difficult to translate (Heliel 1995), are missing from both general and specialized English-Arabic dictionaries of phrasal verbs, e.g.:
- (i) *To patch things up* is literal when it refers to mending with patches with the Arabic equivalent *yartuqu*.
 - (ii) *To patch up* is figurative when it refers to settling differences with the Arabic equivalent *yusawwii*.

Both kinds of verbs will receive particular attention.

7. Definition

In a translation-oriented dictionary like ours, *accurate* and *detailed* definitions in English are badly needed and are important for the understanding of the phrasal verb in its language context and hence for translating it. An all-embracing definition will provide the translator with an important clue to the meaning which is reinforced by synonyms and illustrative examples. Unsatisfactory definitions together with deficient illustrative examples detract from the value of the English section in the dictionary and we will be left with the inadequate decontextualised Arabic equivalents which are of no use to the translator. Compare (a) and (b), (c) and (d):

- (a) *talk away*: force to go away by talking. I'm afraid you won't talk this problem away. He can't talk away his enormous debts.
- (b) *talk away*: to get rid of a feeling with words. Stay with the child as he goes into hospital, you may be able to talk *his fears* away.
- (c) *live down*: survive, manage to forget. You'll never live it down. It will be difficult to live down this humiliation.
- (d) *live down*: If you are unable to live down a *mistake, failure, or foolish action*, you are unable to make people forget that you did it. The story of a *young girl* who finds it impossible to live down *her past*.

The italicized words in (b) and (d) above are important details for conveying the meaning to the translator. Short renderings of phrasal verbs (example (c) above) may lead to wrong Arabic equivalents. The definition of a phrasal verb especially a *figurative* one should be carefully worded as that of the one-word entry in a monolingual dictionary, not to be confined to the label *figurative*. The following example (b), for lack of definition, proved to be difficult to understand:

siphon off remove or transfer by means of a siphon.

- (a) (lit.) Someone has siphoned off my petrol.
- (b) (fig.) The department is good at siphoning off funds from one area to another. They have siphoned the money off for their own use.

7.1 Synonyms

In our proposed dictionary we will make use of all that would help the translator get a better understanding of the phrasal verb. Though synonymy is a relative concept (Cornell 1985: 274) and synonyms differ in collocations, adding them to the definition will be of some help to the translator., e.g.

die away —————> definition + (same as *fade*).

Among dictionaries that proved quite useful as sources for synonyms of phrasal verbs are: The Reader's Digest Oxford Wordfinder (1993), The Merriam-Webster Concise School and Office Thesaurus (1991) and The Concise Oxford Thesaurus (1995).

8. Illustrative examples

Just like synonyms, illustrative examples would help reinforce the meaning by showing how the phrasal verb is actually used in an appropriate context with the words that normally collocate with it. It is important that illustrative examples should justify their existence. Compare (a) and (b), (c) and (d):

- (a) *walk out on*. Don't walk out on me. You can't just walk out on us like this.
- (b) It's the first time my husband's walked out on me.
- (c) *wink at*. Can't you wink at what they are doing.
- (d) Many of the teachers wink at the boys' secret smoking.

In (a) and (c) there are no clues to the nature of the collocate which to a large extent determines the Arabic equivalent. In (b) and (d) the *wife* (implicit object), *teachers* and *smoking* (subject and object) do determine the Arabic equivalents to the phrasal verb.

The decision whether to stick to short phrases and minimal contexts or to use full-length sentences for illustration was difficult to make. It was found out that both approaches are required but for different entries:

- (a) Example of a phrasal verb with a *sentence* as an example:
put through. I rang the bank and asked to be put through to the overseas department.
- (b) Example of a phrasal verb with a minimal context, just a *collocate*:
put forward (idea, suggestion, plan).

In (a) it is thought advisable to use a sentence to reinforce the definition, and convey the idea of connecting two persons on the telephone. In (b) what is put forward, i.e. the collocate, is what is needed.

9. Relatedness of meaning and order of senses

In his attempt to understand the phrasal verb, the translator makes use of the definition, the synonyms and the example sentences provided. Thus the English entry should be a closely-knit unit. Figurative meaning should follow the literal one and the order of senses should be determined by logical relatedness not frequency of occurrence. Compare the following entries chosen from two different dictionaries:

- (1) *put aside* (a) see *put* (b) (fig. abandon) He put aside his anger. (c) (*save*) She has managed to put aside a little money.
- (2) You *put something aside* a. when you separate it from other things for some purpose. *She unpacked her baggage and put aside things to be taken upstairs* b. When you save it for future. *I'd been putting money aside for my holiday.* [same as *save*] c. When you abandon it. *Publishers should put aside their prejudices* [same as *discard*].

10. Arabic equivalents

The meaning of a lexical item which is given in a dictionary entry must be regarded as Jackson (1988) rightly says; "potential", a distillation of the essentials awaiting actualisation in a particular linguistic and situational context". Thus the one-word equivalent in Arabic may not be helpful or might not be the appropriate translation. The translator should be provided with a full range of possibilities in the English section (definition, explanation, synonyms, example sentences, collocates) and a *number* of Arabic synonymous equivalents whether words or phrases from which the translator will be free to choose, e.g.

waste away to become weak and ill gradually. [same as *grow weak, wither*].
Since my aunt's disease, she's been wasting away

[صحتها] تَذْوِي ، تَذْبُل ، تَسْوِء ، تَتَدَهْوَر

[sihhatuhaa] tadhwa, tadhbulu, tasuu'u, tatadahwaru.

It also happens that the translator may select a word in the mother tongue which is not amongst the equivalents given but it fits in perfectly with the text he is translating. In this case the English entry and the Arabic synonymous equivalents with the collocates shown between square brackets will serve as *stimulants* for the right translation.

One of the major weaknesses of English-Arabic bilingual dictionaries is their non-discriminated Arabic equivalents or mere Arabic paraphrases, e.g.

turn aside: حَاد ، عدل ، جَنَح

[ḥaada, 'adala, dʒanaḥa]

The Arabic decontextualized verbs are of no help to the translator. They even lack the prepositions which accompany them and specify their meanings in Arabic.

11. Collocations

Phrasal verbs may be regarded as not collocations but a verb sub-type that occurs within a collocation. There are frequently strong collocational associations between phrasal verbs and other words that determine the Arabic equivalent (Heliel, 1995). Whether in the English or Arabic section collocations should be indicated, e.g. *fight back* (tears/ emotions/ doubts/ fears/ desires), where for each collocation there is a particular Arabic verb equivalent.

Whether short phrases or full-length sentences are used in the English entry, what is extremely important is focusing on the collocate (subject, object or both), not translating the example sentence, e.g.,

tone down (radio / intensity of colours)

يُخَفِّضُ صَوْتِ [الرَّادِيُو] ، يُخَفِّفُ مِنْ [حَدَّةِ اللَّوْنِ]

yukhaffifu, ṣawta [al-rāadyoo] yukhaffifu min [al-lawn]

12. Conclusions

An English-Arabic dictionary of phrasal verbs is an indispensable tool for the Arab translator. In such a translation-oriented dictionary, certain features should be brought into focus:

(i) On the English side

(a) The morphological and syntactic aspects of the phrasal verb.

- (b) Selection which should be based on the frequency of use in the language.
 - (c) Accurate and detailed definitions including synonymous equivalents and illustrative examples.
 - (d) The network of relations the phrasal verb forms with other words (collocations).
 - (e) Order of senses and logical relatedness of meaning.
- (ii) On the Arabic side:
- (a) Arabic equivalents with synonyms to help the translator choose from amongst them, the one that fits in well with the text he is translating.
 - (b) The collocates that go with the phrasal verb and determine its meaning.

References

- Britten, Donald and Gwendolyn Dellar. 1991'. *Using Phrasal Verbs*. London: Prentice Hall.
- Cornell, Alan. 1985. Realistic Goals in Teaching and Learning Phrasal Verbs. *IRAL* 23 (4): 269-280.
- A Dictionary of English Idioms: English-Arabic*. 1985. (Compiled by a team of university teachers of English.) Beirut: Librairie du Liban.
- Goldie, Jane. 1989. *Harrap's English Verbs*. London: Harrap.
- Goodale, Malcolm. 1993. *Collins COBUILD Phrasal Verbs Workbook*. London: Harper Collins.
- Heliel, M.H. 1995. Towards a Dictionary of English-Arabic Phrasal Verbs for the Arab Translator. A paper presented at the 2nd International Maastricht-Lodz Duo Colloquium on "Translation and Meaning", Lodz, 22-24 September 1995.
- Jackson, Howard. 1988. *Words and their Meaning* London: Longman.
- Jowett, W.P. 1951. On Phrasal Verbs. *ELT* 5(6) : 152-157.
- Khalaili, K. 1979. *English Phrasal Verbs in Arabic*. London: Hodder and Stoughton.
- Kirkpatrick, Betty. 1995. *The Concise Oxford Thesaurus*. Oxford: Oxford University Press.
- Mazhar, I. 1949. *A Dictionary of Sentences and Idioms: English-Arabic*. Cairo: The Renaissance Bookshop.
- McArthur, Tom and Beryl Atkins. 1974. *Dictionary of English Phrasal Verbs and their Idioms*. London: Collins.
- The Merriam-Webster Concise School and Office Thesaurus*. 1991. Springfield: Merriam-Webster.
- Murphy, M.J. 1983. *Understanding Phrasal Verbs*. London: Hulton.
- Rastall, P. 1994. The Prepositional Flux. *IRAL* 32 (3) : 229-231.
- Ryan, S.M. 1984. The Frequency of Phrasal Verbs in Spoken American English. Los Angeles: University of California, MA thesis.
- Saadat, A.E and K.A. Naser. 1993. *Dictionary of Phrasal Verbs: English-Arabic*. Riyadh: Mars.
- Sinclair, J. et al. 1989. *Collins COBUILD Dictionary of Phrasal Verbs*. London: Collins.
- Tillock, Sarah. (Ed.). 1993. *The Reader's Digest Oxford Wordfinder*. Oxford: Clarendon Press.
- Wright, W. 1967. *A Grammar of the Arabic Language*. Cambridge: Cambridge University Press.

Dictionaries and Orthography in Modern Africa

M. Lynne Murphy, *Department of Linguistics, University of the
Witwatersrand, Johannesburg, South Africa*

Abstract: This paper considers the issues and options in creating orthographies for languages without alphabetic traditions. In particular, it looks at issues that face linguists and anthropologists in creating such orthographies for preliminary descriptive work on a language. It is often the case that these orthographies by linguistic outsiders later serve as the bases for literacy-aimed orthographies. Thus, the lexicographer is constrained not only to present the language accurately, but to not undermine indigenous literacy projects for that language. This paper looks at the particular case of devising an orthography for Beng, a Southern Mande language of Côte d'Ivoire, in order to illustrate the problems inherent in alphabetising an unfamiliar language.

Keywords: AFRICAN LANGUAGES, ALPHABETISATION, BENG, CÔTE D'IVOIRE, DICTIONARY, FIELDWORK, INTERNATIONAL PHONETIC ALPHABET, LEXICOGRAPHY, LITERACY, MANDE, ORTHOGRAPHY

Opsomming: Woordboeke en ortografie in Moderne Afrika. In hierdie artikel word die kwessies en opsies in die skepping van ortografieë vir tale sonder alfabetiese tradisies oorweeg. Daar word veral gekyk na kwessies waarmee linguïste en antropoloë te doen kry by die skepping van sulke ortografieë vir voorlopige deskriptiewe werk aan 'n taal. Dit gebeur dikwels dat hierdie ortografieë deur linguïstiese buitestaanders as basis dien vir ortografieë wat op geleterdheid gerig is. Gevolglik word die leksikograaf gedwing om nie net die taal akkuraat weer te gee nie, maar ook om nie inheemse geleterdheidsprojekte vir daardie taal te ondermyn nie. Hierdie artikel ondersoek die spesifieke geval van die ontwerp van 'n ortografie vir Beng, 'n Suidelike Mande-taal van die Ivoorkus, om probleme inherent aan die alfabetisering van 'n onbekende taal te illustreer.

Sleutelwoorde: AFRIKATALE, ALFABETISERING, BENG, IVOORKUS, WOORDEBOEK, VELDWERK, INTERNASIONALE FONETIESE ALFABET, LEKSIKOGRAFIE, GELETERDHEID, MANDE, ORTOGRAFIE

1. Introduction

This paper considers contemporary methods for assigning alphabets and orthographies to languages without an alphabetic tradition. Issues relevant to orthographic assignment are surveyed with particular attention to issues for African languages. The ultimate goal of the paper is to assess the possible ramifications of such assignments for lexicographers in the modern African context.

These ramifications are discussed with reference to the orthography for Beng (a minority language of Côte d'Ivoire) created by American scholars for a Beng-English ethnographic dictionary.

Throughout history, conquering peoples have set the languages of their victims to print in writing systems mimicking the conqueror's own. This continues to be true in the present day, when former colonising and imperialist cultures exert orthographic influence (by means of missionaries and academics) on the same peoples they once conquered. European (including Euro-American) missionaries have frequently pursued alphabetisation of the local languages of former colonies, but academics, especially linguists and anthropologists, and governments have also played rôles in developing orthographies. In the present day, national governments and international organisations (such as UNICEF) are acting upon an increased interest in promoting literacy in developing nations. This interest presses for further development of orthographies for languages with no writing customs, and the application of existing European alphabets to these languages continues for both pragmatic and political reasons.

For outsiders involved in the development of orthographies and literacy, several ethical dilemmas present themselves. By introducing writing to a non-writing culture, members of a writing culture seem to make an implicit judgement: that the non-writing culture is inadequate and must be repaired through the introduction of writing. Thus, the introduction of writing to a culture may simply be a means of proselytising European (or Arabic or Chinese) values and culture. However, the introduction of writing by former colonial forces may instead be viewed as partial restitution for the injustices of colonialism: a) by allowing the victims of colonialism to share in the economic system that has benefited the colonial powers, and b) by helping to preserve languages whose livelihoods have been threatened by the intrusion of a colonialist language and the economic and political changes brought forth by colonialism.

Promoting literacy and non-traditional (i.e., Western-style) education produces more ethical and practical problems. One can make many good arguments for promoting literacy in local language. Vernacular literacy allows more people to attain higher levels of education and may preserve languages that would otherwise die out due to the predominance of official (often European) and national languages in the educational and economic system.¹ But literacy can also be detrimental to traditional village life. As young people become literate, especially as they become literate in national languages, they often become dissatisfied with traditional life in agricultural villages and leave to find work in cities.² Whether the cities of developing nations can support an influx of new literate people is definitely questionable.

As part of a literacy program, orthographic development should involve native speakers of the language in order to preserve the language users' rights of self-determination in linguistic and educational matters, as well as for practical reasons. (Literacy efforts are facilitated if the indigenous population sup-

ports the chosen writing system.) However, often some record of the language is made for the benefit of outsiders (e.g. academics) prior to mass literacy efforts. This paper aims to discern how academic orthographies affect future attempts at literacy-aimed orthographies and whether these effects may be harmful to attempts at literacy.

This paper consists of six sections. The remainder of this section explains the terms and typographical conventions used. Section 2 briefly raises questions concerning the nature of writing systems. The third section concerns orthographic design as it applies to previously unwritten languages, especially those of Africa, discussing first the selection of graph style, then orthographic choice. Section 4 outlines other orthographic issues: word division and capitalisation conventions. Section 5 concerns the treatment of the Beng language of Côte d'Ivoire in an ethnographic dictionary published in the United States. The final section discusses the rôle of dictionaries in establishing a practical orthography.

1.1 Terminology and conventions

In this paper, the term GRAPH refers to a minimal unit of writing which by itself represents some part of a language. For example, the 26 letters in English orthography are graphs, which may represent sounds. In the Chinese writing system, graphs semantically represent morphemes. ALPHABET refers to a set of graphs that represent sounds, though not necessarily the same sounds in every language that it serves. A SYLLABARY is a set of graphs that represent syllables, usually consonant-vowel combinations. DIACRITICS are symbols used in conjunction with alphabets or syllabaries in order to make more precise sound differentiations than the graphs alone. They differ from graphs in that they cannot stand alone in writing — they must affix to some graph. Accent marks are well-known examples of diacritics. An ORTHOGRAPHY is the writing system of a language, consisting of a set of graphs and conventions regarding the relation of those graphs to sounds or words, the division of words, capitalisation, *et cetera*. Although languages may share almost identical alphabets, their orthographies may differ. For example, according to the orthography of English, pan indicates the pronunciation [p^hæn], while in French orthography, pan ('bread') indicates the pronunciation [pā]. The term ALPHABETISATION refers to the process by which a language is assigned an alphabet and orthography. This should not be confused with the French term *alphabetisation* (used often in the materials cited here), which refers to the process of literacy education. Nor should it be confused with the term's use elsewhere to mean 'alphabetical ordering'.

In keeping with linguistic tradition, phonetic representations appear in brackets and linguistic examples are underscored. Exemplary graphs are represented in boldface.

2. The nature of writing

While the languages of most former European colonies (with the exception of some Asian colonies) have been considered by Westerners to be unwritten, some of these languages do have historical writing systems that have been ignored or deemed unsatisfactory by the colonial and post-colonial powers-that-be. Defining writing system is in itself a problem. The writing systems of Europe have almost one-to-one correspondence between spoken and printed words. Yet other systems, such as Chinese, do not practise such phonetic / orthographic consistency. In Chinese, the character represents a concept that is realised with different sounds in the different Chinese languages. Certainly, other cultures have means of representing ideas in concrete symbols. These symbols might be incorporated into the design of objects, in the patterns on clothing, in house decoration, or in paintings and sculpture. They might be used to record historic events, to revere heroes or deities, or to record debts or crop information. Are these writing systems? If we were to view such symbolic systems as writing, would our views toward the introduction of European-style writing systems have to change? In a perfect world, perhaps they would. But in a world made smaller through imperialism and subsequent world-wide economic dependencies, it may be too idealistic to suggest that orthographers should accept local symbolic systems as writing systems comparable to those of Europe and Asia. The writing systems discussed in this paper conform to Western standards; however, the reader should keep in mind the possible inadequacy of this standard.

3. Common goals of orthographic design

In designing or choosing an alphabet and developing an orthography for a language, many practical and political issues must be considered. However, these practical and political concerns may point to conflicting orthographic choices. In this section, I consider first the choice of graph style (roman, indigenous, original, etc.) and then consider the development of orthographies using the selected graph style, focusing on the roman graphs which are most often chosen for African languages.

3.1 Graph style selection

Most orthographies developed in Africa and other former European colonies have been based on roman graphs. While this choice of graphs is rarely questioned or challenged in orthographic literature, this selection raises many practical and ethical questions. Export of these western European graphs to colonies around the world smacks of cultural and linguistic imperialism. In fact, many

Christian missionaries have seen roman alphabetisation of local languages as the means by which to introduce their religion to the peoples of Africa and America. Richard Lepsius, in the advertisement for his own international orthography, insisted that, "in the case of Africa, [use of a standard system based on roman graphs is necessary, for] in this way only can we hope for Evangelization of that vast continent" (1863: v). This sentiment is echoed today in the actions of the Summer Institute of Linguistics, a Protestant missionary organisation that develops literacy in vernacular languages in order to translate and distribute the Bible and other Christian texts.

3.1.1 Indigenous graphs

In the European fervour to "evangelise" or "civilise" colonised peoples, European orthographies have sometimes supplanted indigenous orthographies. In western Africa, indigenous orthographies were developed for several languages prior to or separate from European influence. At the turn of this century in Cameroon, King Njoya of the Bamum developed an ideographic script of 466 characters which was used in writing the history of the Bamum and for royal correspondence. In Sierre Leone and Liberia, syllabaries for five Mande languages were developed by native speakers between 1830 and 1940. Each has been used for correspondence and record-keeping, and one (Vai) has been standardised and used in translations of the Bible and Koran. These systems were developed without overt influence from non-African cultures, although some similarities have been found between the Mande syllabaries and a secret writing system used to transcribe Arabic by the Hodh of Mauritania (Dalby 1986). This similarity of writing systems may indicate that a common writing system was used by the medieval Mandin empire, which spread the linguistic ancestor of the Mande languages (Dalby 1986). Thus, these cultures, often described as having unwritten languages, may have centuries-old traditions of writing.

For languages with extant or historical writing systems, the introduction of a new orthography would seem redundant. Furthermore, replacement of indigenous orthographies by imported systems is blatant cultural imperialism. But despite the fact that syllabaries existed for several African languages when missionaries contacted their speakers, the westerners considered it necessary to introduce roman writing systems. This was not entirely due to western ignorance of the indigenous systems. In his advertisement for his standard alphabet, Lepsius (1863) mentions the Vai syllabary as evidence that African people are adept at teaching each other writing systems, ignoring the point that, for the Vai, his alphabet is redundant.

Dalby (1986 — one of the few orthographic histories that discusses native West African scripts) gives no indication that any of these indigenous writing systems flourish. Because mass literacy was not established with the invention

of the Bamum or Mande systems, they may be destined for extinction. Few people already know the indigenous systems, but many indigenous people know the writing system of the language used in government or church-run schools, often the official language or a regional *lingua franca*. Widespread use of the official writing system hinders the survival of the indigenous system because of the time, expense, and commitment necessary for teaching and maintaining two systems. So, while vernacular literacy programs could preserve and even promote indigenous writing systems and related cultural traditions, these efforts instead ensure that the indigenous writing systems are abandoned.

3.1.2 Entrenchment of roman graphs

The persistence of European official languages entrenches roman graphs in former European colonies. As well as preventing the revival of extant indigenous systems, this prevents the introduction of original graph systems from within the language community. For members of the language community to create their own orthography from scratch would be the utmost in self-determination in the promotion of literacy. However, such an option is barely considered because of the practical matters of printing technology and managing language policy in multilingual states.

In African nations with a European official language,³ the use of roman graphs for indigenous languages allows for the use of the same printing equipment to produce materials in more than one language. Thus, the reasons for the choice of roman-based systems are usually economic in nature: a) printing equipment for roman graphs is readily available, and b) use of the same graphs for local and official languages avoids the economic and mental burdens of maintaining separate equipment and educational efforts.

In this age of printed communication, a primary goal of orthographic design is that the graphs be readily convertible to type. In the First World, where computer-based printing technology is very accessible, invention of new graphs poses few problems for the publishing industry and increasingly fewer problems for those who have abandoned typewriters for word processors. However, in developing nations, such technologies (and the skills for using them) are rarely available or affordable. Because new writing systems are often developed as parts of national literacy campaigns, they must be printable immediately, for use in primers and later in other educational, political, medical, and agricultural materials. For these reasons, officials often seek to utilise extant alphabets for the new orthographies, so that existing equipment may be used in printing.

However, in practice, application of the roman alphabet to non-Indo-European languages is often difficult and necessitates added expense. Because languages vary in their inventories of consonants and vowels, assigning

African significances to roman graphs is not a simple matter. Even if the sounds of the language number fewer than 26 (the number of roman graphs used in English), a basic principle of orthographic practice bars any assignment of graphs to sounds that fundamentally differs from their assignment in other languages (for example, assigning the vowel sound [ɔ] to the letter *b* is *verboten*).⁴ For this reason, non-roman graphs are often introduced to a basically roman system. Often, these graphs are modified roman characters, such as *ŋ* or *ε*, or are borrowed from other alphabets, such as *θ* and *ð*. Once non-roman characters are introduced, the adoption of a roman-style system is no bargain. New typewriters and printing equipment must be designed to service the language.

Although new printing equipment must be manufactured, economic efficiency and political expedience have led many African nations to adopt official alphabets that include roman and non-roman graphs. While the modified roman alphabets involve an initial expense, cost-efficiency can be attained by applying the same graphs to all indigenous languages. For example, the Nigerian National Language Centre has consolidated the writing systems of over 250 indigenous languages and introduced a pan-Nigerian typewriter which contains all of the letters of the roman alphabet except *x* as well as nine modified roman characters and a number of diacritics (Dalby 1986: 28). While adoption of any writing system involves expense, this tactic dilutes the expense by applying the same printing equipment to many languages and thus creating a large market for the new equipment, allowing it to be mass-produced.

3.1.3 Conclusion

Modern, state-supported orthographies for African languages are most commonly based on a roman-style graph system. This holds true despite the existence in some areas of indigenous writing systems. Harmony among the writing systems of various indigenous languages and (official) European languages is usually encouraged. Proponents of this method cite economic reasons for the choice of roman graphs as well as pedagogical reasons (discussed further below). While the involvement of roman script in non-Indo-European languages is problematic, both from orthographic and political / ethical standpoints, it is now so deeply entrenched in African political and educational systems that alternatives, such as invention of new alphabets, no longer seem viable.⁵

3.2 Phoneme / graph assignment

Once the roman style of graphs has been selected for a particular language, planners must determine how roman graphs will represent the sounds of the language in question. Two methods are most popular: (1) mimicry of the ortho-

graphic system of another language, and (2) adoption of a standardised orthography which has been created in order to represent the sounds of many languages. The borrowed orthography or standard orthography is usually used as a model for a new orthography, rather than adopted without changes. Thus, the systems are tailored to fit the sounds of the languages they are intended to serve. Orthographies that adequately represent the sounds of the language are often adapted through use so that the characters are easier to produce or to read or so that less phonetic detail is represented. Standard orthographies have come to surpass the orthographies of European languages in popularity, but both are discussed below.

Orthographic theorists generally agree on several criteria for an ideal orthography (see Kemp 1981: 16*, IPA 1957: 1-2, Mann and Dalby 1986: 207, Winter 1983). Foremost is the "One Sound / One Sign" principle⁶ which holds that for any phoneme (minimally distinctive sound unit), one and only one graph should consistently represent that phoneme and no other. This principle is found in every list of orthographic principles in the literature cited here and is usually the first principle mentioned, as well as the most discussed. Notably, this principle is not followed by most languages with a long history of writing, such as English or French, because natural, diachronic sound change has little regard for spelling. However, One Sound / One Sign is the goal of most current orthographies and is claimed to benefit both the language scholar, who demands consistency and precision in phonetic description, and the native speaker, who is attempting literacy (the more simple and elegant the spelling system, it is reasoned, the easier the acquisition of reading skills). The adherence to One Sound / One Sign is further discussed below with respect to various orthographic systems.

3.2.1 Borrowed orthographies

In the earliest language-recording efforts by European colonists and missionaries, the orthography of the European's native language (or some other European language) was most often used for transcription. This tendency can be attributed to the lack of international standard orthographies and to simple ethnocentrism. These early orthographies persist in some areas where the original colonial forces have remained and where continual efforts have been made to record or promote literacy in the local language. (Note, for example, the French spelling of [u] as *ou* in *Ouagadougou*.) Such orthographic borrowing may be instigated or perpetuated today by those who believe that borrowing holds some pedagogical value.

For many individuals, literacy in a local language is seen as a stepping stone to mastery of a European language (usually the official language of the nation) and thus to greater economic mobility. Modelling the local language's orthography after that of a European language is claimed to aid in the develop-

ment of bilingual literacy. However, such orthographic mimicry in the name of educational principles may be misguided. As Yapita Moya (1981) points out for the Bolivian language Aymara, the adoption of Spanish orthography confuses the Aymaran learner of Spanish as well as the Spanish-speaker learning Aymara, as the sound systems of the two languages are different enough to cause problems for orthographic sharing. Aymara has three vowels, while Spanish has five. Yet in three historical orthographies of Aymara modelled on Spanish orthography, all five of the Spanish vowel graphs are utilised, violating the One Sound / One Sign principle, causing confusion to native speakers who cannot discern a logic to the assignment of vowel graphs to sounds, and causing many learners of Aymara to believe that five vowels actually exist in Aymara (and that they are stupid for not being able to distinguish them).

Yapita Moya argues that the use of Spanish orthography for Aymara promotes a feeling of intellectual inferiority among the Aymara people. Because their language is so easily read by Spanish speakers (whether or not the Spanish speaker can understand the content of the text), Aymarans assume that their language is "simple" compared to the Spanish language, with which they often have difficulty. Here, we can find not only abstract philosophical reasons to oppose the culturally imperialistic imposition of European orthographies on non-European languages, but also evidence that such imperialism is damaging to educational efforts, to the emotional well-being of the individuals of a culture, and to their cultural identity.

3.2.2 Standard / international orthographies

Standard roman orthographies for non-European languages gained popularity among European scholars during the 19th century. The aim of these systems is to serve more than one dialect, if not all languages, by applying graphs to sounds universally. The One Sound / One Sign principle is intended to hold not only within one language, but across many. If a sign *b* represents a bi-labial, voiced stop in language A, then it serves the same purpose in language B. If there is no bi-labial, voiced stop in language C, then the sign *b* should not be evident in the writing of that language. The efficiency of this system is affected by the differences and similarities among the languages that are served, the consistency of the application of graphs to sounds, and the number and range of graphs and diacritics included in the system.

The colonial history of Africa has resulted in the creation of many competing orthographies. Orthographies of individual African languages were modelled on many different European orthographies because different European nations had colonies in Africa, and because the colonies of any particular European nation were not necessarily contiguous and did not honour language or language-group boundaries. Thus, the name of one language family might be spelled three different ways in three different colonial territories, for exam-

ple, Bantoe, Bantou, Bantu (Wharton 1925: 76). The shift to standardised alphabets did not appreciably alter this situation because of the number of such alphabets that have been proposed and, to some extent, applied. Wharton (1925) mentions 146 transcription schemes that had been proposed from the 17th through the 19th centuries, and he bemoans the strife caused to librarians by the variant transcriptions of a single language name, for example, chi-Suahili, kiswahili, Swaheli (p. 78).

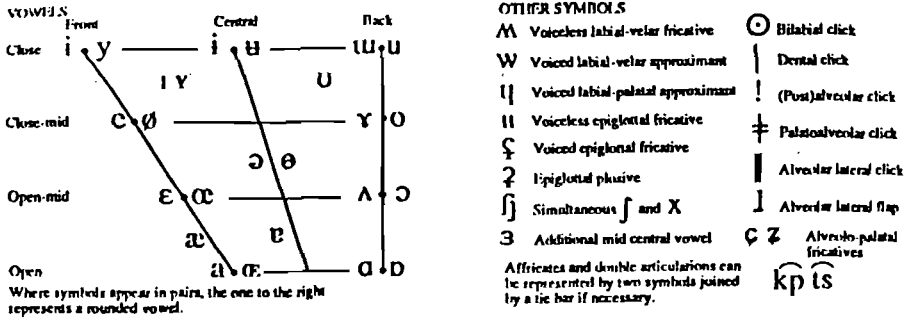
The twentieth century has seen progress in the standardisation and harmonisation of writing systems through the introduction of the International Phonetic Alphabet (IPA) in 1888 by the International Phonetic Association (Association Phonétique Internationale). The IPA is illustrated below in Figure 1.

Figure 1. International Phonetic Alphabet (reprinted from Ladefoged 1990: 551)

CIPHERS																							
	Labial		Labiodental		Dental		Alveolar		Postalveolar		Retroflex		Palatal		Velar		Uvular		Pharyngeal		Glottal		
Plosive	p	b				t	d			ʈ	ɖ	c	ɟ	k	g	q	ɢ			ʔ			
Nasal		m		ɱ			n				ɳ		ɲ		ŋ		ɴ						
Trill							r										ʀ						
Tap or Flap							ɾ				ɽ												
Fricative	ɸ	β	f	v	θ	ð	s	z	ʃ	ʒ	ʂ	ʐ	ç	ʝ	x	χ	χ	ʁ	ħ	ʕ	h	ɦ	
Lateral fricative							ɬ	ɮ															
Approximant				ʋ			ɹ				ɻ		j		ɰ								
Lateral approximant							l				ɭ		ʎ		ʟ								
Ejective stop	pʼ						tʼ				ʈʼ		cʼ		kʼ		qʼ						
Implosive	ɓ	ɗ					ɠ	ɡ					ɥ	ʙ	ɟ	ɠ							

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Figure 1 (continued)



This alphabet has had several advantages over previous alphabets such as Richard Lepsius' Standard Alphabet (1863) or Alexander Melville Bell's Visible Speech (1870). First, the IPA is a co-operative effort by an international association, unlike the previous efforts of individuals. Thus, from its inception, the IPA has had broader-based support than earlier efforts. In addition, perhaps because of the collectivity of the effort, perfection or completion has never been claimed of the IPA. Rather, it is a perpetual work-in-progress, subject to changes as new phonetic and phonological facts are learned or as problems in its practical application are discovered (see, e.g., International Phonetic Association 1989). The IPA is in this way unlike the previous works of individuals, which were often advertised as the solution to all the problems of transcription and transliteration, and which regularly fell out of fashion when they failed to adapt to new tasks.

The IPA has also benefited from the timing of its introduction. During the end of the 19th century, linguistics began to establish itself as a scientific discipline. The IPA was introduced to an academic community which had already begun to be convinced of the value of linguistic study.

The IPA allows for transcription at various levels of phonetic accuracy. This feature makes the IPA appropriate for many different orthographic tasks and audiences. Diacritics and other symbols marking non-contrastive phonetic features may be employed by linguists studying a particular language, but may be left out in the transcription of texts. Markers of features such as tone may be employed at various stages in literacy training, but may be left out by more experienced writers where the context is sufficient to make clear the relevant features. Thus, the IPA is useful as a tool of linguistic study, but may be applied such that non-linguist users are not burdened with an abundance of phonetic information.

While these features of the IPA have allowed it more success than previous alphabets, they also prevent it from being a truly One Sound / One Sign system. Among users of "the IPA", there is much variation among symbols,

such as *i* - *ɪ*, *ʃ* - *ʂ* and *tʃ* (*tʂ*) - *č*. Because most languages have few distinctive liquids, *r* is often used, for typographical ease, to represent whatever liquid exists in a particular language, regardless of whether the phoneme is better described by *R*, *ʀ*, or *ɹ*. Furthermore, some phonemes present in African languages, such as prenasalised consonants ([*mb*], [*ŋg*]) and labio-velar stops ([*gb*], [*kp*]), do not have single-graph representations in the IPA.

If the IPA is used as a basis for a literacy orthography, these orthographic / phonetic inconsistencies are not problematic, despite the failure of the One Sound / One Sign principle. As a universal standard alphabet, the variation of sign-symbol correspondence is only problematic for a particular language if variant forms are used within that one language. Such random assignments of graphs are rarely included in the orthography of a single language, unless competing orthographies are designed (in which case, one form will eventually prevail). Furthermore, universality in graph assignments does not necessarily benefit the languages in which literacy is being established. For example, Winter (1983) argues that differences in orthographic convention for long vowels in two related American languages, Walapai and Havasupai, are helpful to the speakers of the languages (and perhaps to linguists as well) because they visually differentiate the two languages (which are otherwise very similar), helping to preserve the strong linguistic pride and identification which exists in both cultures despite the near-identity of the two languages in linguistic terms.

Other breaches of principle in applying the IPA, such as the use of *r* to represent different sounds in different languages, are beneficial (or at least, not harmful) to the language community for reasons of economy and ease in printing. The most problematic breach of One Sound / One Sign is the use of digraphs (combinations of two graphs) to represent one sound in the labio-velar stops of many African languages. The linear order of the two signs is misleading: *gb* is not a series of sounds, [*g*] followed by [*b*], but instead it is one stop with double articulation (i.e., simultaneous closure at two points of articulation). The IPA's failure to manage a single symbol for a single labio-velar phoneme may reflect the Indo-European bias of the International Phonetic Association. Having spent its efforts and graphs on the languages more familiar to its members, it fails to allow West African languages one sign for each sound. Because these African languages are not apt to have consonant clusters which would be orthographically similar to the double-articulated consonants, the use of digraphs should not be any more problematic to the native speaker of Beng than the use of digraphs such as *ch* is to native speakers of English, French, or German. However, the use of digraphs such as *gb*, *kp*, and *ŋm* is misleading to speakers of West African languages who are already familiar with the roman alphabet. Because a European alphabet is involved here, learners and speakers may tend to treat Europeans as the experts on the language, as the ones who know the "right" way to read, and this may under-

mine native speakers' confidence in their own language, language use, and intelligence, as was the case for Aymara.

3.2.3 Recent developments in Africa

In Africa, the desire for standard alphabets has grown among language and education policy-makers in order to simplify the educational needs of multilingual nations. Within languages and countries, and across the continent, standard alphabets have been proposed and legislated as official alphabets. In Tanzania, the alphabet of Kiswahili developed in pre-independence days was retained after Kiswahili was declared the official language of the nation. This system retains many orthographic anglicisms, including digraphs such as *ng* ([ŋ]) and *th* ([θ]) (Dalby 1986). In Central African Republic, an official orthography for the *lingua franca* Sango has been developed with a reduced roman alphabet, 22 letters and no digraphs.

Alphabets for multilingual use have been developed and adopted officially in several West African nations. In Ghana, a cross-lingual script has been developed which introduces non-roman graphs and avoids the need for digraphs (Dalby 1986). Other nations, such as Benin (Tchitchi and Hazoumé 1983), Côte d'Ivoire (Kokora 1979), and Mali (Dalby 1986), have adopted alphabets involving both digraphs and non-roman graphs. However, at least in the case of Côte d'Ivoire, the official alphabet has had little effect on languages for which orthographies had previously been established (except in certain official efforts to utilise the official alphabet), thus failing to harmonise the diverse orthographic systems of the nation (Alma Gottlieb, p.c.). Furthermore, the official orthography has not affected the current alphabetisation of languages, judging from the disparity between the official alphabet and the efforts of the Summer Institute of Linguistics (1984).

In spite of the establishment of so many official alphabets at the national level, a pan-African orthography has been attempted. The African Reference Alphabet (ARA) was developed through a series of conferences in Niamey, Nigeria with the support of the International African Institute and UNICEF. Devised in 1978 and revised in 1982, the ARA includes 60 graphs, many of them roman, roman-derived, or Greek, and many of which exist in the IPA and are given phonetic values similar to those of the IPA.⁷ However, these graphs do not represent fixed phonetic values in and of themselves. Instead, phoneme-graph assignments are made on a language-by-language basis, in accordance with the following principles: (1) one sound / one sign (no digraphs); (2) graphs included in a language's orthography should be maximally distinct; (3) the same graph should represent the same sound for all the languages within a single country; (4) a language spoken in several countries should have the same alphabet in each country; (5) the same sound should be represented by the same graph within each subregion of Africa; (6) diacritics

should be avoided (Mann and Dalby 1987: 207). Figure 2 shows the corpus of graphs in the ARA.

Figure 2. African Reference Alphabet (reprinted from Mann and Dalby)

a.r.a consonants												
	bilabial		labiodental		labiovelar ¹		alveolar		alveopalatal		pharyngeal	
					dental ²		lateral		palatal		glottal	
							postalveolar/ retroflex		velar		uvular	
nasals ³	m		ɱ	n ²	n		ɳ	ɲ	ɳ	ŋ	ɴ	
plosives	p b		ɸ β	t d ²	t d		[ɖ]		c j	k g	q	ʔ
implosives	ɓ			ɗ	ɗ				ɟ	ɠ		
ejectives/ aspirates ⁴	ɸ			ɗ	ɗ				ɟ	ɠ		
clicks ⁵	ɓ			ɗ	ɗ	ɬ				ɠ		
affricates ⁴								c j				
fricatives	ʃ ʂ	f v		θ ð	ɬ ɮ	ʃ ʂ	ʃ ʂ	ʃ ʂ	ç ʝ	x ɣ	χ ʁ	ħ ʕ
roll/flap ¹⁰					ɾ							
continuant	w	ɔ ¹¹				l	ɭ		y			

a.r.a vowels		
i	ɪ	u
e	ɛ	o
ɛ	ɛ	ɔ
ɐ	ʌ	ɔ

Adoption of a pan-African alphabet is indeed practical in economic terms: typewriters and other printing equipment including these 60 characters would be marketable in any part of Africa and available to users of any African language. The ARA is also practical in its flexibility. It attempts harmonisation within languages and regions only. Thus, it does not require a separate graph for every phoneme in Africa and can easily avoid the use of diacritics for distinguishing, for example, the dental [t] of one language with the alveolar [t] of another. However, given the investments that nations have already made in their own official alphabets and other political and cultural differences amongst the African nations, widespread adoption of the ARA seems unlikely. It is likely that the ARA project was in large part an academic and bureaucratic exercise.

4. Further orthographic concerns

4.1 Division of words

One of the most basic problems in setting a language to writing is the division of speech into word units. In order to discuss such a problem, it would seem

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2011)

necessary to define word in linguistic, non-orthographic terms. Yet the concept of word-hood is one that often eludes linguists. We are accustomed to thinking of words as units unto themselves. This perception is aided by our experience with words as printed units with spaces between them. In ordinary speech, however, there are no spaces between words. These spaces must be devised when we put the language to writing.

Units of meaning may be discernible in speech, but all units of meaning are not necessarily words. The minimal units of meaning, morphemes, are of two types: bound and free. Bound morphemes appear in speech only if juxtaposed with some other syntactically and semantically appropriate morpheme. Free morphemes can occur in utterances without affixing to something else. For example, the plural marker -s is a bound morpheme in English; it must attach to a count noun. House is a free morpheme. Free morphemes are not necessarily words, however. While house is a free morpheme, it is only a part of a word in houseboat. Furthermore, bound morphemes may be written as words. For example, in orthographies of Bantu languages, one sees a great deal of variety in the treatment of verbal prefixes, such as subject concord, incorporated objects, tense and aspect marking. So, subject concord markers are not treated as separate words in Lingala, but are treated as separate words (and are frequently referred to as 'pronouns') in Venda.

Wolff (1962) discusses three strategies for word division. Alternative A does not allow bound morphemes to be written as separate words — they are written as prefixes or suffixes to free morpheme words. Alternative B is to use a European orthography as a model, and to divide words on analogy with European words. Alternative C is a compromise strategy in which one tries to represent bound morphemes as parts of larger, free morphemes, while avoiding words that are excessively long. Wolff develops and argues for this last strategy, intending that it serve the structural reality of the language as much as possible while guarding the language's legibility. This seems practical and linguistically sound for most African languages (although difficult and controversial to apply in heavily incorporating languages). This, of course, requires extensive grammatical study of the language prior to alphabetisation.

4.2 Capitalisation and punctuation

Usually, capitalisation and punctuation conventions are modelled on the conventions of the colonial language or other European language. Since capitalisation is associated with importance and sacredness, this Eurocentric approach risks cultural insensitivity. If those things that are sacred to the European are capitalised, but those things that are sacred in the indigenous culture (e.g., for the Beng, the sky) are not, the orthography works a not-so-subtle devaluation of the indigenous belief system.

Aside from the ethical considerations, practical matters of graph formation create problems for capitalisation for orthographies with non-roman graphs. For example, in systems based upon the IPA, upper case characters must be designed for graphs such as ŋ. This adds to the expense of including non-roman graphs, requiring means for printing two distinct characters for each non-roman graph included. Furthermore, some IPA or pseudo-IPA graphs, such as ɪ and ʀ, resemble the upper case forms of other graphs. Thus, I (the lax vowel) and i/I are likely to be confused, and an uppercase form must be invented for (small cap) I that does not resemble (upper case) I.

In order to avoid these ethical and practical problems, the ARA uses no capitals and a modified European punctuation system.⁸ This strategy avoids the value judgements inherent in capitalisation decisions by avoiding the decisions themselves. Thus, it is not controversial whether names of persons, sacred objects, deities, or days of the week have initial capitals, since no capitals exist.

Several national alphabets are like the ARA with respect to capitalisation. For example, the Nigerian alphabet, for which a typewriter has been produced, includes no upper-lower case distinctions among its graphs, using mostly graphs that correspond with lower case roman graphs. However, in practice, people who are literate in European languages often copy their capitalisation conventions into languages with non-capitalising alphabets. For example, the Ivoirian alphabet (Kokora 1979) includes no upper case forms. However, initial capitals at the beginnings of sentences and in proper names are used in a journal published at the University of Abidjan in the Baule language (with French translations on opposite pages), which purportedly employs the official Ivoirian orthography (Loucou 1981). However, because certain graphs included in the Baule text have no upper case forms, this capitalisation convention could not be universally applied. Furthermore, initial capitalisation is irregularly applied to the titles of articles and the journal itself, as well as to the names of ethnic groups. With such irregular observation of a capitalisation convention, it is difficult to understand the motivation for using capitals in the first place. Perhaps the authors and editors of the journal thought that a journal without capital letters might look "unacademic", or perhaps the overwhelming availability of French printed materials, in comparison to Baule printed materials, has instilled in the authors a subconscious bias toward capitalisation in the French fashion. If a system without capital letters is to survive as such, it is necessary that literacy education in that system include some sort of reinforcement of the idea that orthographies without capitalisation are as worthy of print as those with capitalisation, and that the sacred objects of cultures with capitalisation are not more sacred because they are capitalised. (These are real problems because of the use of European languages in schools. Any person with more than a couple of years of formal education will know of the capitalisation conventions in some European language.)

5. The Beng case

I now turn to a case study and critique, which concerns the following questions: Will (and should) development of an orthography for the purpose of western study of an African culture affect future attempts at alphabetisation and literacy in that language? For the Beng language of Côte d'Ivoire, what is the best orthography that can be developed for linguistic and ethnographic materials? The answers to these questions often became clear too late to influence the orthography used in the Beng-English dictionary (Gottlieb and Murphy 1995). Therefore, some of the following discussion is more rationalisation than criticism of the orthography chosen for Beng. However, close attention is paid to the usefulness of this orthography for the purpose of linguistic study and establishment of literacy.

The Beng are an ethnic group of approximately 10,000 people, divided into two kingdoms which use slightly different dialects of their language, Beng (Gottlieb 1992).⁹ The language belongs to the Southern Mande group, which also includes Wan and Toura.

Information for the dictionary was first collected in 1979-80 and again in 1985 by Alma Gottlieb during anthropological field study. The result of these visits was a collection of index cards carrying over 2000 words plus hundreds of additional lexicalized phrases, transcribed in an IPA-like style. As linguistic study was not the foremost purpose of her visit, and as she is not trained in linguistics, Gottlieb's transcription of Beng words was often irregular; the sketch of Beng phonetics presented here is based on her transcriptions and a list of Beng phonemes produced by a member of the Summer Institute of Linguistics. The remainder of this section consists of subsections on the following topics: graph assignment for consonants and vowels, capitalisation, word division, and adaptation of borrowed words.

5.1 Consonant graphs

The bulk of the Beng consonants are common to West African languages. However, the written materials from Gottlieb's field research leave some questions as to the exact consonant inventory of Beng. Table 1 shows the graphs for the consonants of Beng used in the forthcoming dictionary, in contrast, where appropriate, with graphs of the IPA, the Ivoirian national alphabet, and the ARA.

Table 1. Beng consonants

	BILABIAL	LABIO-DENTAL	LABIO-VELAR	DENTAL	PALATAL	VELAR	GLOTTAL
STOP							
[-VOICE]	p		kp ¹	t	c	k	
[+VOICE]	b		gb ²	d	j	g	
[+NASAL]	m		mgb ³	n	ny ⁴	ŋ	
FRICATIVE							
[-VOICE]		f		s			h*
[+VOICE]		v		z			
LIQUID				l/r			
TRILL				r ⁵			

* used extremely rarely, mostly in onomatopoeia

¹ ARA = β

² ARA = β

³ IPA, Ivoirian = ŋm; ARA = ~

⁴ IPA, ARA = ŋ, Ivoirian = ny/ŋ

⁵ Because of uncertainty as to true nature of this phoneme, equivalents cannot be given for other orthographies

⁶ IPA, ARA = r; not differentiated from other central liquids in Ivoirian alphabet

The dictionary orthography has a few problems that might have been prevented by a more thorough phonological study of the language. The first of these problems involves the possible existence of a labio-velar nasal stop, which was not included in Gottlieb's original estimation of the consonantal system. Instead, this phoneme was treated as a prenasalised /gb/ and was spelled **mgb**.

However, on closer examination of the other possible prenasalised stops, i.e., those consonantal clusters spelled as a nasal followed by a stop, the existence of prenasalised consonants in Beng becomes doubtful. There are no word-initial **mb** clusters, and the recorded **nd** and **nt** clusters involve the contraction of the first-person singular possessive pronoun, **ŋ**, and a kinship term that begins with [d] or [t]. In these cases, the **nd** does not represent an instance of prenasalisation, but of a syllabic [n] followed by a stop, i.e., two separate phonemes (belonging to different morphemes). Yet there are several instances of initial and medial clusters recorded as **mgb** (also **ngb** in Gottlieb's field notes). In these cases, contraction of two morphemes is clearly not the case. SIL work on Beng (1984) also posits a labio-velar nasal consonant.

The Ivoirian alphabet and IPA represent such a phoneme with a digraph, **ŋm**, while the ARA utilises a single graph which resembles a compact, oversized tilde. The dictionary persists in using a trigraph, **mgb/ngb** because of the

perceived explosive nature of the consonant; at the outset, the phoneme is nasal in pronunciation, but it is released with a [b]-like sound.

The trigraphic choice is deserving of some criticism. Not only does it carry us far from the One Sound / One Sign principle, but it is misleading in its suggestion of a *prenasalised* consonant, rather than as a nasal consonant. The explosive [b]-like ending of the phoneme is most likely a sub-phonemic, and perhaps allophonic, characteristic of the nasal phoneme. Including the **b** in the orthographic representation of the phoneme may be helpful to westerners attempting pronunciation of Beng, but it would be unnecessary for the Beng, should a popular orthography be developed from this dictionary. The Ivoirian and IPA representation, **ɲm**, would serve as a phonemically accurate representation for the sound and would be recognisable as that phoneme to any native Beng speaker who acquires literacy in Beng (a Beng speaker would not require so much phonetic detail as the **b**). By employing **ɲm**, we would also reduce the trigraph to a digraph, increasing readability. I have not lent serious consideration to the ARA single graph, as it is so unavailable in printing and typing equipment. If the ARA gains acceptance and use in western Africa before serious attempts at Beng alphabetisation for literacy are made, the ARA graphs would provide a more elegant representation of the labio-velar nasal than the suggested digraph.

The second major phonemic problem in the Beng orthography is the use of both **l** and **r** for what appears to be only one phoneme. A thorough investigation of the possible complementary distribution of [l] and [r] cannot be made from the written data; however, the written data do provide several clues which indicate that the two graphs represent the same phoneme. A first piece of evidence is the absence of words beginning with **r**, while many words begin with **l**. Furthermore, intervocalic **l** is much more prevalent than intervocalic **r**. Consonant-liquid clusters also reveal evidence of complementary distribution. The dictionary includes several words with initial **sr** but none with initial **sl**. The graphs **l** and **r** were used interchangeably in many words in Gottlieb's vocabulary cards and field notes, especially in stop-liquid clusters. Gottlieb (p.c.) confirms that [l] and [r] use seemed interchangeable to her in many words and among individual speakers.

One could rationalise that the **l/r** graph distinction is useful in recording allophonic distinctions. This rationalisation is actually somewhat rational for the purposes of this dictionary. As it is to be published by a linguistic organisation (Indiana University Linguistic Club), linguists will presumably be the dictionary's major audience. It is possible that our **l/r** distinctions will enable some linguist to determine the distribution of liquids in Beng. But should this orthography have an effect on future alphabetisation of Beng for popular literacy, the **l/r** distinction could be problematic. The excessive number of graphs for a single phoneme could make spelling difficult for native Beng writers, as the **l/r** distinction might often seem quite arbitrary.

The use of digraphs (**gb**, **kp**, **ny**) is forgivable despite the blatant disregard for the One Sound / One Sign principle. Utilising digraphs for the labio-velars simply follows the convention established by the IPA and adopted by most others, including the Ivoirian national alphabet. Though the ARA does include single-graph representations of the labio-velars, these graphs are unavailable in typing and printing equipment (except that which is especially made for the ARA), and therefore it would be burdensome for general use as well as for use in our dictionary.

The palatal nasal, which the dictionary represented with **ny**, does have a single-graph representation in the IPA and elsewhere, **ɲ**. In the Ivoirian alphabet, either **ny** or **ɲ** may be used, presumably depending upon one's typographical capabilities. Such a flexible solution seems appropriate for this sound, as the IPA graph is not only uncommon in typing equipment, it is difficult to differentiate from other nasal graphs in the handwriting of many. If the Ivoirian alphabet gains wider acceptance before Beng is alphabetised, it is reasonable to expect that both of the variant graphs will be employed by Beng-speaking individuals.

5.2 Vowel graphs

Beng has seven vowels, as illustrated in Figure 3. Two of the graphs used in the dictionary are not contained in the roman alphabet, but are used in many African orthographies (including the Ivoirian system and ARA) and in the IPA and are therefore somewhat available for typesetting and on African typewriters.

Figure 3. Beng Vowels



One controversial element of Beng phonetics is the status of nasalised vowels. The SIL missionaries researching Beng (1984) identify four nasal vowel phonemes in Beng, whereas Gottlieb and Murphy treat nasal vowels as pronunciation variations of vowel-nasal clusters (the nasal consonant of these clusters is deleted in certain contexts, leaving a nasal-coloured vowel). However, the dictionary treatment does not address these situations in a regular manner. The tendency is to use a vowel graph + nasal consonant graph in word-final

contexts and to indicate nasal(ised) vowels elsewhere by a diacritic tilde (~) over the vowel.

The SIL treatment of nasal vowels posits four nasal vowel phonemes: /ɨ/, /ū/, /ē/, and /ɜ/. In this situation, the nasal consonant (usually an /ŋ/) that follows these vowels in many contexts is explicable as some type of *liaison*. The lack of /ē/ and /ɜ/ in the SIL system concurs with the apparent lack of distinction between medial lax and tense nasal vowels. The Beng orthography makes a distinction between the lax and tense vowels, based on distinctions in Gottlieb's field notes. This distinction is likely to be meaningless, however, considering that many languages have fewer nasal vowels than oral vowels, since fewer vowel distinctions can be made with a lowered velum.

SIL's orthography uses digraphs (vowel + n) in describing the nasal vowels, despite the phonemic status that SIL accords the nasal vowels. Considering the popular use of the tilde as an indication of the nasal feature worldwide, the SIL's use of a digraph is puzzling and misleading. By its *sequential* ordering of the graphs that usually stand for (oral) vowels and nasal consonants, the SIL suggests that these two sounds are produced sequentially.

Before mass literacy in Beng is attempted (and before further lexicographical work in Beng is done), the status of the nasal vowel must be investigated. If the nasal vowels of Beng are best interpreted as different phonemes from the oral vowels, then regular use of the tilde-marked vowel graph should be employed. In any case, regularisation of the treatment of nasal vowels is necessary, but further phonological study of Beng is needed in order to decide the best course of action.

5.3 Representation of tones

In Beng, as in many other African languages, words carry lexical tone. The style of marking of tones reflects a prevalent custom in Africa and in the romanisation of Asian languages. This custom involves iconic diacritics borrowed from European orthographies, where they serve as accent marks. The five tones of Beng are represented as the following:¹⁰ \checkmark (mid — no diacritic), $\acute{\checkmark}$ (high), $\grave{\checkmark}$ (low), \frown (falling), $\breve{\checkmark}$ (rising). Other options for marking tone have been established, but they involve either marking some tones by sub-graph diacritics or placing the tonal symbol before or after the syllable instead of placement on the vowel (IPA 1957). While these practices may make symbolic or phonetic sense, popular preference for the iconic accent system did bias the choice, as did the relative ease with which the accent diacritics can be produced in writing, typewriting, or word-processing. Use of this system should not be problematic in either academic work or in a future alphabetisation of Beng for literacy purposes. In the latter situation, tone-marking might be eliminated in most contexts, save for those in which lack of tone-marking induces a pragmatically plausible ambiguity. Our dictionary should not be used as a model for tone-marking in Beng, however, because many tones may be missing or

incorrect. Because Beng tones undergo changes dependent on neighbouring tones and because many lexical items in the dictionary were recorded from experiences of them in sentences, the underlying tone of any word may have been misinterpreted by the American recorder.

5.4 Adaptation of borrowed words

Because our dictionary attempts a phonetic orthography, borrowed words have been spelled as they sound when pronounced by a Beng speaker, not as they are spelled in their original language, for example, Beng karfür for French carrefour. For personal names borrowed from neighbouring languages (such as Baule), the dictionary indicates the popular spelling of the name within the entry. While it may seem unorthodox to use a spelling other than the spelling that is used locally, the very French-style spellings of the names represent much different sounds in the Beng orthography than the names actually have (e.g., nesan replaces the spelling Nguessan and yakuba replaces Yacouba(h)). This juxtaposition of phonetic spellings with the popular spellings (which are somewhat official because of their use in identification papers) may help distinguish the dictionary's orthography as a phonetic, descriptive orthography, which is not in itself a practical orthography for popular use.

5.5 Capitalisation

The dictionary follows the ARA in its non-use of capitals in all contexts, even personal names. This strategy avoids forcing western value judgements upon the orthography of Beng. Thus, it does not allow westerners to decide which (if any) sacred objects, characters, or traditions should have capitalised names. The non-use of capitals also reinforces the idea that ours is a phonetic orthography designed for linguistic study, not for popular writing. If the ARA system or the non-capitalised Ivoirian alphabet becomes popular in Côte d'Ivoire, future orthographies may not include capitals as well. If not, choices must be made for a popular orthography as to which words deserve initial capitalisation. Ideally, the Beng people themselves will be accorded the choice of capitalisation conventions.

5.6 Word division

Decisions as to which morphemes and strings of morphemes would be treated as words depended upon both knowledge of Beng morphology and English-speaking biases, an accidental adherence to Wolff's (1962) compromise strategy C (see Section 4). Bound morphemes are treated as affixes, such as the passive

suffix -le (used in the verbal citation form, as in trile 'to be / make black' < tri 'black' + -le) and the 'one removed' temporal suffix -ze (gbleze 'day before yesterday' < gble 'yesterday' + ze). As a matter of policy free morphemes are treated as words in any case where sound changes have not rendered the morpheme orthographically dissimilar to its pronunciation in isolation. So, for example, blonyile ('to be happy' < blon 'liver' + nyile 'to cool') is treated as one word because of the loss of the [ŋ], while a compound without sound change is treated as two words, such as yiru yali ('hyena' < 'night' + 'walker'). (One exception to this is the set of patriclan names, which all end in the morpheme len 'child'. For consistency across patriclan names, all are written as compounds, regardless of whether any sound change is evident in the constituent morphemes.)

Apostrophes indicate certain contractions, most notably the address forms of many kinship terms, which contract the first-person, singular, possessive pronoun with the reference form of the kinship term. In many of these cases, some sound change does occur in the contraction, usually the deletion of (or assimilation of nasal qualities in) the initial stop of the kinship term and place-of-articulation assimilation in the possessive pronoun (e.g., n + da → n'a '(my) mother'). Use of the apostrophe differentiates the contractions from similar, uncontracted words (e.g., n'a 'mother (address)' vs. na 'wife (reference)') and does not misleadingly suggest the existence of a prenasalised consonant phoneme (as in n'toma '(my) namesake'). Because the pronoun does not undergo assimilatory sound changes in other contexts of use, possessive pronouns are treated as words in other contexts.

These choices in word division seem morphologically and phonologically well-motivated, however at least one other choice may deserve some criticism and revision. The postpositions of Beng, on analogy with the prepositions of English, were treated as separate words from their objects. It is possible, however, that these postpositions might be better treated as case-markers (or semantic rôle markers), and therefore (on analogy with languages with rich case-marking systems) would be better represented as suffixes, rather than words. Furthermore, many of the postpositions vary in pronunciation, depending on their phonological contexts (e.g., o/wo 'in'). Thus, in treating postpositions as words, we have broken our own rule of word division according to phonological stability. Because no serious grammatical study of Beng has been attempted, it is impossible to draw firm conclusions about the syntactic function of the so-called postpositions, but such premature assessment of the situation may bias future studies or literacy efforts.

6. The dictionary as orthographic model

Any lexicographic study of a non-alphabetic culture risks affecting that culture. If people of this culture are interested in having an orthography for their lan-

guage, but feel they lack the skills to create one, the lexicographer's orthography may be seen as a solution to their problems. For forces outside the culture that are interested in seeing the language alphabetised (e.g., government, missionaries), exposure to the lexicographer's orthography may (correctly or misleadingly) give the orthographers linguistic information with which to work. As discussed above, the lexicographer, in creating an orthography, makes many judgement calls about the language: what is a word, how is a particular sound represented, what is worthy of initial capitalisation. These decisions provide a medium by which biases toward the lexicographer's native (or preferred) tongue or his / her misunderstanding of the target language may be carried into the writing system of the language and ultimately into indigenous and external popular (and even academic) perceptions of the language.

English-speaking culture's attitudes toward dictionaries indicate the reverence we have toward their entries. Because most standard English dictionaries only record the terms of the higher social registers of English, the dictionary is used as a tool of linguistic discrimination. The grade school teacher's cliché, "Ain't isn't a word — it's not in the dictionary," demonstrates the authority that English speakers give to dictionaries (and lexicographers). Dictionaries are respected as fonts of linguistic "truth" and are expected to be very conservative. The esteem in which we hold the dictionary is evident in references to *the dictionary* (like *the Bible*), in which we ignore the fact that competing dictionaries exist and may represent different "truths" about our language.

If the same sort of authority is given to a preliminary dictionary of an unalphabetised language (such as Beng), the consequences could be disastrous if popular use of the orthography prescribed by the dictionary is attempted, since the dictionary is based on an outsider's perceptions of the language and not on native speaker intuitions (although it may be produced in consultation with native speakers). Therefore, the description of a language through a dictionary should be as accurate as possible, not only for the benefit of western science, but so as not to sabotage later efforts at establishing literacy in the language.

In the case of the *Beng-English dictionary*, thorough linguistic study could not be made prior to publication. The ethical issues surrounding the alphabetisation of Beng with only a cursory study of its linguistic structure conflict with the concern that other English-speaking academics have access to information about the Beng.¹¹

The sins in the *Beng-English Dictionary's* alphabetisation of Beng are mitigated by a number of factors. First, the text of our dictionary is in English, which is not widely used in Côte d'Ivoire. Thus, Beng speakers and Ivoirian educators will have limited access to the dictionary and will be less likely to use its orthographic model than if we had written in French (the official language of Côte d'Ivoire). The fact that the dictionary is published by a small organisation in the United States adds to the dictionary's inaccessibility as a model (although at least a dozen copies are now available in Beng villages). As discussed above, some of our orthographic conventions (non-capitalisation, phonetic spelling of

borrowed names) reinforce the perception of this orthography as a scientific endeavour, not a first step toward a Beng written literature. Disclaimers in the dictionary front matter may also serve to lessen the perception of our dictionary as authoritative.

Upon hearing of the imminent publication of this dictionary, some Beng people expressed the hope that this dictionary might help them learn English (Philip Graham, p.c.). It is interesting to note that Beng people who are literate in some language (most likely French) have assumed that they will be able to recognise Beng in print. The fact that this orthography is based upon the IPA may make this task more difficult than these Beng individuals imagine (they may be expecting the French-based type of spelling one often sees for their neighbouring languages, including Baule / Baoulé). This dictionary was not intended for such a task, and therefore would be of very limited use to anyone wishing to learn English. However, with the input of a literate Beng person, it may be possible to revise the dictionary, first through more careful linguistic study, then by the addition of an English-Beng section, so that this study can serve the Beng people.

Notes

1. Arguments that literacy aids intellectual development have been challenged in recent years. Eisenstein (1979) traces European history to the conclusion that many of the social advances thought to have been caused by the invention of writing were actually products of the invention of printing technology. Scribner and Cole (1981) argue that while the development of certain intellectual habits is frequently attributed to literacy, these habits are instead the product of the discipline of formal education. In their study, individuals who attain literacy skills in the informal village settings in Liberia and Sierre Leone did not attain the same intellectual skills characteristic of those who gained literacy through formal schooling.
2. According to Bendor-Samuel and Bendor-Samuel (1983), this problem has been encountered by missionary literacy workers in Ghana. Although village schools taught students to read and write in English, the missionaries could not find literate, adult villagers. This was because those who completed their educations quickly became unsatisfied with agricultural life and moved to the cities to find means of employment that would allow them to utilise the skills gained in school.
3. Of the 43 African nations discussed in Chrystal (1987), only seven do not have European official languages. Of these seven, six have Arabic as an official language. Only Ethiopia has only an indigenous language (Amharic) and its writing system predates European influence (Dalby 1986).
4. For further discussion of orthographic principles, see International Phonetic Association (1957), Williamson (1984), Mann and Dalby (1987: 207-211).
5. Other alternatives are apparently considered so unviable that they are not even discussed. In *Practical Orthography in Nigeria*, Kay Williamson' (1984:13) takes "for granted that the Latin

- alphabet, particularly as it is applied to English, is to form the basis of the representation of Nigerian languages."
6. This principle is also known as "One Phoneme/One Graph" and "One Phoneme/One Grapheme." In order to avoid the debate over the usefulness of the concept 'grapheme', I have chosen to use a more neutral, albeit less precise, label for the principle.
 7. The ARA's Principle 9, with relation to existing orthographies, states "any a.r.a letter that coincides with an i.p.a letter should have as its most frequent realisation a value concordant with its i.p.a value, or other widely established usage" (Mann and Dalby 1987: 208, punctuation and capitalisation as in original).
 8. See the quote in footnote 7 for an example of ARA-style punctuation and (non-)capitalisation conventions.
 9. The Beng have been referred to by other sources as the Ben, the Ngan, the Gan, and the Ngen (Mundt 1987: 31). The name Beng is preferred here as it is an auto-ethnonym.
 10. The *v* in these examples is intended to serve as a generic symbol for "vowel" and is included in order to demonstrate the placement of the diacritic above the vowel graph.
 11. The desire to produce a record of Beng is not completely self-serving for the academic community. With the epidemic of language death around the world, it is vital to record minority languages now in an effort to retard their demise or at least have a cultural record in case of the language's death. (See Linguistic Society of America 1993.)

Acknowledgements

I am most grateful to Alma Gottlieb who has taught me everything I know about Beng and who has patiently put up with my incessant questions. Virtually all of the information concerning Beng in this paper is derived from my conversations with Alma, however, I bear all responsibility for the interpretations of her work here, and I sincerely hope that I have not misrepresented Alma or the Beng in any way. Parts of this work were presented at the *Annual Conference on African Linguistics in East Lansing, Michigan* in 1992, and I am grateful to the audience there for their comments.

References

- Bell, Alexander Melville. 1870. *Explanatory Lecture on Visible Speech, the Science of Universal Alphabets*. London: Simpkin, Marshall, and Co.
- Bendor-Samuel, David H. and Margaret M. Bendor-Samuel. 1983. *Community Literacy Programmes in Northern Ghana*. Dallas: Summer Institute of Linguistics.
- Coulmas, Florian and Konrad Ehlich (Eds.). 1983. *Writing in Focus*. Trends in Linguistics: Studies and Monographs 24. Berlin: Mouton.
- Crystal, David. 1987. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Dalby, David. 1986. *L'Afrique et la lettre / Africa and the Written Word*. Paris: Fête de la Lettre.

- Eisenstein, Elizabeth L. 1979. *The Printing Press as an Agent of Change*. Cambridge: Cambridge University Press.
- Gottlieb, Alma. 1992. *Under the Kapok Tree: Identity and Difference in Beng Thought*. Bloomington: Indiana University Press.
- Gottlieb, Alma and M. Lynne Murphy. 1995. *Beng-English Dictionary*. Bloomington: Indiana University Linguistics Club.
- Hardman, M.J. (Ed.). 1981. *The Aymara Language in Its Social and Cultural Context*. Gainesville: University Presses of Florida.
- International Phonetic Association. 1957. *The Principles of the International Phonetic Association being a description of the International Phonetic Alphabet and the manner of using it, illustrated by texts in 51 languages*. London: Department of Phonetics, University College.
- International Phonetic Association. 1989. Report on the 1989 Kiel Convention. *Journal of the International Phonetic Association* 19: 6780.
- Kemp, J. Alan. 1981. Introduction to Lepsius's 'Standard Alphabet.' Lepsius, Richard. 1863: ix-99*.
- Kokora, Pascal Dago. 1979. *Une orthographe pratique des langues ivoiriennes*. Abidjan: Institut de Linguistique Appliquée.
- Ladefoged, Peter. 1990. The Revised International Phonetic Alphabet. *Language* 66:(3): 550-552.
- Lepsius, Richard. 1863/1981². *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters*. (Ed. by J. Alan Kemp.) Amsterdam: John Benjamins.
- Linguistic Society of America Committee on Endangered Languages. 1993. The need for the documentation of linguistic diversity. *LSA Bulletin* 141, Oct. 1993: 5.
- Loucou, J. N. (Ed.). 1981. *kokowa: baule amannien fluwa 2* (Université d'Abidjan).
- Mann, Michael and David Dalby. 1987. *A Thesaurus of African Languages: A classified and annotated inventory of the spoken languages of Africa, with an appendix on their written representation*. London: Hans Zell.
- Mundt, Robert J. 1987. *Historical Dictionary of the Ivory Coast (Côte d'Ivoire)*. Metuchen, NJ: Scarecrow Press.
- Scribner, Sylvia and Michael Cole. 1981. *The Psychology of Literacy*. Cambridge, MA: Harvard University Press.
- Summer Institute of Linguistics. 1984. *Kwe Sewe*. 1985. Abidjan: SIL.
- Tchitchi, Toussaint and Marc Laurent Hazoumé. 1983. Le développement [sic] d'un alphabet national dans une communauté multilingue. Coulmas, Florian and Konrad Ehlich (Eds.). 1983: 187-207.
- Wharton, L.C. 1925. Transcription of foreign tongues. *Transactions of the Philological Society* 1921-24: 59-112.
- Williamson, Kay. 1984. *Practical Orthography in Nigeria*. Ibadan: Heinemann Educational Books.
- Winter, Werner. 1983. Tradition and Innovation in Alphabet Making. Coulmas, Florian and Konrad Ehlich 1983: 227-238.
- Wolff, H. 1962. Problems in Vernacular Orthographies: Word Division. *Journal of African Languages* 1(3): 221-31.
- Yapita Moya, Juan de Dios. 1981. The Aymara Alphabet: Linguistics for Indigenous Communities. (Trans. by M.J. Hardman.) Hardman, M.J. (Ed.). 1981: 262-270.

Ancestor Non-worship in Mampruli

Tony Naden, *Ghana Institute of Linguistics, Literacy and Bible Translation,
Walewale, N.R., Ghana*

Abstract: This paper first draws attention to the necessity of careful choice of terminology in ethnological discussions of the areas of culture often termed 'traditional religion', especially in the use of terms like 'worship'. There follows a discussion of some of the methodological difficulties in this area, with the suggestion that a lexicological approach to the vocabulary used by members of the society studied may make a useful contribution to understanding here. The second part presents tabulated lexical data from the Western Oti / Volta languages which illustrates the problems and possibilities discussed.

Keywords: LANGUAGES, LINGUISTICS, WEST AFRICA, GUR, RELIGION, WORSHIP, GOD, GODS, DIVINATION, SACRIFICE, COMPARISON, LEXICOLOGY, SEMANTICS, LEXICOGRAPHY, DICTIONARY, VOCABULARY, WORDLIST

Abstrait: Inapplicabilité du mot 'culte' au respect envers les ancêtres — langue mampruli. Cette étude présente d'abord une esquisse des problèmes et possibilités qu'offèrent les études du domaine fréquemment nommé 'la religion traditionnelle'. La nécessité de choisir une terminologie avec le plus grand soin est signalée, surtout dans l'emploi du terme 'culte' ou ses synonymes. Les problèmes méthodologiques sont considérés, avec proposition que l'analyse sur le plan lexicologique puisse clarifier quelques aspects de ce sujet complexe. Enfin on offre des données sur le vocabulaire des langues Oti / Volta pour illustrer les problèmes ainsi que les possibilités d'une telle approche.

Mots Clef: LANGUES, LINGUISTIQUE, AFRIQUE OCCIDENTALE, LANGUES VOLTAIQUES, RELIGION, CULTE, DIEU, DIEUX, DIVINATION, SACRIFICE, COMPARATISME, LEXICOLOGIE, SEMANTIQUE, LEXICOGRAPHIE, DICTIONNAIRE, VOCABULAIRE, LISTE-DE-MOTS

0. Introduction: Roots of This Study

In violation of all the canons of freshman composition and theme-construction, this paper has two, albeit related, aims. The first is to consider the concept and terminology of the Anthropology of Religion and the treatment of religion as a heading in ethnographic description, with especial reference to the use and usage of the term "worship". The second is to consider the lexical field of what would normally be considered the terminology of traditional religion, in the

Mampruli language of northern Ghana and its congeners. This empirical investigation is based in the Western Oti / Volta Lexinotes¹ on "sun" and "God" (#3) and traditional religion (#14).

As has been pointed out by the terminologist Fred Riggs (e.g. 1986: 2) it is the tradition in the social sciences to use "Delphic" language — to develop terms for new concepts by redefining and reusing existing or ordinary-language terms rather than inventing neologisms. In many cases this only results in global polysemy of terms, local ambiguity usually being avoided by explicitly giving the definition currently in use *in loco* (Riggs quotes Humpty Dumpty — "When I use a word ... it means just what I choose it to mean, neither more nor less"). However there is a problem in some cases where important paradigm- or worldview-presuppositions are embedded in the existing or ordinary language use of the term; in these situations there may be an unconscious carryover of conceptual baggage either into the thinking of the investigator reusing the term or — contrary to the writer's intention — into the way in which the audience understands the discourse in which it is used. This is surely a danger with the term "religion" and some of its hyponyms, amongst which I am particularly concerned here with "worship". Assuming the use of English as the language of communication (let's not open any more cans of worms!) the term "religion" for a Marxist, a humanist, a Roman Catholic or Evangelical Christian, a Muslim, Buddhist and so on, has a meaning, in several senses of the word "meaning", loaded with all sorts of different salient components, and where the communicator and the receptor differ in their personal positions the success of the communication is obviously in the greatest jeopardy.

The situation is complicated in anthropological discussion because, in addition to the connotations of the terminology to the speaker and hearer, there is also the fact that the topic of the discourse is some social group which has its own worldview which is probably different again. Whether or not we jump on currently popular bandwagons and wave the banner of postmodernism, it is certainly generally recognised these days that in presentations in the social sciences of particular societies it is necessary to be more attentive to the voice of the people concerned, the 'stories' or viewpoints of the members, or whatever our terminological predilection, and not allow our interpretation and analysis to be wholly coloured by our own worldview and agenda. This is of course by no means wholly new, though it may be more frequently and pugnaciously stated in current discussion of anthropological practice. In my own field of West African societies, and specifically those speaking Gur languages (Naden 1989) in Northern Ghana and adjacent areas, we find Jack Goody in the 'fifties referring to "interpretations of symbolic meanings which are not demonstrably explicit to the actor" (1956 [1961]: 75, fn.) and emphasising that these need checking empirically.

The danger is of swinging the pendulum too far in the opposite direction. Some development projects aim for appropriate technology and non-intrusive methods and end up leaving people with nothing beyond their traditional

methods which have kept them poor and undernourished. [We had neighbouring agricultural volunteers who wanted to introduce donkey carts instead of internal-combustion transport: then they rejected factory-built metal-framed carts in favour of locally-made wooden ones; then they wanted to get rid of pneumatic-tyred steel car wheels and use wooden wheels, but there is no local tradition of wheelwrighting, no knowledge of whether local timbers are suitable in local climatic conditions, so the wooden wheels broke and people were back to carrying things on their heads or running diesel tractors as local public transport.] Comparably if the social scientist merely presents an unedited authentic member text with a take-it-or-leave-it stance the audience remains as lacking in understanding of the society covered as before, or makes a do-it-yourself interpretation without the benefit of the training and exposure of the researcher.

These tensions are well illustrated in the field particularly under scrutiny here, that of ritual practices usually lumped by the ethnographer under the heading of 'religion', and specifically such studies in traditional African societies². On the one hand there is a wide disparity in the personal ideologies of the researchers, and on the other a considerable world-view gap between them and the peoples studied. At first sight we might consider the latter gap to be bridged by publications by educated members of African ethnic groups, such as the well-known work of John Mbiti (e.g. 1969, 1970) for East Africa, or Peter Sarpong (e.g. 1974) in my own area. *Quos honoris causa nomino* — they do at least as good a job as non-African ethnographers in dealing with a very difficult problem. However, further consideration shows that a person who is a published author in a world language and a prominent academic, let alone a Roman Catholic bishop, is rather far from any traditional African roots his family may once have had. There is also the tendency to feel that because one is a genuine indigene one can generalise one's own experience beyond the confines of one's actual observation, so that Mbiti's Kenyan background is made to speak for Africa, and Sarpong projects his Akan experience to be that of Ghana as a whole. Even nearer home, I may cite Sylvester Nkrumah's combination of useful concrete observations on a society near his home area with vague statements that extrapolate to 'African' worldview (1993). A "native anthropologist" (his own term) who shows a good appreciation of this problem is Abasi (1995, 448-9), recognising that even being a native speaker of the language can be a double-edged weapon.

My contention here impinges on the discussion initiated by Kopytoff (1971 — see, e.g., Mendonsa 1976, and, most recently, McCall 1995) — with whom I largely agree as far as my understanding goes of the Mamprusi and similar societies considered here. At least the onus of proof would seem to be on anyone who would wish to stress the appropriateness of using terms like 'supernatural' to describe the traditional attitudes of an African group to the ancestors, and would require that that proof should include volunteered statements of non-deculturated members of the society, not merely responses to research-

ers' questions. However the focus in this paper is slightly different, as it is particularly the use of the term 'worship' with which I am concerned here — my title is not 'non-(ancestor-worship)!'

What I would like to suggest is that a careful lexicological study of the vocabulary used by members of the society under consideration can help towards objectivity in such situations. At the very least the bilingual lexicographer is conscious of the pitfalls awaiting the student of cross-cultural meaning, and in addition will be aware of various techniques of analysis and presentation which may help to avoid the worst distortions. With a view to illustrating this contention, part 2 below presents the relevant vocabulary of the Mamprusi people whose language and culture is my primary area of research, in conjunction with the vocabulary of the related languages of the Western Oti / Volta (W.O/V) subgroup of Gur, largely based on the Lexinotes mentioned above.

1. Religion and Worship in Anthropology

1.1 Religion: The starting point for an attempt at a modern 'etic' definition of terms such as "religion" is Rudolf Otto's 'Idea of the Holy' (Otto 1923 / 1950). This derives the origin of religion (in the evolutionistic model which was the paradigm of choice at the time) from the feeling of awe in the face of the unusual or incomprehensible (*'mysterium tremendum'*) which before any personification or theologising may be described as recognition of 'the numinous'. With this we may compare the OED³ definition of "Religion" (sb.3): "action or conduct indicating a belief in, reverence for, and desire to please, a divine ruling power; the exercise or practice of rites or observances implying this." Here the numinous object of awe has become "a divine ruling power".

A further degree of interpretation and specification is found in 'Notes and Queries' (R.A.I. 1951: 175): "Religion is characterized by a belief in, and an emotional attitude towards, the supernatural being or beings, and a formal mode of approach — ritual — towards them". Here the somewhat-imprecise 'divine' has become "supernatural" which certainly introduces the western post-Renaissance contrast of nature and supernature. The subsequent sections mention "Beliefs concerning Man" (soul — death — after-life — abode of the Dead — fate of the Dead — reincarnation), "Beliefs concerning Supernatural beings and Agencies" (spirits — cults — officiants — sacred places) and "Forms of Ritual" (prayer — offerings — actions — taboo — oracles — divination — ordeal — oath — ritual language) — (ibid.: 176-187).

There follows a major division which deals with "Magical Beliefs and Practices", stating that "In magic, no appeal is made to spirits" (ibid.: 187). This is all very clear as long as we are confident that we can identify a 'spirit' when we (don't) see one! My main contention here is that in many parts of northern Ghana sacred places, oaths, ritual, taboo, divination and other things listed in

the 'Religion' sections above do not involve anything that can be appropriately designated 'spirit', 'divine / supernatural being', or 'spiritual X'.

This key problem of terminology is specifically seen in the (topic- and geographical-) area in focus here in the case of the constant use by English-speaking ethnographers of the terms "spirit", "spiritual" and derivatives (similarly French *esprit...*, German *Geist...*). The only term in the local languages rendered by speakers in English as "spirit" is *SI- -GA / -SI (MP *siiya / siisi*; DB *shii / shihi*; KLA *siig / siis*; BL *chiik / chiisa*). This is 'spirit' in the sense of that immaterial part of a human being which is present in the living but absent in the dead; this is what has the experiences contained in dreams, and witches' spirits (in some areas) travel, or send out a familiar, or use some apparatus to catch and consume (parts of) people's spirits causing otherwise inexplicable lingering illness and death. On the other hand, when scholars offer translations like "shrine-spirit/s" (for *BAG- or *BUG-: 2.2.3 Vocabulary #10 below), "spirit of the earth(-shrine)" (for *TIN-, *TIN-GBAN: 2.2.2 Vocabulary #6), and "guardian spirit" or "spirit guardian" (for *SIG-: 2.3.2 Vocabulary #12), or refer to the "spiritual" activities, forces, or nature of what is being done in divination and sacrifice, they are not reflecting members' ways of speaking of these things but adding a reinterpretation. And the standpoint of the reinterpretation is the western scholar's materialism, so that to call something "spiritual" is equivalent to calling it "imaginary", "non-existent" or "unreal". Thus Fortes (1980: vii) describes the scholarly stance as "agnostic", but it is clear from the rest of the piece that it is rather a dogmatic materialism — "Krishna is not really brought down amongst his devotees and Tallensi ancestors are not really present on or in the shrine dedicated to them" (*ibid.*: xii).

The other problem with the 'Notes and Queries' definition of magic, which also states that "the desired end is believed to be achieved directly by the ritual technique itself, *i.e.* by the use of the appropriate actions, objects or words", is that it could perfectly well cover the priming of a pump, bleeding of brakes, tuning of a radio set or surfing the internet. 'Magic' is the science of people I despise. Old-time explorers laughed at the tendency of 'primitives' to label modern technology as "white man's magic", but one might look at it the other way and see that the colonised were recognising the colonialists' technology as the same sort of thing as their own techniques which the 'white-man', feeling superior, labelled as "magic" or "superstition". What I am arguing for is not radical cultural relativism, but the recognition that in exploration of meanings crossculturally — which is a large part of the work of both the anthropologist and the bilingual lexicographer — we need to be very slow to slap off-the-peg labels from one language on the concepts of another, let alone import all their connotative baggage along with them.

1.2 Worship: This *caveat* is particularly necessary when we come to a term like "worship". The normal use of the English term has a rather specific set of senses and connotations: note that even French, for instance, has no real

equivalent, both 'culte' and 'vénération' only covering limited parts of the area designated by the English word; the former particularly, along with the Dutch 'godsdienst' and 'eredienst', corresponding rather to the English "(religious) service", and focussing more on ritual practice than on attitude. "Worship", however, has its etymology in 'worth', the OED defining "Worship" (v.1) as "To honour or revere as a supernatural being or power, or as a holy thing; to regard or approach with veneration; to adore with appropriate acts, rites and ceremonies", where only the third clausula corresponds to "culte", "eredienst", and etymologically we compare the entry under (sb.1) — "The condition (in a person) of deserving ... honour". The English term "worship", therefore, primarily refers to a virtue in a person or object which elicits a feeling of respect or awe (as in Otto's 'numinous'), and only then to appropriate behaviour. Against this sort of background of associations, a literary / theological writer like Tolkien can even define "Divinity" as "the right to power (as distinct from its possession), the due of worship, in fact 'religion'" (1964: 27).

My argument here, then, is that, in writing in English of unfamiliar cultural systems, the word "worship" should be related to recognised or ascribed 'worth', and is not to be used as a mere synonym of "ritual". I would further suggest that for the Mamprusi, and probably for a number of neighbouring and related ethnic groups, the 'culte' of the ancestors is more an extension of the 'culte' of respect and prestations offered to living elders, rather than being rendered, by the mere fact of death⁴, a totally different relation of 'worship' of the 'supernatural'. If one challenges an accepted usage one is open, of course, to the demand that one offer a better alternative: one possibility is "ancestor veneration" which McCall uses (1995: 257) in his summary of Kuper 1982 — 'veneration' can apply to the living, but also has "religious" overtones (and moreover has a corresponding French term which is no *faux ami*).

1.3 Methodological problems: The above discussion is theoretical — theological even — but there are practical methodological problems in ethnographic and linguistic research in this area of culture which are intensely practical and pressing for the field-worker.

1.3.1 The first type of difficulty, both in the chronological and in the logical sense, may be labelled as elicitation problems. How does a researcher in a culture of which he is not a native — or to which he is returning after years of alien residence and education (Abasi 1995), or of which he is a member but needs to determine how far his understandings are typical of those of fellow-members or are rather idiosyncratic — find out what people think and do about such matters? This subject area is particularly apt to be treated as esoteric — these things "are not talked about" from conviction that these are arcane matters, in which revelation of secrets to outsiders may lead to retribution being visited on the source or recipient of the revelation, or may simply deprive the practice concerned of its efficacy, either at the time or permanently. On the

other hand it may be that they "are not talked about" from lack of interest outside the actual performance itself. The Mamprusi and their neighbours are not particularly prone to secrecy and protection of esoteric doctrines and practices by deviousness or threats⁵. On the other hand the researcher may not be prepared for a non-ideological approach to life, where even in areas we may want to label 'religious' there is no system of doctrine or myth underlying the practices. This is much more the situation in this area⁶: people can describe a festival, ritual, or indeed a technological procedure, if they are just about to be involved in it (or have just been involved), but not at other times. Even the traditional month-names (Naden forthc.) can be given in correct order for the current season by many people who seem unable to give them for the other end of the year.

The danger in both these cases is the leading question. Where people do not volunteer information it must be elicited, but until you know a good deal about the matter under investigation, how do you know what to ask? And in this area, as quite widely in Africa, people are anxious to please and very ready to agree with any statement the researcher offers which is not totally repugnant to their views (in the latter case they will dodge the question). This is partly from a general desire to please and promote good social relations, partly from a cultural value of hospitality to 'strangers', and partly from a feeling that the researcher — whether outsider or local 'évolué' — is a person of wealth and power who can bring benefits if pleased and cause trouble if dissatisfied.

This is a positive approach for lubricating the wheels of social life, but it becomes a challenge to those of my generation, brought up on a modernist belief in the existence of objective truth and the possibility and desirability of seeking it. We are not necessarily speaking of what the members of the society, or even the fair-minded outsider, would consider lying. If the enquirer is worth any of his or her salt the leading question will not lead in a totally alien direction. But as an instance we may cite the tendency, in this part of Africa, for German and particularly French researchers to report a systematic myth-and-ritual complex in which the sky is considered a male being who in the rain impregnates the female earth and brings fertility of crops, animals and humans, while sacrifices and offerings at rain-shrines and earth-'groves' are worship offered to the Sky-god and Goddess of the Earth to propitiate them and promote this fertility-bringing intercourse. I am sure that many people in these areas, if presented with this scenario, would assent to it as a nice idea — indeed probably many people in the postmodern 'West', those not committed to one of the Abrahamic faiths or to ideological materialism, would do the same. But as we shall see it is noticeable that those who have spent the longest time in the study of individual peoples — Fortes and J. Goody (and, in fact, the present writer) — have not found these views expressed as such, unprompted, by the members of the relevant societies.

With regards to oral evidence we may in general be more confident that we can get an insight into something which is emic — an insider view — in the

society studied from what people *do* say, rather than from what they *can* say or are willing to agree to. This of course requires a longer time and greater patience from the researcher; maybe only those privileged to spend a long time in the 'target' area — 'native anthropologists' or those like myself who work in a long-term development project and study anthropology, as it were, on the side, or maybe academics who study a society within easy visiting range of their 'home' institution — will be able to just wait until they hear somebody say something relevant.

1.3.2 Somewhat similar to the contrast between what people *do* say and what they *can* (be made to) say is the contrast between what people *say* they do and what they *do* do. Again, observation of the more-reliable alternative is frustrated by time pressures, often due to financial constraints but also due to the non-correspondence of the dates available for field study and the occurrence of the behaviour to be studied. [For instance, Drucker Brown's excellent study of Mamprusi royalty (RAMK — see References) is slightly impaired by the fact that the culminating events — the death of a 'King' and the accession of the next — cannot be planned or foreseen and, in fact, took place just after the researcher had come to the end of her fieldwork.]

The comparative objectivity of observations of behaviour is also rendered less valuable, particularly in a semantic field like that presently under discussion, by the fact that one can observe *what* is done but not *why* it is done. Again I can only speak for the north of Ghana, but cultural communication patterns hereabouts make "why?" questions particularly ineffective for the purposes for which the apparent equivalent is used in European languages — *i.e.* obtaining information on motivations and purposes of actions⁷. A possible strategy is to ask what would happen if the activity was omitted, or was done differently. The imagining of hypothetical situations, however, seems to be something more appropriate in some cultures than others, and even where appropriate there may only be some individuals who perform well in such tasks. [Of course part of research design and its implementation is the identification and recruitment of assistants who have some skill at what they are being asked to contribute to the project, but if the task involved is too eccentric to local norms there is the suspicion that one is getting the assistant's idiosyncratic perceptions rather than what is typical of the society.] Also the answer may be too general — the result of *any* hypothetical omission or alteration of any ritual would be 'very bad'!

1.3.3 We therefore need a response to these weaknesses inherent in the method of research which begins with selecting and employing as reliable an 'informant' as can be found and then asking him (very rarely her) what people customarily do, or what they are being observed to do, and why? In fact the questions (not necessarily to a professional assistant) "What is that?" and "What is / are ... doing?" may very often yield information that might have been expected from a "Why...?" — as linguists know to their cost, as they often may

elicit lengthy functional explanations when they had hoped for a simple noun or verb. "What is the name of that thing?" is often a better question for obtaining a noun: verbs may be more difficult and the method of eliciting them more language-specific (where nominalisation is possible "What is the name of his / her / your action?" may be possible). One response which may be effective is to combine the methods so as to allow the strengths of one to minimise the weaknesses of another. For example, asking an informant what people normally think, say, or do can be given an objective anchor by starting from an observed non-elicited remark or behaviour, particularly if the assistant also observed it and knows the full context. Detailed transcription of recorded texts in the language/s used by members of the society studied can also yield supplementary insights. To focus on linguistic matters with an assistant who is able to do so can serve to defuse any embarrassment or reserve which might attend direct questioning on such topics. I use 'texts' here in the sense used in field linguistics, referring to any piece of language-use produced spontaneously by speakers of the tongue concerned.

This underlines the need for a multi-pronged approach including listening and everyday observation as well as questioning, recording of explicitly 'ritual texts', and participant observation of explicitly ritual occasions. This extends to questioning about the everyday observation, keeping an eye / ear out to observe in appropriate social contexts whether members actually do and say what informants have claimed, and observance of repeat performances of ritual texts to check whether the version recorded is invariant, typical or aberrant. Then where there are inconsistencies, the researcher has to find very delicate methods of drawing attention to the discrepancies in a way which is non-threatening to the informant (and not disruptive of relations between assistants): if this is possible, attempts to explain the different accounts may be extremely revealing.

1.3.4 In this nexus of interlocking methods, one valuable strand is the structural analysis of relevant areas of vocabulary. Lexicological analysis, particularly in the form of a lexical-field study, has the advantage of working with what people *do* say because it is their words which are the immediate object of analysis, and also shares with text study mentioned above (1.3.3) the advantage of removing the focus from behaviour to linguistics which may avoid some of the reserve which informants may feel in answering direct questions on these topics. Thus the analytical work which is a necessary foundation for bilingual or cross-cultural lexicography can also be of assistance in the ethnological understanding of the cultures which are expressed in those languages. From my special perspective I might add that by the study of several culturally- and / or linguistically-related groups comparative and etymological analysis may show up emic clusterings and partitions within lexical / semantic fields as they are preserved across language-change (Naden 1993). As an example, section 2

of this paper offers the Western Oti / Volta Lexinote (originally Number #14) on terms in the semantic area in focus here.

1.4 Practical conclusions: In consideration of a particular society's culture and language, we then need to be very careful in our use of terms like 'god', 'deity', 'divinity', 'spirit', so that each word used in the language of description is related in some systematic way to a word or idiom in the language of the members of that society, and ideally the demarcation of topic areas should be justifiable lexicologically and / or in explicit member statements. To give a concrete example, we should not say "in my worldview witchcraft is part of the belief in non-material forces which I label 'Religion' (and don't believe in)", but we examine whether there is any slippage or overlap between vocabulary, personnel, equipment or techniques between the areas of witchcraft and, say, ancestor veneration: if we find none then we do not set up a category of Religion of which these are sub-fields.

It is very difficult to achieve this degree of objectivity: the desirability is beginning to be recognised — McCall says that in considering ancestors "we must discard the boundaries of 'cult' and 'religion'" (1995: 258), but he still slips into terminology like "priestess" and the "*spiritual* and mundane functions of the shrine" (ibid.: 260 — my emphasis) in his account of Ohafia, without apparent original-language justification (elsewhere he gives local linguistic equivalents for any remotely technical terms). Similarly even Abasi uses "the underworld" (1995: 465) and "the afterlife" (ibid.: 466) in a way which is suspect of deriving from the English in which he is writing rather than the Kasem of which he is writing.

This is not to deny the existence of covert categories in languages, groupings which do not have a lexicalized label but nevertheless are part of the system. [The W.O/V languages characteristically lack lexical terms for any intermediate-level hyperonyms in the taxonomy of natural kinds (Brown 1984, 1986; Rosch et al. 1975). Thus in Mampruli there are syntactic and lexical (both morpho- and semo-) patternings which clearly demarcate a field of colour terminology comprising basic 'black', 'white' and 'red', and derived "yellow" ('dawadawa⁶-powder') and "(blue-)green" ('bean(leaf)water'), but no lexeme for "colour" or "coloured" or "to colour / be coloured / change colour".] If, however, we want to use "supernatural" or some such word to represent such a covert category, then we need to introduce the term by a presentation of the analysis which justifies the usage — and in fact this reasoning will be an important part of the substance of the description. Equally we may wish to use 'Delphically' an English word which is broader in sense than the local-language equivalent; this is valid as long as the precise situation is made clear at the outset. Thus in discussing W.O/V languages we tend to follow the practice of English-speaking members of these societies in using "spirit" to represent the *SII-GA / SI lexeme (1.1 above): this is fine as long as we make clear at the outset that this only refers to the non-material part of a human being, and we do

not perform sleight-of-hand exposition by slipping "spirit" in with wider reference, where local languages would not use *SII-GA / SI.

2. W.O/V Vocabulary of Traditional Religion⁹

In most of the groups considered here, the influence of world religions is small: my informal estimate is that the number of Christians and serious Muslims (most people follow some Muslim practices sometimes) amongst the Mamprusi is about 10% each: the remainder follow the traditional religion¹⁰. Islam may have a somewhat greater hold amongst the Mossi, Wala and Dagomba, but also much less among the Balsa, Dagaba, Frafra, Hanga (roughly speaking, the Muslim Hanga have hived off as the KaMara — Prost 1974a, Naden MS), Kamma, Nabdem and Tallensi; Christianity may be somewhat stronger amongst the Dagaba. The general aspect of traditional religion is similar among all these peoples, as is the vocabulary, though as in other cultural spheres there are distinct variations when one scrutinizes the details.

The foregoing discussion has shown that religion is a difficult area to delimit and that I suspect that much of what appears to Westerners as religion or magic, functions for the members of these societies as science — techniques which one takes on trust from experts and hopes will be effective in dealing with a practical problem ('magic' is what you call the science of people you despise?). The researcher is easily confused by his or her own beliefs and values, both personal and theoretical, or even by connotations and meanings attaching to English vocabulary, and it is very difficult to discover, or even conceive, the best way of expressing in academic, say, English, what a Mampuriga or a Frafra 'really believes' or 'means' by his words or actions in this cultural area. The objective correlates provided by a lexical study of the words used in these areas may, therefore, be particularly beneficial. In a Lexinote it is, of course, the vocabulary which is in focus.

2.1 Divination

It is logical to start with divination. The normal procedure in these matters is that one suffers, or fears, some serious problem (infertility or sickness in man or domestic animals, crop failures, drought ...), or one plans some major step (marriage, seeking a chiefship, a business venture...) and consults a diviner (Ghana English 'soothsayer'): the diviner performs his art and then indicates the necessity of making, or obtaining, or making offerings at, one or more of the various cult objects described later.

There are four major lexicalized items in this sub-field, the verb "to divine" (or "consult a diviner"), the name of the practitioner "diviner, soothsayer", the name of his apparatus "diviner's paraphernalia", and a fourth item which may

have reference to a shrine (other than the actual apparatus), the power / spirit(s) ... which reside in the shrine and / or apparatus and / or guide the divination process, or any combination thereof. If anyone can furnish precise definitions of any of these terms in a particular language-community it would be enormously helpful if they could publish or communicate them.

The clearest, most consistent item in the lists is "to divine", *BUG-:

Vocabulary #1¹¹

DIVINE, TO: DB *bugi, bagsi*; MP *bugi*; HG *bogi*; KLA *buk / bugid*; FR *buge*; MR *boge, bughe*; LW (G:MBG) *buu / buura*; BL *bogi, bog gbanta, parese, diiri*

The diviner's apparatus is usually carried in a skin bag and "diviner's (or 'divinatory') skin bag" is often used for the whole apparatus, or even (often in MP) for the man himself. I have not observed, nor have I found any detailed description of observation of, the diviner's equipment or methods in the W.O/V-speaking area of Ghana. J. Goody (SOL 97) refers to "the fall of the diviner's cowries"¹², and Deschamps (1972) mentions a small stick (*baquette*). Mossi diviners are said to observe water in a container, pigeons, or an iron 'butterfly' (*pilimpika*) which is guided by the 'fairies' (2.5 below) (Tiendrebeogo and Pageard 1975). Compounds with *BAG are common:

Vocabulary #2

DIVINER'S APPARATUS: DB *bagri, bagayule* "sacrifice ... ancestors" — Oppong; MP *ba'akɔligu / ba'akɔla, -kɔna, ba'ayuya* (lit. 'names'); KLA *ba'akolug / ba'akon*; FR *baga, bakologo*; MR *bagh-waogba*; BL *baan-nyaani (baan -yui "-bag", -yak "-rattle", -doari; gban-doari "wand"*;

I have followed this order of presentation because in several languages the practitioner himself is called, literally, a 'diviner's bag' or the 'owner' of divination. One would like to postulate a basic *BAG-A / -BA "diviner" (Manessy 1975 reconstructs O/V *BO, but forms like KN *ubɔɔ* can be derived from *BAG — cf. Vocabulary #10, fn. 19, below — more easily than *baga* from *BO in the absence of **ba* and **bog*¹³). Other items could then be derived from *BAGA; however only in MR (and Rattray's TL) is this the normal surface form:

Vocabulary #3

a) DIVINER (1): DB (F) *baga / baganima*, (Oppong) *bagsi*; MP *ba'akɔligu daana*; KLA *ba'akolug daan, ba'abugud, ba'amaan*; TL (R) *baga / bagaba, bakologo*, (F) *baya*, (GG) *ba'a /*

ba'anam; NB (R) bay, (N) bakulug / bakul; FR baga; MR baga / bagba ("féticheur" — Canu); DR (R) bogubura, (G) bagbuura / bagbuuri; LW (G) бəббуура; BR (K) бəббөрə; SF baya / bayaba; BL baano / baanba;

- b) DIVINER (2): KLA ba'abugud; MR (Yadre) bougesa / bougesanamba; NT бəккə / бəккiba; LB (R) boga; DR (R) bogobura; BL bogroa ("client"); KM buguru / buguti, bugurum;
- c) DIVINER (3): KR viisra / viisriba (viisi "examine, scrutinize"); MR geta / getba;

The "diviner's shrine", "divinatory spirit/s", or "divination" terms are not so often cited, and one cannot be sure that these glosses are selected with precision:

Vocabulary #4

DIVINATION (-SHRINE, -SPIRITS): DB bagtali; TL bayer, bakologo "divinatory ancestors"; MR bagre; LW (G) бə, bay (MBG: 58 "it was repeatedly denied to me that Bagre was an association of diviners"); BL baan-nyaani ("divinatory objects"), jadok-bogluk ("shrine of a spirit that often materializes in an animal ... may become a divinatory shrine")

The *BAG-RI/-A word also appears in "family shrine" (2.3 below): it may be that originally the lexeme meant "ancestral spirits (givers of revelation and receivers of sacrifice)", but that gloss is not appropriate to most of the specific terms used in these languages today.

2.2 Sacrifice: Group Cults

The next series of items concerns mainly "shrines" which receive sacrifice, mainly to avert actual or potential harm. The languages divide into two main groups according to whether they have a specialized verb "to sacrifice" *KA(N)B, or whether they use the ordinary verb which I translate as "to fix" in the American sense of "make or mend" ('fix breakfast / a flat tire'). The former focusses on the killing of the animal; the latter has as its grammatical object, normally, the shrine at which the offering is made, and includes the whole ceremony (libation, prayer, preliminary killing of chicks, actual sacrifice, butchering and distribution of meat). Thus one might say MP N ya m-maali m ba'ari, "I am going to make an offering at ('fix up') my ancestor-shrine"; at a funeral it

is announced that so-and-so is giving, for example, a sheep *m-maali u ba* "to 'fix up' his father" (or whatever relation the deceased was), though it is also common to say *n-ti (u ba)* "to give" which may be simply "for...". The same verb is used for any ceremony: Mamprusi greet *Ni i soosa maaligu*, "For your yesterday's 'fix-up'", as a greeting on the day after the person saluted was involved¹⁴ in any ceremony, dance, festival, etc., whether he or anyone else made a sacrifice or not.

Vocabulary #5

- a) SACRIFICE, TO (1): KLA (? *kanb* "to singe"?); TL *kaab*; FR *kaanbe*; (? cf. MP *ka* "to bless, pray" = FR *kan'* [n. *kan'an*]; LW *kaab* [n. *kan'an*]; LW *kaab* [n. *ard* v.]?)
- b) SACRIFICE, TO (2): DR (H) *bug*, (K) *буга*, (G) *bagre*; LW (G) *бугре (бугре* = DR *bagre* "initiation rites of a secret society" — SOL: 52); BR (K) *бугр*
- c) SACRIFICE, TO (3): DB *mali*, [n.] *malgu*; MP *maali / maanni*, [n.] *maaligu*; HG *maali*; KLA *maal*, [agent n.] *maan / maana*; MR *mane*, [n.] *maoongo*;

The *MAL 'fix' root occurs elsewhere without being said to be used for "sacrifice":

- c') 'FIX' (≠ SACRIFICE): KR *maali*; TL *maali*; NB *maane*; FR *maale*; NT *mani*; DR *mal*;
- d) *sara* in DB / MP (also Bisa — so maybe in MR and KLA as well) is a funeral sacrifice (and accompanying ceremonies); it is said to be from Hausa — ? sense "cut" or "custom"?

2.2.2 Public Sacrifice: the Earth

Under "land" in Lexinote #7 there is a discussion of the land in its 'religious' aspects, listing "Earth-Priest" and "Earth-Shrine, grove". The relevant items are repeated below:

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2011)

Vocabulary #6

GROVE, EARTH-SHRINE: DB (FS) *tingban* (-dana); MP *tingbanni* / *tingbana*; KR (P) *temgbaō*; TL *tungban*, *tongban*; NB *tengboŋu*; FR *tengŋo* / *tengono*, (Rapp) *tenganne* / *tengana* ("sacred grove"); *Tengongo* Earth-Goddess", *tenkuggere* / *tenkuga* ("on which the Tendaana makes offerings"); MR *teŋa*, *teŋkuga* ("sacrifice place"); DG *figan*, *tēgan*; DR *tingan*; LW *tengaan*; BR (K) *tēgan*; BL *tanggbain* / *tanggbana* ("earth-shrine [n.]; spirit of the earth; earth-shrine; sacred place of the earth-divinity ...")

This word is a compound of *TIN-/TEN-, with a range of meanings in the area of "earth, ground, down; land, soil, sand; country, town, village" (Lexinote #7), and *GBAN- of which the meaning "skin" is presumably basic, with derivatives "leather, leatherwork" and "flat things — paper (> letter > book), sheet of (various things)". It is only distinguished, if at all, by nominal declension from "region, country, piece of land" — e.g. MP *tingbanni* / *tingbana* "grove" | *tingbanŋu* / *tingbana* "region".

As to the personification of the Earth, Franz Kröger (p.c.) describes being told that a certain Buli *tengbani* could not receive sacrifices in the afternoon as 'he is not at home now'; and Patrick Adongo (p.c.) mentions a Frafra saying, when a squall of rain is seen passing, that 'the *tengŋo* is going on a visit and cloaks himself in the storm'. For DG Pénou-Achille (1982: 70) translates *figan* as "dieu de la terre"; for LW, J. Goody quotes of the *tengaan* that "he (o) loves people", and as a female derives fertility from sleeping with the male sky (*saalō*) — though in a fn. he cautions "I do not wish to suggest it is a recurrent theme in LoWiili speech, but it appears in the origin myth which I recorded". For DR / WL Tuurey (1982: 16) speaks of "tutelary spirits, the chief among them being a female earth deity known as Tenge\... The *tenga* has many *tengbama* (shrines)". In the case of MR we are told that (Somé 1970: 9) "the Earth (*Tenga*) is considered as the wife of the heavens ... has an altar of sacrifice called *Teng kuga*"¹⁵. Rapp (1966) gives for FR *tenga* "earth...", *Tenga* "Earth-goddess", *tengongo* "Earth-goddess, wife of the sky-god Yinne"¹⁵.

In TL, however, although Fortes mentions that "the Earth is commonly referred to by the third person singular pronoun *u*" (i.e. the personal pronoun, not inanimate *ka* as for common noun "earth" — see Naden 1982), he states that "there is no convincing evidence for the Earth being personified as a 'Goddess' in the Northern Territories of the Gold Coast — certainly not among the Talensi" (1945: 175, fn.), and this seems correct for at least MP, DB, KL and HG. The 'earth-as-wife-of-sky' motif therefore seems to correlate within W.O/V with just the linguistic subgroup Northwestern — MR, FR and the DR-group, though Zwernemann (1961) cites the concept from Southwest Grusi (Sissala

etc.) and Kasem-Nuni, so perhaps a geographical East / West division is in question.

Somé (ibid.) also says for MR that the country is dotted with "sacred groves inhabited by good spirits which protect man. These spirits are embodied in (*matérialisés par*) animals (snakes, crocodiles etc.)"¹⁵, with which we may compare the statement quoted by DruckerBrown (RAMK: 100) from the daughter of the Mamprusi earth-priest chief of Gambaga: "Gambaraana has earth-shrines. They are all chiefs. They are not human beings but snakes ... They aren't people of this world ... They just help to guard the world." In the neighbouring Kasem region (north of BL, west of FR, south of MR — their language is Gur but of the (Northern) Grusi subgroup, Naden 1989) "The *tangone* is a spirit which inhabits a grove of trees near the Kalogo stream, but the term seems to be used both of the spirit and of the grove indiscriminately." (Teviu / Callow n.d.: 85 fn.) — cf. FR Adongo 'the *tengona* is the trees etc.'

While a diviner can in a specific case prescribe practically any course of action, including an earth-sacrifice, it is in general true that the earth-shrine is the concern of the residents (or descendants of one-time residents) of the stretch of earth, the 'parish' (Fortes, J. Goody), covered by the particular shrine.

Fortes, however, describes a kind of 'public' shrine which he calls the "External boyar"; unfortunately he never indicates what the TL translated "External" is, or whether the Tallensi do not themselves make the distinction, but it rather arises from the anthropologist's own analysis. This shrine receives sacrifices not only from all Taleland but also from outsiders from other parts of Ghana who come to ask its help, particularly in cases of childlessness. Certainly the Earth-shrine serves a wider 'congregation' than the other shrines, and it is the 'groves' which are primarily the locus of any sacrifices not called forth by crises in the life of an individual or a family but conducted at regular seasonal intervals, such as the 'beer of the grove' *tengaan dāā* pre-farming-season rogation sacrifice of the LoWiili (SOL: 92 — "the annual sacrifice at the earth-shrine"), and among the Tallensi the comparable *Golib* (Fortes 1945: 104) and the 'beer of the shrine' (*Bɔyaraam* ibid.). Comparable regular festivals seem less common elsewhere (Naden [ed.] forthc.), though it is hazardous to argue from silence in existing reports: the Balsa have a great Harvest Festival called the *Fiok*, for instance, but it is sufficiently independent of the 'ancestor worship' complex not to require mention in Kröger 1982. Remarkably for an agricultural people, there is certainly no annual seasonal festival amongst the Mamprusi: all their festivals are based on the (non-seasonal) lunar calendar (Naden forthc.), and the Dagomba, at least, are presumably the same. In Zuarungu the Frafra refer to festivals of other peoples (TL, NB) as one form of seasonal reference (Schaefer / Adongo forthc.) so presumably do not have comparable festivals of their own, nor do the Kusaas seem to have prominent seasonal rituals. Both Fortes (1945: 104 ff.) and J. Goody (SOL: 92) note that the rogation festival outranks the harvest celebration.

The 'grove' of the Balsa sounds much more of a family possession than a 'public' shrine in the discussion of Kröger 1982, but this may be because of the particular focus of that work.

Sacrifices at the 'grove' are determined and led by the "earthpriest", 'owner of the earth(-shrine)':

Vocabulary #7

EARTH-PRIEST': DB *tindana*; NN (Sk) *tindaana*; MP / KR (N) *tindaana* / *tindaandima*; KR (P) *tembaōdaane* / *-daamba*; KLA *tendaan*; NB *tindaan* / *-daannam*; TL *tendaan* / *-daannam*; FR *tendana* / *-daama*; NK (R) *tendana*; DR (R) *tendagena*, (H) *tindaḡna*; MR *teḡsoba*; LB (R) *teḡsob*; BR *teḡgaansob*; SF *tiḡsuba*; BL *teḡnyono* / *-nyam*; KM (South) *tintiḡ* (=Sisaala); KM (North) *tindaana* (=MP)

The customary glossing of this as 'Earthpriest' is based on the reasoning: 'kills animals and speaks to the earth(S-shrine), this is sacrifice and prayer, that is religion (therefore not real) — a religious functionary is a priest'. The form is *TIN- / TEN- as above, plus the item which is semantically the same though formally different in the various languages, and is glossed by English-speakers as 'owner'. This latter form has a meaning / usage similar to the Arabic *dhu* and Hausa *mai*: it is a person (human or personification) related to, or defined or distinguished by, something which is characterisable by a noun — either a 'person, place or thing' or a grammatical nominalization. This may be used in a context like *Zupiliga daana!* "Hey, you with the hat!" (all these examples are MP); and often refers to a hawker of some particular item — *Kōddu daana be ni?* "Is there a banana-seller around?" / *Kōddu daana!* "Hey, banana girl!". It is semantically, though not formally (unless in the Northwestern *SOB- seen in MR, LB, BR, SF in Vocabulary #7) related to a verb *SU which means to be 'in charge' of something (again often glossed 'to own') both in the sense of having privilege over and of having responsibility for it. Hence *ḡḡni n su kuuri wa?* "Who 'owns' this hoe?" may imply "... if it's his we must give it back to him" or "... I'm going to give him what for — leaving it here for me to trip over!". The person 'in charge' of it in this sense is its *daana* (*soba*, *nyono*) — not "owner" in the legal sense (words like MP *dinni* "own thing" and *toori* "very own" carry this meaning), the person concerned may have borrowed the hoe, but he is currently responsible for it. At the other extreme *Daana* is used in the titles of senior elders at major chiefly courts — *KpanaRaana*, *WudDaana* (literally 'Horse.Owner', 'Spear.Owner') — and chiefly titles, *GbanDaana* being the chief of *Gbarni*, and so on. The chief does not have feudal or autocratic 'ownership' of the village, but has both privileges and responsibilities in regard to it (see

also Naden [ed.] 1984: 4f.). This would seem to be applicable in general to the relationship between the **Tindaana** and the Earth.

Mamprusi **tindaandima** are not typical — they may be priest-chiefs with ritual functions vis-à-vis the 'King' (Drucker Brown 1975: 92-3), may also divine (94), and may have special functions like neutralising witchcraft (2.5 below) or operating an oath-verifying ordeal, rather than being primarily responsible for a specific piece of land. Conversely there are many places in Mampurugu which do not belong to any earth-'parish' jurisdiction.

2.2.3 Family Sacrifices

The ancestor sacrifices *par excellence* are those offered at clan or family shrines. These are 'public' in the sense that they concern all those in the relevant genealogical grouping; they are 'private' in the sense that there is a defined group and outsiders have no part in them. The senior male of the relevant group is the custodian and person responsible for the cult, but it is not his property, and he is under pressure to maintain the ritual correctly, both from family members and from society, especially manifested through the diviners and the ill-luck which leads people to consult them ('society' includes both the living and the dead, the latter being specially interested and potent in this area of life). The sacrifice is made to particular (deceased) ancestors: there is usually no lexical difference between "ancestor" and "grandparent":

Vocabulary #8

- a) GRANDPARENT, ANCESTOR (1): DB **yaba** / **yanima**; MP **yaaba** / **yaadima**; HG **yaa(ba)** / **yaadima**; KR (P) **yargo** / **yar-godima**; KLA **yaab** / **yaanam**; TL **yaab** / **yaanam**; NB (R) **yab(a)**; FR **sɔyaaba** / **sɔyaabdoma**; MR **yaba** / **yabramba**;
- b) GRANDPARENT, ANCESTOR (2): FR **sɔyaaba** / **sɔyaabdoma**; NT **cakotuwo** / **cakoto**; LB (R) **sakum**; DR (R) **sankuma**, (G) **sāākum** / **sāākumine**; WL (R) **sankuma**;
- c) GRANDPARENT, ANCESTOR (3): KLA **ba'ar** / **bada** "ancestral spirits" (S); NT **cakotuwo** / **cakoto**; DR (G) **kpiin** ("the dead"); BL **kokpieng** / **kokpiengsa**, /**koba**, **chue**, **ngiak** / **ngaasa**, **kpilima** ("the dead")

The rendering (Arozenara: 1975) '*Mânes des Ancêtres*' — "ancestral spirits, shades" reads too much connotation into the term. No word which would be rendered "spirit" is used in the local languages for any of these phenomena. Note Kröger, 1982: 6:

... we must find out to which aspect of the ancestor's personality a sacrifice is offered. It is not offered to his body, nor to his soul ... All sacrifices are offered to the ancestor's *wen* which has never been inside a living person.

The negative part of this is true throughout the area; the positive idea that the sacrifices are to the 'god' (2.3.1 below) is, as far as I know, only BL. The importation of terms like 'spirit' ('ghost', 'shade' ...) into this discourse reflects the Western dichotomy between the living and the dead, the material and the immaterial, the natural and the supernatural. A speaker of one of these languages simply says, for instance, "I am going to give water to my grandfather", and there is no linguistic distinction between giving a living old man a drink and pouring a libation to a deceased ancestor.

Although the words spoken during the rituals are directed to ancestors, one normally talks about "sacrificing to / 'fixing up' the *shrine*". Thus if a Mamprusi says *N ya maali m ba'ari*, this can mean "I am going to sacrifice to my ancestors" or "I am going to repair (the physical structure of) my ancestor-shrine". I witnessed the *maali* -ing of a *ba'ari* where the sacrifice was made in the middle of the entrance yard of the Paramount Chief's palace, where there was no form of physical 'shrine' or 'altar' at all: by eighteen hours later the blood, feathers, and millet flour placed on the ground as a gift for the ancestors had been trampled and scattered so that no sign of the sacrifice remained. Most commonly, however, the shrine is — or is contained in¹⁶ — a hump, dome, cylinder or column of clay or gravel. In addition to its use for 'shrines' or 'altars', this same physical form of construction may be used to make a place for sitting or putting things on (a low, domed, 'earth-bench' and a cylindrical clay 'occasional table' are usually to the right and left of the doorway, just inside, as you enter a Mamprusi woman's room — the door is hinged on the left so that it reveals the seat for you to sit and greet the occupant, while concealing the property on the 'table'). It may also be used to form a larger 'dais' for a chief or elder to sit on:

Vocabulary #9

EARTH-BENCH, DAIS, ALTAR: DB *bimbini* / *bimbina*; MP *bimbinni* / *bimbina*; KLA *bimbim* / *bimbima*; TL *bimbimər*; FR *bimbine* / *bimbina*;

However, the term for "shrine" is usually used for the physical structure as well as the ritual function: it may be noted in passing that in general, while not in use, the shrine is not specially 'holy', there is no sacrilege involved in touching it, and so on¹⁷. The MP shrines are often minimal — either nothing, or a low hump; BL mediumhigh, often cylindrical or domed (also TL, KL) or conical (TL): the FR pillar-shaped, tall cones. While anthropologists often leap to assigning such shapes 'phallic' significance, I am not aware of any such associations being expressed by members of the societies. Other physical forms include clay pots (BL, FR, KL, TL), and pots, stones, etc. supported by a branching stick (0.5-1.0 mtr. high) — MP, TL. While in MP and many of the other languages there is a different term for each functional type of shrine, the BL word **bogluk** seems to have been generalized to function very much as I am using 'shrine' — you can have, for instance, a **wen-bogluk** ("god'-shrine" — MP ***wun-bugri** --**ba'ari** would be impossible), a **tiim-bogluk** ("medicine'-shrine" — see 2.4 below) and so on. Similarly in FR **bagre** can be "anything you sacrifice to (including a talisman or drum)" (Nancy Schaefer, p.c.).

Vocabulary #10

- a) FAMILY SHRINE (1): MP **ba'ari** / **ba'a**; HG **begiri**; KLA **ba'ar** / **ba'a**; TL (GG) **ba'ar** / **ba'a**, (F) **bayer** / **ba'a** ("the generic term for any ancestor or medicine shrine" 1945: 52); NB **ba'ar** / **ba'a**; FR **bagre** / **baga**, **bakologo** / **bakon**; DR **bagre**¹⁸; LW (SOL) **bɔɔre**¹⁹ (cf. BL **ma-bage** of an ancestress: "bage is always a sacred horn" Kröger 1982: 2);
- b) FAMILY SHRINE (2): DB (FS) **bugule** / **buga**; KLA **bugur** ("you go and get it from your mother's people" — S. Agumah p.c); TL **bɔɔar** / **bɔga**; DR **bɔɔre**; BL **bogluk** / **bogluta**²⁰;

Only the **Daɔaba** seem to have human effigies (**kpiin-**, **sāā-daa** "ancestors' stick") of the ancestors (SOL: 88). The existence of secret societies is only noted, in W.O/V, for DR and MR, masking only in MR — in northern Ghana the only other groups with masking are the DR-contiguous Grusi groups **Tampulma** and **Vagla**; note also that the sexual union of sky and earth is a DR-MR-FR subgroup concept, 2.2.2 above. Shrines for family ancestor cults seem to be located at the interface between the house interior and the outside world, just within, or on the outside flanking, the gateway of the rooms-and-courtyard house of the eastern area (Naden [ed.] 1976: 35, and Map 2.c, 33), in the cattle-room or 'byre' which is the only room in the western fortress-style house which communicates with the outside (J. Goody SOI: passim; plan p. 39 — note how the "medicine shrines" room connects to the interior of the house, the 'byre' outwards) — normal access being by ladder over wall or roof.

2.3 Sacrifice: Personal Cults

2.3.1 The sacrifices so far considered have concerned groups — people of a given 'Earth Parish', or all members of a clan, family, or other genealogically-defined segment. There are also shrines which are the personal concern of an individual. One of these is the 'god' -*WIN / WUN / YIN: "each person's Yin belongs uniquely to that person. It is a symbol of his individuality" ... "his personal Destiny" (Fortes, 1949: 227). Only in BL is the 'god' of an ancestor described as continuing to receive sacrifices from descendants (quote, 2.2.3 above), **bogluk** and **wen** seeming to be used almost indiscriminately in the naming of BL shrines (Kröger 1982: *passim*, e.g. pp. 16ff.).

The Balsa **wen** (a stone) comes from the sun to every (?) person, "usually during childhood"; amongst the Mamprusi only some people have **wuna**, which are clay pots or other objects, made for them in infancy. In contrast, the TL **yin** comes, normally, from the bush, and usually to a young man, in the form of an animal he hunts, or a found object, often a piece of metal, perhaps a tool or weapon. The diviner reveals that there is a specific group of ancestors who "demand that he should accept them as his Yin and sacrifice to them on this shrine" (1949: 228); in KL the 'god' may derive either from the father's or the mother's clan and seems to function mainly in terms of preventing the person from doing something important (e.g. marriage) until it is propitiated (S. Agumah, p.c.). Fisch (1913) describes DB **wune** as "object of idolatrous worship in the house, made of clay: household idol"¹⁵ (but see in 2.3.3 below). The vocabulary, originally in Lexinote #3 'Sun and God', is updated as:

Vocabulary #11

- a) GOD: DB **wuni / wuna**; MP **wunni / wuna**; HG **wunni / wuna**; KLA **win / wina**; TL **yin**; NB **yin / yina**; FR **yeni / yena**; MR **wendi, winde**; NT **wenne**; DG **wen**; LW **wen**; SF "God" (Naa)**wuni**; BL **won, wen**; BM **yenu**;
- b) GOD (2): MR (R) **nwendi**; DG **mwin**; DR **ɲmin**; WL **ngmini**; M **ɲmin**;

2.3.2 Spirit-Guardian

A very similar concept seems to be that of 'spirit-guardian'. This tends to be defined as something like "any 'religious' power which takes a particular interest in an individual". In LoWiili this "tutelary or guardian spirit" may be "special guardianship of the ancestors" for an individual or clan, or may be the Earth or any of the other shrines: it is usually associated with a special place

(the tutelary of one section 'is' a baobab tree) — SOL: 63, 50, 99, 76. The BL guardian spirit always has a man-made shrine: usually it is one of the household gods (e.g. an ancestor, a **jadok**, a medicine, or a **tenggbani**) — Franz Kröger p.c. It is the father who entrusts the child to its spirit-guardian which is often his (the father's) **Yin** — TL (Fortes 1949: 165, 212, 229), and in KL they name the child after it and then he has to sacrifice to it, "it can be from his grandfather, so it would be a stone", it may be in a room or 'by' a room in a **bimbim** (Vocabulary #9).

Vocabulary #12

SPIRIT-GUARDIAN: DB **sugri / suga**; MP **sigri / siga**; KLA **sigir / siga**; KLT (R) **segri / sega**; TL **seyər**; NK (R) **segere / sega**; MR (**sig** "pierres de sacrifice" — Canu); LW (J. Goody) **sigra**; LB (R) **sera / sega**; DR (R) **segera**, (G) **sigra**; DR **siura**²¹ ("LoDagaba" — J. Goody); BL **segi / sega**;

Unfortunately this item does not appear on the wordlists of the linguists: most of the attestations not from the well-founded anthropological studies of Fortes and J. Goody are from the dubiously reliable Rattray. In DB and MP the term is used in personal names — (Na-)Sugri (Na-)Sigri; for DB this is said to be connected with identification with a reincarnated ancestor, and this may be also the case for MP though I was told that the etymology is rather from **sigri / siga** "first rains". The MP **wunni** (Wunni also used as a personal name, like Sugri it may be used for man or woman) fills more of the function attributed to the *SIG-RI elsewhere.

2.3.3 Personal shrine

In MP the family **ba'ari** (2.2.3 Vocabulary #10 above) is clearly distinct from the **bugri**, the concern of an individual, which he may obtain for himself, or else it may be inherited, or obtained for him / her at birth or in childhood (the details of this need checking) especially if his / her sickness or the death of previous children should cast doubt on his / her viability (the child, of either sex, may then be named Bugri, and in that case may receive a 'proper' name later, at a safe age like 9-12 years). One significant objective criterion of the distinction in MP is that sheep are the only 'small cattle' acceptable as a sacrifice to a **ba'ari**, while the **bugri** receives goats. Thus Drucker Brown quotes the Gambaraana as sacrificing *goats* and fowls "and make **ba-yu-bug-shia**" (RAMC: 94): "make" here is presumably **maali** "sacrifice (to)" (Vocabulary #5.c above); **bayubugri** would be a bugri of the name(s) of father(s)²² — only a noun like **bugri** can be compounded (in its root form) with the indefinite **-seaa** ('shia'), not the verb

m-bugi "to divine" (Vocabulary #1 above) as implied by Drucker Brown's translation "sacrifice to the ancestors indicated by divination".

In its individual nature the MP **bugri** seems to share some of the functions of the *WEN and the *SIG (2.3.1-2 above). It is also used of a few 'public' shrines, such as that at Bugiya 8 km ESE of Walewale which enforces oaths and is particularly used in finding stolen property, and its oath as an ordeal for suspected thieves: this usage seems comparable to Fortes' "External **bɔyar**". The KL **bugir** may particularly be something you "go and get from your mother's people" (Agumah p.c.) for yourself or your child. The clear lines drawn in MP do not seem to be general for the group: the root, presumably, appears in "to divine" (Vocabulary #1) and in some words for "family shrine" (#10.b). The BL cognate is, as we have seen, a general term for 'shrine', and the TL one refers to a public or clan shrine. Fisch, for DB glosses "divinity (*Gotttheit*), a spirit worshipped in some kind of object" (op.cit.¹⁵) and gives quotations that the **tindana** looks after the **bugule**, it is a stone or trees, at night he walks about, he kills people, he lives in the bush, he is a 'fairy' (2.5 below). This sounds confused: Christian missionary Fisch with Muslim language-helpers (he adds in this entry "My informant's statement is influenced by Islam") tends to give a jaundiced view of traditional religion as already seen in 2.3.1 (a DB quotation in the entry for **tingbandana** includes the words "he is a bad man"): however it may show a generalisation of the term to include the 'grove' and other shrines (or show that the MP specialisation is a recent innovation).

Vocabulary #13

SHRINE: DB **bugli** / **buga**²³; NN (Sk) **bogole** (**tindaana** sacrifices to it); MP **bugri** / **buga**; HG **bugiri** / **buga**; KLA **bugir** / **buga** (matrilineal); TL **bɔyar** / **bɔya**²³; NB **bugur** / **buga**, **wook** / **woot**²⁴; FR **bugure**²³ ("shrine-spirit"), **bagare** / **baga**²³; MR **bugli**; DR (H) **bɔg**, **bɔgr** / **bɔgɔ**; BL **bogluk** / **bogluta**²³; KM **bo-** / **buguli**, **buguti**;

2.4 Acquired 'Power'

This division may seem arbitrary, but there appears certainly to be a clear distinction between all the foregoing shrines and those to which we come now. The former are treated, as it were, as existing in their own right and laying the obligation to cultivate them upon the person / group concerned (cf., e.g., Kröger "if the **tengbain** spirit has no **bogluk**, i.e. if he does not receive any sacrifices, he may cause harm to human beings" (1982: 85); and Fortes on the TL Yin quoted in 2.3.1 above). The latter look towards powers or forces which individuals set out to acquire in order to fend off existing or feared misfortune, to

gain strength for a venture, or to attack an enemy. Normally the acquisition of these objects involves going to a practitioner and purchasing them, often as a straight commercial transaction, or by providing an animal whose slaughter may in some sense be a sacrifice and a source of power, but of which the practitioner keeps the meat as (part of) his fee, and parts of the animal — horns, bones, skin, blood ... — may be used in making the object. The general term in this area is 'medicine' — a very ancient lexical root probably related to "tree"²⁵.

Vocabulary #14

- a) MEDICINE: DB tiim / tiima; MP tiim / tiima; HG tiim / tiima; KR (N) tiim / tiima, (P) tinse "un remède"; KLA tiim / tiima; NB tiim / tiima; FR tiim; MR tiim / tiito, -tibo / -tido; NT tiem; LW (G) fiil, tiib / tibe ("medicine shrine"), -tiib / -tibe, -fiil; BR (K) satēe ("medicine to treat someone struck by lightning"); SF tiim; BL tiim / tiita; KM tiim / tiiti;

The expert who has the power to make and dispense medicines of this type is a "medicine-man" (lit. 'owner')

- b) MEDICINE-MAN: DB (N) tiimalana, (FS) timlaana / timlaanema; KLA tiimadaan; MP tiindaana, tiimadana; MR tiimsoba / tiimdamba; BL tiimnyono / tiimnyam; BR (K) gyeborɔ;

As the translation suggests, a frequent reason for seeking such power is to cure a disease, and the verb "to heal, cure" is related:

- c) HEAL, CURE: DB / MP / HG tibi; KLA ti'eb; KLT tieb; FR tebe; BL tebi (TL tiba, FR teba "medicine-man")

As something purchased from an expert, *TII-M covers both 'magic' medicines and modern, pharmaceutical medicines. The modern practitioner is usually distinguished by being described by a form from English "doctor". There is additionally, certainly in some of the languages, a term for herbal medicines, which anyone can go to the bush and find and prepare for themselves. The form is usually etymologically "tree-roots"²⁶, though in fact it is more often bark and leaves which are used:

Vocabulary #15

- HERBAL MEDICINE: MP danya'a; KLA danya'a; FR dɔn'ɔnrɔ; DR daanyuur;

'Medicine' represents such a wide spectrum of powers, offensive and defensive, healing and protecting, that the term 'overflows' to other parts of the field. Fisch quotes statements that the **tindana** has much medicine (but he also confuses **tindana** and **timlana** etymologically), the Muslims also have a little, and so do witches (2.5 below). Jack Goody speaks of a clan which "owns the rain because of its powerful rain medicine" (SOL: 80), hunting medicines (including **wefii** — "bush (hunters') medicine", see Lexinote #8), newhouse and clan shrines (**yipaala tiib**, **ḍḍḍḍ tiib**). 'Medicines' are often contained in pots (MR **tiimruko** "jarre sacrée contenant des produits mystérieux": BL (**tiim**)-**bogluk** "often a clay pot filled with medicine") but are also wrapped in skins, ropes and / or pieces of paper / plastic and hung up in or on rooms or granaries, or as skulls, horns or sticks of various forms thrust into the thatch or the space between the thatch and the top of the wall. The word is also used for modern "chemicals", especially insecticides but also the chemicals used in photographic processing and similar procedures.

Perhaps to be treated as a major subcategory of 'medicines' are talismans and amulets. These are certainly etically and objectively distinguishable as they are carried or worn about the person — on necklaces, bracelets, rings or attached to clothing — rather than being in a fixed location. All the other objects mentioned are installed in a specific place, though they may occasionally be transferred, for instance when a house is abandoned and built elsewhere; there do not seem to be any ritual objects which are periodically moved as part of their use — spending part of the year in one place and part in another, for instance, or being taken in procession as part of the rite (except, perhaps, the spears and other objects, including the 'King' — RAMK: 156 foot, used in MP etc. kingship rituals — RAMK 58-59, 96). These and other medicines may have many different specific names according to their particular purpose, manner of acting, or ingredients (e.g. FR **iila** "horns", **ligiri** "cowries" etc.; MR **tiimnugbinsi** "magic rings"). General terms are not widely recorded — information please!

Vocabulary #16

AMULET, TALISMAN: DB **sabli / saba**; MP **sabri, sabli / saba**; FR **bagri / бага**; MR **sebere / seba**; WL **/seba**; BL **bage / бага, poali / poala**;

In general a medicine or amulet does not receive any cult; having been acquired it just does its work. However, anything of special significance may receive a sacrifice — musical instruments (MR — Arozenara; BR Steve Batstone, p.c.) especially drums (FR Nancy Schaefer, p.c.) and talismans (FR *ibid.*), medicines, blacksmith tools, and so on. Some medicines may thus be 'topped up' to restore

their power, but many dusty, cobweb-covered 'medicines' hang in rooms from year to year without any apparent attention being paid to them.

2.5. 'Fairies', Witches

Two other major concepts seem to outsiders to be relevant to this semantic area, though they may not be emically 'the same sort of thing' to local members. The first is labelled in the local English by the English term 'fairy', and this is a convenient label as the alternatives are cumbersome and do not fit the variety of concepts which are attached to what is evidently a cognate set of terms in the languages. The other is universally labelled with the English term 'witchcraft' though the African concepts are at least as far distant from any original English sense of the word as in the case of 'fairy'.

2.5.1 'Fairy'

The personages usually described in Ghana English as 'fairies' are in most cases mischievous bush-sprites — little men with red hair and beards, with feet on back-to-front so that their footprints mislead trackers. They encounter lone people in the bush and disorient them so that they are not found until some time later, and maybe some distance away, dead, mad or temporarily dazed and amnesiac. However in LoWiili they are "hill and river sprites" (SOL: 38: also "bearded dwarfs" p. 91) who may receive sacrifices and may be an individual's spirit-guardian (2.3.2 above). In MR they may also receive sacrifice (Some 1970: 10) and guide one form of divination (2.1 above). Mendonsa (1973, 1982) mentions a new and marginal 'fairy'-based divination for the Sisaala, a Grusi group living between the Balsa and the Dagaari peoples. Fisch was told by his informant of the *bugele* that was cared for by a *tindana* that "he is in the bush, he is a fairy" (translated by Fisch as 'a Devil, evil being'). The word seems particularly prone to phonological variations while remaining recognisably the same lexical item:

Vocabulary #17

FAIRY: DB *kulkparga* / *kulkparsi*, (N) *kpilkparga* / *kpilkparsi*;
 MP (east) *ku(r)kpaariga* / *-kpaarisi*, (west) *kpikpaariga*;
 HG *kikpaarigu* / *kikpaarigusi*; KLA *kikirig* / *kikiris*;
 TL *kalkpaareg* / *kalkpaaris*; FR *kenkirego* / *kenkito*;
 NK (R) *kyinkyirega*; MR *kinkiriga* / *kinkirisi*;
 LB (R) *kɔntɔn*; DR (K) *kɔntɔn*; LW (G) *kɔntɔmo* /
kɔntome; BL *kikiruk* / *kiki(r)ta*, *kikerik* / *kikerisa*;

2.5.2 Witchcraft

Witchcraft is a sphere which is certainly not emically linked to the other notions: only to outsiders does it seem to belong in the same category, as involving powers and happenings which are not explained in his materialistic terms and are regarded by him as 'supernatural' or 'spiritual' (and therefore unreal).

When someone unexpectedly gets sick or dies, one explanation which may be given by *divination*²⁷ is that witchcraft is responsible. The diviner or public opinion points to somebody as the witch — usually a woman, often an elderly relative or household-member: if the suspect does not confess, there may be a witch-finder who can 'see' and indicate the culprit. The witch confesses and / or is 'shamed' by being beaten and howled at (MP *n-wooli* — "to go 'Woo, woo!' at somebody") by a circle of villagers. The witch is held to have 'chewed' the 'spirits' (MP *siiya / siisi*) of people, trapping them with some apparatus or 'flying' to batten on them at night, and the victims though outwardly unaffected may then sicken and die.

The witch is usually compelled by ostracism if not force to leave the house and village. In MP the chief of Gambaga has 'medicine' which neutralizes witchcraft, so he has a quarter in his town which forms a 'city of refuge' where accused witches can live unmolested. The whole preceding account is MP, but probably the same is true in outline for the whole area (cf. the essentially similar account for TL — Fortes 1949: 32ff.). Though Fisch quotes "a witch has *tim* and casts enchantments" there is no evidence that anyone actually consciously practises witchcraft (in contrast to, e.g., the Tiv), and it is sometimes treated more as a dangerous disease which infects the witch than a culpable offence.

The form is *SO-, in Noun Class I for the witch as a person; I cannot explain the nasal which is present between the root and suffix in some languages, particularly in DB and MP where it is only in the singular:

Vocabulary #18

WITCH: DB *sɔnya / sɔba*; MP *sɔnya / sooba*; HG *sonye / sonyesi*; KLA *soenn / soonb*; TL *sɔi / soob*; FR *sɔnya / sɔmba* (R) *soa / soba*; MR *swoya / sweeba*; LW *sɔɔr*; BR (K) *sɔɔba* (? = pl. ?); BL *sakpak*, (lam-)yiik / *yiisa, nyaarɔa / nyaarɔaba* "seer";

Fortes connects with *soog* "uterine kin" (TL, FR) and says that it is inherited matrilineally, which is not the case in MP, at least.

2.6 Conclusion

Note the two roots cropping up throughout this semantic field:

*BAG: Vocabularies 1, 2, 3.a, 4, 5.b, 8.c, 10.a, 13, 16

*BUG: Vocabularies 1, 3.b, 4, 5.b, 10.b, 13

Notes

- 1 The Lexinotes project is described in Naden 1993.
- 2 It might be urged that the established ethnographic practice of concentrating on 'traditional' cultures is both patronising and irrelevant in the modern world. However, most people outside major cities are living nowadays in a cultural continuum which blends the traditional and the 'modern', and it is valuable to have analyses of the traditional component in order to study the blends, this being the more difficult both to isolate and to understand.
- 3 The whole dictionary was published in the same year as the translation of Otto, though the "Religion" article was in a fascicle published in 1906.
- 4 This is clearly consistent with the finding that "Bantu terms used to refer to ancestors were identical to those used to denote living elders" (Kopytoff, 1971 apud McCall 1995: 256).
- 5 With regard to ritual matters, at least; for secrecy in historical traditions see Davis 1992: 67-8.
- 6 The lengthy cosmogonic and cosmological text from the Bagre, and its related ritual, recorded by J. Goody (MBG — see References) is very untypical for the north of Ghana, let alone the Oti / Volta groups; one wonders whether it is evidence for a westward-looking stance, influenced by Lobi and Senufo peoples.
- 7 On cultural specificity of the function of questioning see Esther Goody, 1978. In Mampruli and its congeners "Why do / did you...?" is functionally 'Why the Hell do / did you...?' — *i.e.* a rebuke rather than a request for information.
- 8 The sweet yellow powder surrounding the seeds in the pods of the *Parkia* tree (*P. biglobosa* or *P. Clappertonia*), a species of which many parts are put to various economic uses (French *nééré*).
- 9 This section is largely based on Lexinotes (see Naden 1993) numbers 3 "Sun and God" and 14 "Traditional Religion". The Lexinotes #3 ("god / God") and 7 (section on "earth-priest", "grove") sparked off an epistolary discussion between myself and (particularly) Franz Kröger as to how far one can rightly speak of "deities", "spirits" or "supernatural beings" (rather than "(non-material) powers, forces") in connection with many of these areas. Some of the discussion and data used here arises from this.
- 10 Now see also Barker 1986.
- 11 Vocabularies re-edited June 1993.
- 12 Wumbillah (n/d) notes for Kusaas that a cowri that falls mouth-up is said to 'smile', and combinations of smiling and non-smiling outcomes in the fall of three or four cowris have significance as responses to the diviner's or client's questions.
- 13 But see Vocabulary #3.b entries.
- 14 Or greet *Ni ti* ... "For *our* ..." if the greeter was also involved.

- 15 My translation.
- 16 'Contained in' in cases where the actual sacred object is a stone or piece of iron but is incorporated in the 'altar'.
- 17 Naden [ed.] 1985, Kröger 1982: 8.
- 18 Also glossed "sacrifice" (113) and "initiation rites of a secret society" (52).
- 19 J. Goody identifies *bɔɔre* and *bagre* (SOL: 52); the dialect of his LoWiili shares the feature that "the voicedness of /b/ and even of /g/ undergoes regressive assimilation to a preceding /a/"¹⁵ as in BR and in the DG of Pénou-Achille from whom the quote comes: it gives its grave character to yield a back / round vowel — Pénou-Achille's examples {dab+lu} -> /dɔɔlu/ "manliness", {pog+θ} -> /pɔw/ "woman".
- 20 The -l- in BL may be a thematic extensor consonant (Manessy 1975: 133f.) or a second morpheme: the latter could explain the change of class from II to IV, but note that in "fire" *BUG—*LAM with the second member of the compound attested in the E.O/V languages (Bieri (P) *daam*, (Dp) *lam*) yields BL *bolem* with loss of the -g-.
- 21 *-ŋ- -> -iu- see note 19.
- 22 Or is it *ba'ayubugri*? — *ba'ayuya* being used for "divinatory" (Vocabulary #2 above), *i.e.* "(sacrifice to his) divinatory shrine".
- 23 Also more general reference.
- 24 Could this be *WO-, *WE-, or *YO-KU/-TI — 'bush' (hunting) shrine?
- 25 Under PWN *TI Mukarovsky (1977) lists both **ti-ka* "tree" forms and **ti-m* "medicine" words: Common Bantu *-ti also yields both senses (presumably with different class prefixes).
- 26 I therefore suspect the separateness of the two Grusi reconstructions "root" *LI/U and "medicine" LI/U(D) in Manessy 1969.
- 27 In many groups the funeral rites incorporate some ceremony to divine whether the death was caused by witchcraft.

References

A Sources of vocabulary in Part 2

- (J) Lists in JWAL article, Swadesh et al. 1966
- (R) Lists in Rattray, 1932
- (N) personal research (Naden)

- DB Dagbani: printed sources, (FS) Fisch 1913, (N), (OP) Oppong, (O) Okraku
- NN Nanun: personal research, (Sk) Skalniņ 1983
- MP Mampruli: personal research, Drucker Brown RAMK
- HG Hanga: printed sources, (H) p.c. Geoffrey Hunt
- KR KaMara: Bouna dialect Prost 1974b (P); Ghana dialect (N), (H) p.c. Geoffrey Hunt
- KLA Kusaal, Agole (Eastern) dialect: printed sources, (N)
- KLT Kusaal, Tɔɔndi (Western) dialect: notes by David and Nancy Spratt (S), (N)
- NB Nabit: (N), (R), (J)

- TL Talni: Fortes 1945, 1949 (F), (N), written list by Abraham Manga for Naden (A), (GG) two lady informants recorded by Naden
- FR Frafra: written sources, p.c. Bob / Nancy Schaefer
- GN Gurenne: (J) — origin unrecorded
- NK Nankan: (N), (R) (may not be same dialect)
- MR Moore: printed sources, Hall n/d (1948) (H), Canu, Tiendrebeogo / Pagéard
- NT Nootre: Prost 1971, 1974a, 1975
- DG Dagara: Burkina Faso (J), Pénou-Achille
- DR Dagaari: (J), Ed Hall notes / files (H), (G) J. Goody — (SOL, MBG)
- LW LoWiili: (G) J. Goody — (SOL, MBG)
- NR Nura (Dagaari of Lawra): (J), Kennedy 1966 (K)
- LB Lobiri(? = NR): (R)
- BR Birifor: (J), (K) Larry Kuch (p.c.)
- WL Waali: GILLBT files
- SF Safalaba: (J), (N)
- BL Buli: (FK) Franz Kröger — (AW) 1982, (D) 1992, GILLBT files (Ian Gray, Todd Poulter, Paul Dancy), (N), Melançon / Prost 1972 (P)
- KM Kɔnni: (N), Naden 1986, Cahill
- BM Bimoba: published sources, Norman Price p.c., GILLBT files (Jill Jacobs, Sonia Hine)
- PWN Proto Western Nigritic: reconstructions of Mukarovsky 1977, 1976
- W.O/V Western Oti / Volta: reconstructions and summary forms of Manessy 1975, Naden 1989, Lexinotes, personal research.

B Works Cited

- Abasi, Augustine Kututera. 1995. LUA-LIA, the "Fresh Funeral": Founding a House for the Deceased among the Kasena of North-east Ghana. *Africa* 65(3): 448-475.
- Arozenara, Pierre. 1975. Notes à propos de quelques instruments de musique mossi. *Ouagadougou. Notes et Documents Voltaïques* 3: 39-53.
- Barker, Peter. 1986. *Peoples, Languages and Religion in Northern Ghana — A Preliminary Report*. Accra: Asempa / G.E.C.
- BIBLE (MR). 1983. *Wɛnnaam Sebre ne Moore: La Bible en Mooré*. Lomé: Société Biblique en Afrique Occidentale.
- Bourdillon, M.F.C. and M. Fortes (Eds.). 1980. *Sacrifice*. London, N.Y.C.: Academic Press.
- Brown, Cecil H. 1984. *Language and Living Things: Uniformities in Folk Classification and Naming*. New Brunswick, NJ: Rutgers University Press.
- Brown, Cecil H. 1986. *The Structure of Folk Biological Categories and the Perceptual Basis of Life-form Classes*. Manuscript.
- Cahill, Mike. 1993. What! No Prepositions? Cahill, M.: (Ed.). 1993: 7-9.
- Cahill, Mike. 1993. Vowel Harmony and Underspecification in Kɔnni. Cahill, M. (Ed.). 1993: 10-17.
- Cahill, Mike (Ed.). 1993. *Papers from G.I.L.L.B.T.'s Seminar Week: 1-5 February 1993*. Tamale, N.R.: G.I.L.L.B.T.

- Canu, Gaston.** 1968/71. Gùrènnè et Mò:rè. *WALS / SLAO* 1968/71: 265-283.
- Davis, David C.** 1952. "They Sing Our Origins": A Study of the Lungsi Drummers of Mampuru. *African Music* 7(2): 58-71.
- Deschamps, Hubert.** 1972. *Les religions de l'Afrique noire*. Paris: PUF.
- Drucker Brown, Susan.** 1975. *Ritual Aspects of Mamprusi Kingship (RAMK)*. Leiden: Afrika-Studiecentrum / Cambridge: African Studies Centre.
- Fisch, Rudolf.** 1913. Wörtersammlung Dagbané-Deutsch. *MSOS* 16: 113-214.
- Fortes, Meyer.** 1945. *The Dynamics of Clanship among the Tallensi*. London: OUP.
- Fortes, Meyer.** 1949. *The Web of Kinship among the Tallensi*. London: OUP.
- Fortes, Meyer.** 1980. Preface: Anthropologists and Theologians: Common Interests and Divergent Approaches. Bourdillon, M.F.C. and M. Fortes (Eds.). 1980: v-xix.
- Goody, Esther N.** 1978. *Towards a Theory of Questions*. E. Goody (Ed.). 1978: 17-43.
- Goody, Esther N. (Ed.).** 1979. *Questions and Politeness: Strategies in Social Interaction*. Cambridge, Cambs.: CUP.
- Goody, Jack.** 1956. Second edition. 1961. *Social Organisation of the LoWiili (SOL)*. London: OUP for IAL.
- Goody, Jack.** 1972. *The Myth of the Bagre (MBG)*. London, Oxford: OUP.
- Hall, John F. s.a. (1948).** *Dictionary and Practical Notes, Mossi-English Languages*. Ouagadougou: Mission des Assemblées de Dieu.
- Hunt, Geoffrey R. and Rosemary Hunt.** 1981. *A Phonology of the Hanga Language*. Legon: IAS.
- Jacobs, Gillian G.** 1970. *Bimoba Syntax*. Legon: IAS.
- Kennedy, Jack.** 1966. *Phonology of Dagaari*. Legon: IAS.
- Kopytoff, Igor.** 1971. Ancestors as Elders in Africa. *Africa* 41(2): 129-142.
- Kröger, Franz.** 1982. *Ancestor Worship among the Balsa of Northern Ghana*. Höhenschäftlarn: Klaus Renner.
- Kröger, Franz.** 1992. *Buli-English Dictionary*. Münster, Hamburg: Lit.
- Kuper, Adam.** 1982. Lineage Theory: A Critical Retrospect. *Annual Reviews of Anthropology* 1: 71-95.
- Manessy, Gabriel.** 1969. *Les langues gurunsi* (2 vols.). Paris: SELAF.
- Manessy, Gabriel.** 1975. *Les langues Oti / Volta*. Paris: SELAF.
- MBG see Goody, J. 1972.
- Mbiti, John S.** 1969. *African Religions and Philosophy*. London: SPCK.
- Mbiti, John S.** 1970. *Concepts of God in Africa*. London: SPCK.
- McCall, John C.** 1995. Rethinking Ancestors in Africa. *Africa* 65(2): 256-270.
- Melançon, Lucien and André Prost (R.PP.).** 1972. *Dictionnaire buli-français*. Dakar: Univ. de Dakar.
- Mendonsa, Eugene L.** 1973. *Divination Among the Sisala of Northern Ghana*. Unpublished PhD dissertation, Univ. of Cambridge.
- Mendonsa, Eugene L.** 1976. Elders, Office Holders and Ancestors among the Sisala of Northern Ghana. *Africa* 46: 57-65.
- Mendonsa, Eugene L.** 1982. *Politics and Divination*. Berkeley: Univ. California.
- Mukarovsky, Hans.** 1976. *A Study of Western Nigritic II*. Wien: Institut für Ägyptologie und Afrikanistik der Universität Wien.

- Mukarovsky, Hans.** 1977. *A Study of Western Nigritic I.* Wien: Institut für Aegyptologie und Afrikanistik der Universität Wien.
- N&Q** see R.A.I. 1951.
- Naden, Anthony J.** 1982. Class Pronoun Desuetude Revisited. *JWAL* XII(1): 34-42.
- Naden, Tony (Ed.).** 1976. *Kinship Terminology: and Some of the Social Correlates or Outworkings of the Kinship System in Ghanaian Culture.* Legon: IAS.
- Naden, Tony (Ed.).** 1984. *Law in Some North-Ghanaian Societies.* Legon: IAS.
- Naden, Tony.** 1986. Social Context and Mampruli Greetings. Huttar, George and Kenneth Gregerson (Eds.). *Pragmatics in a Non-Western Perspective.* Dallas: S.I.L. / University of Texas at Arlington [SILPL #73]: 161-199.
- Naden, Tony.** 1988. Gur Languages. Dakubu, Mary Esther Kropp (Ed.). *The Languages of Ghana.* London: Kegan Paul International for IAI / WAL: 12-49.
- Naden, Tony.** 1989. Gur. Bendor-Samuel, J.T. (Ed.). 1989. *The Niger-Congo Languages.* Lanham, MD.: University Press of America: 141-168.
- Naden, Tony.** 1993. From Wordlist to Comparative Lexicography: The Lexinotes. *Lexikos* 3: 167-190.
- Naden, Tony.** Forthcoming. *Time and Calendar in Mampruli.* Naden (Ed.). Forthcoming: 25-42.
- Naden, Tony (Ed.).** Forthcoming *Time and the Calendar in Ghana.* Legon: IAS.
- Naden, Tony.** Manuscript. Notes on the KaMara of Larabanga.
- New Testament (KL).** 1976. *Wina'am Gbon Nonapaal: (New Testament) Kusaal.* Wycliffe Bible Translators / World Home Bible League.
- New Testament (HG).** 1983. *Naawun Bukupaali: The New Testament in Hanga.* International Bible Society / World Home Bible League.
- New Testament (FR).** 1986. *Na'ayene Gɔɔpaalom: Farefare.* International Bible Society.
- New Testament (DB).** 1982. *Naawuni Kundi Palli: New Testament in Dagbani.* Accra: Bible Society of Ghana.
- Newell, W.H. (Ed.).** 1976. *Ancestors.* The Hague: Mouton.
- Nkrumah, Sylvester.** 1993. *Concepts of the Supernatural Among the Akyode.* Cahill (Ed.). 1993: 45-64.
- OED.** 1933. *A New English Dictionary on Historical Principles.* London, Oxford: Clarendon, OUP.
- Okraaku, John S.** 1917. *Dagomba Grammar (with exercises and vocabularies).* Oxford: OUP.
- Oppong, Christine.** 1973. *Growing Up in Dagbon.* Tema: Ghana Publishing Corpn.
- Otto, Rudolf.** 1923 / 1950. *The Idea of the Holy.* London: OUP.
- Pénou-Achille, S.** 1982. *Systématique du signifiant en dagara.* Paris: L'Harmattan.
- Prost, André (R.P.).** 1971. Nôte et moore. *Notes et Documents Voltaïques* 5(1): 25-30.
- Prost, André (R.P.).** 1974a. Notes sur le kamara de Bouna. Abidjan: *Annales de l'Université d'Abidjan série H VII:* 75-93.
- Prost, André (R.P.).** 1974b. Les langues de l'Atakora: VII. Le nôtre, langue des Nôtèba dits Bulba. *BIFAN* 36, B(3): 628-659.
- R.A.I.** 1951¹⁶. *Notes and Queries on Anthropology.* London: RKP for Royal Anthropological Institute.
- RAMK** see **Drucker Brown.** 1975.
- Rapp, Eugen L.** 1966. *Die Gurenne-Sprache in Nordghana.* Leipzig: VEB Verlag Encyklopädie.
- Rattray, R.S.** 1932. *Tribes of the Ashanti Hinterland.* (2 Vols.). Oxford, London: OUP.

- Riggs, Fred W.** 1986. *The Interdisciplinary Tower of Babel*. Prepub. draft.
- Rosch, Elinor, Carolyn B. Mervis, Wayne Gray, David Johnson and Penny Boyes-Braem.** 1975. *Basic Objects in Natural Categories*. Berkeley: University of California at Berkeley.
- Sarpong, Peter.** 1974. *Ghana in Retrospect*. Accra: Ghana Publishing Corporation.
- Schaefer, Nancy and Patrick Adongo.** Forthcoming. Notes on the Frafra Time System. **Naden (Ed.)**. Forthc.: 19-24.
- Schaefer, Robert L.** 1975. *The Phonology of Frafra*. Legon: IAS.
- Skalnik, Peter.** 1983. Questioning the Concept of the State in Indigenous Africa. *Social Dynamics* 9(2): 11-28.
- SOL** see **J. Goody.** 1956 / 1961.
- Some, Bernard B.** 1970. La religion traditionnelle mossi comme source de valeurs de civilisation politique. *Notes et Documents Voltaïques* 4(1): 3-31.
- Spratt, David and Nancy Spratt.** s.a. a. *Kusaal-English Dictionary*. Tamale: Ghana Institute of Linguistics (now GILLBT).
- Spratt, David Nancy Spratt.** s.a. b. *An Introduction to Learning Kusaal*. Tamale: Ghana Institute of Linguistics. (now GILLBT).
- Spratt, David and Nancy Spratt.** 1968. *The Phonology of Kusaal*. Legon: I.A.S.
- Spratt, David and Nancy Spratt.** 1972. *Kusaal Syntax*. Legon: I.A.S.
- Swadesh, Mauricio, Evangelina Arana, John T. Bendor-Samuel and W.A.A. Wilson.** 1966. A Preliminary Glottochronology of Gur Languages. *J.W.A.L.* III (1): 27-65.
- Teviu, Adjewo and John C. Callow.** s.a. (1969?). *The Founding of Paga*. Legon: I.A.S.
- Tiendrebeogo, Yamba and Robert Pagéard.** 1975. Dictons mossi sur les devins. *Notes et Documents Voltaïques* 8(3): 37-8.
- Tolkien, John R.R.** 1964. *Tree and Leaf*. London: Allen and Unwin.
- Tuurey, G.** 1982. *An Introduction to the Mole-speaking Community*. Wa, UWR: Wa Catholic Press.
- WALS / SLAO.** 1969 / 1971. *Actes du huitième congrès international de linguistique africainen*. Abidjan: Annales de l'Univ. d'Abidjan, 1971, Série H: fasc. hors série.
- Wumbillah, Joseph B.** s.a. (1970s). *The Kusaas People of Northern Ghana*. Unpublished MS.
- Zwernemann, J.** 1961. *Les notions du dieu-ciel chez quelques tribus voltaïques*. *Etudes Volts.* 2: 71-100.

Die inkortingsneller *moet* en die woordeboek

A.F. Prinsloo, *Afrikaanse Taaladviseur, Radio Sonder Grense, SABC,
Auckland Park, Suid-Afrika*

Abstract: **The Morphological Reduction Trigger *Moet* and the Dictionary.** The Afrikaans auxiliary verb *moet* ("must") can, within certain contexts, act as a trigger for the morphological reduction of a compound main verb of the sentence without changing the semantic content of that sentence (a condition, in fact, for all morphological reduction). The circumstances under which this occurs are examined, and lexicographical evidence and implications are discussed.

Keywords: AGENT, AUXILIARY VERB, CARD COLLECTION, COMPOUND MAIN VERB, DAT-RESUMPTION, DETERMINER, LEXICOGRAPHY, MAIN VERB, NEGATION, NON-REDUCIBLE, REDUCED SENTENCE, REDUCIBLE, REDUCTION, REDUCTION RESULT, REDUCTION TO THE LEFT, REDUCTION TRIGGER, SENTENCE CONTEXT, SYNTAX, UTTERANCE SITUATION, VERB INVERSION

Opsomming: Die Afrikaanse hulpwerkwoord *moet* kan binne sekere kontekste as inkortingsneller ten opsigte van 'n saamgestelde hoofwerkwoord optree, waardeur laasgenoemde ingekort word sonder dat die semantiese inhoud van die sin verander ('n voorwaarde, trouens, van alle inkorting). Die omstandighede waaronder dit geskied, word ondersoek en leksikografiese getuigenis en implikasies word bespreek.

Slutelwoorde: AGENS, BEPALER, DAT-VERVATTING, HOOFWERKWOORD, HULPWERKWOORD, INGEKORTE SIN, INKORTBAAR, INKORTING, INKORTING NA LINKS, INKORTINGSNELLER, INKORTINGSRESULTAAT, KAARTKORPUS, LEKSIKOGRAFIE, NEGATIVERING, NIE-INKORTBAAR, SAAMGESTELDE HOOFWERKWOORD, SINKONTEKS, SINTAKSIS, UITINGSITUASIE, WERKWOORDINVERSIE

Die hulpwerkwoord *moet* vereis normaalweg die optrede van 'n hoofwerkwoord in enige gegewe sin, bv.

- (1) die seun *moet* die fiets *koop*

waar *moet* as hulpwerkwoord die optrede van 'n hoofwerkwoord soos *koop* verpligtend maak. Vgl.:

- (2) *die seun *moet* die fiets.

Die vorm waarin die hoofwerkwoord hom in die sin manifesteer, kan egter wissel, veral by saamgestelde werkwoorde. Die sin

- (3) die skoene *moet skoongemaak word*

is semanties gelykstaande aan

- (4) die skoene *moet skoon*.

In (4) neem die hoofwerkwoord *skoonmaak* die vorm aan van 'n adjektief maar nou in 'n verbale funksie.

1. Inkortingsnellers

In hierdie opsig kan *moet* bestempel word as 'n *inkortingsneller*, d.w.s. die teenwoordigheid daarvan veroorsaak dat die hoofwerkwoord inkorting kan ondergaan sonder dat die semantiese inhoud van die sin gewysig word — 'n voorwaarde vir suksesvolle inkorting. Dit blyk dat hierdie inherente inkortingsvermoë binne sekere kontekste in meerdere of mindere mate, buiten in *moet*, ook in die volgende Afrikaanse hulpwerkwoorde skuil: *kan, mag, sal*:

- (5) die glase *kan blink* (< blinkgemaak word)
(6) die deure *mag toe* (< toegemaak / -trek word)
(7) die veters *sál los* (< sál losgemaak word).

In die geval van (7) blyk dit dat *sal* as inkortingsneller (hierna afgekort IS) gereideliker optree as daar buitengewone klem in die sin is, wat nie die geval is by

- (8) die veters *moet los* (< losgemaak word)

nie.

Die hulpwerkwoord *wil* tree ook anders op:

- (9) die werk *wil *klaar* (< *klaargemaak word)

lewer 'n stervorm op.

Die rede hiervoor is dat *wil*, anders as *moet*, 'n spesifieke agens vereis, d.w.s. iemand of iets as onderwerp van die sin wat *wil* uitvoer wat in die hoofwerkwoord beskryf word. In die geval van *wil* word die inkortingsresultaat ook verkry uit 'n hoofwerkwoord wat in die bedrywende vorm is, en nie in die lydende, soos by die voorbeeldsinne vir *kan, mag* en *moet* nie. Vgl.:

- (10) Adrian *wil klaar* (< klaarmaak)

met

(10a) die werk **wil klaar* (< klaargemaak word).

Die inkorting met *moet* as hulpwerkwoord kan saam met 'n agens as onderwerp voorkom as die saamgestelde hoofwerkwoord in die ongekorte sin in die infinitiewe vorm is, en daar geen direkte voorwerp teenwoordig is nie:

(11a) *Kiwa moet klaar* (< klaarmaak)

maar

(11b) *Kiwa moet die werk *klaar*.

(11c) Die perde *moet uit* (< uitkom / uitloop)

maar nie

(11d) *Annelie moet die perde *uit*

nie.

Omdat leksikografiese getuienis vir hierdie soort inkorting in woordeboeke bykans totaal ontbreek, en die voorbeelde met *moet* die beste in die ondersoekte kaartkorpus van die Woordeboek van die Afrikaanse Taal (WAT) verteenwoordig word, word hoofsaaklik dié hulpwerkwoord se optrede eksemplaries behandel.

2. Die optrede van *moet* as inkortingsneller

(a) Morfologies of sintakties

Die inkorting van die hoofwerkwoord kan morfologies-sintakties of sintakties wees, maar aangesien dié onderskeid dikwels slegs op skryfwysekonvensies sal berus, is dit nie werklik van belang vir dié artikel nie. Vgl.:

(12) die glase *moet skoon* (< skoongemaak word *naas skoon kom*)

(13) die klas se punte *moet hoër* (< hoër gemaak word / styg / opgestoot word).

(b) Negativering

Die negativering van die stelling in (12) is moontlik, maar verg dikwels 'n sinsklemverandering:

(14a) die glase *moet ?nie skoon nie*

lyk¹ minder bestaanbaar as

(14b) dié glase *moet nie skoon nie* ("maar die ander wel").

(c) **Waarskynlikheid**

Dit hang egter af van die "waarskynlikheid" van die stelling — dit lyk onlogies dat glase vuil gelaat moet word, maar in 'n sin soos

(15) die veters *moet nie los nie*

is die situasie van vaste veters normaalweg 'n meer gewenste toedrag van sake, en lyk dié uiting dus bestaanbaar (kyk ook onder 4. **Beperking op inkorting vanweë onwaarskynlikheid** hier onder).

(d) **Grade van vergelyking**

Die inkortingsresultaat (d.w.s. die ingekorte hoofwerkwoord) kan in die vergrotende en die oortreffende trap wees:

(16) die veters *moet los / lossier / *losste* (< losgemaak / lossier gemaak word).

Let op dat die oortreffende trap *nie losste* is nie, maar **die losste**:

(17) die veters *moet die losste*.

Uit (17) blyk dat *moet* saam met die oortreffende trap waarskynlik net soos *sál* (vgl. (7) hierbo) 'n sinsklemwysiging vereis:

(18) dié veters *moet die losste* ("en nie daardie nie")

(19) dié glase *moet die blinkste* ("en nie daardie nie").

In hierdie verband kan daarop gewys word dat woordeboeke in die algemeen die oortreffende trap van adjektiewe sonder die bepaler *die* aandui, wat misleidend is (vgl. WAT deel IX by ¹los). *Die* vorm 'n integrale deel van die oortreffendetrappvorm in Afrikaans, in teenstelling met Engels, en behoort dus deel van die aanduiding van die oortreffende trap in woordeboeke te wees.²

(e) **Geen agens**

Soos wat vroeër terloops vermeld is, kan 'n agens nie as onderwerp saam met 'n direkte voorwerp by hierdie soort inkorting figureer nie:

- (20) Henning *moet* die glase **blink*
- (21) Johann *moet* die werk **klaar*.

Dit geld ook waar die direkte voorwerp as onderwerp van die lydende vorm optree:

- (22) Die werk *moet* deur Johann **klaar*.

By sekere deeltjiewerkwoorde wat uit voorsetsel plus werkwoord bestaan, kan die ingekorte sin uit onderwerp + *moet* + voorwerp + voorsetseldeeltjie bestaan, bv.

- (23) Nico *moet die grens oor*
- (24) Hans *moet die huis in*
- (25) die dokter *moet die land uit*.

In hierdie gevalle is die ingekorte sin waarskynlik die resultaat van inkorting van hoofwerkwoordestukke soos onderskeidelik (23) *oor die grens gaan*, (24) *in die huis gaan* en (25) *uit die land gaan*.

In 'n sin soos

- (26a) Boetie *moet die berg op*

kan die ongekorte sin wees:

- (26b) Boetie *moet teen die berg opgaan*.

(f) **Dat-vervatting**

Die IS tree op in die *dat*-vervatting:

- (27) Tiaan sê dat die tente *moet op* (< opgeslaan / opgerig word)

maar weereens kan geen onderwerpsagens in dié sin optree nie:

- (28) Tiaan sê dat Jan die tente *moet *op*.

'n Interessante punt hier is dat die inkorting veroorsaak dat die normale inversie van die werkwoordstukke na *dat* nie plaasvind nie: *moet op* teenoor *opgeslaan moet word*.

(g) Vraagsin

Moet tree ook as IS op in die vraagsin:

- (29) *Moet* die skoene *skoon*? (< *skoon* gemaak word?)

(h) Verledetydsvorm *moes*

Die IS-kenmerke van *moet* geld ook die verledetydsvorm *moes*:

- (30) die skoene *moes skoon*
 (31) die damme *moes vol* (< *vol* loop / gemaak word).

(i) Inkorting na links

Die IS *moet* het altyd inkorting na links tot gevolg. Dit is as gevolg van die saamgestelde aard van die hoofwerkwoord waar die eerste deel die kernbetekenis van die werkwoord dra, terwyl die tweede deel hoofsaaklik as semanties "leeg" beskou kan word — soos hieronder met voorbeelde betoog word. Inkorting na regs is dus nie moontlik nie, omdat dit aan die semantiese inhoud van die sin sal verander — en dit druis in teen die voorwaarde waarop inkorting plaasvind:

- (32) die tente *moet *geslaan word* (< *opgeslaan*)
 (33) die tydelike kamers *moet *gerig* (< *opgerig*) *word*.

(j) Inkorting na *wat*

Dieselfde toedrag van sake geld hier as by die *dat*-vervatting soos bespreek in paragraaf (f). Daar is ook nie die normale werkwoordinversie na *wat* nie. (Kyk hiervoor die eerste voorbeeld onder paragraaf 5A, sin (48).)

3. Kategorieë van hoofwerkwoorde waarmee *moet* optree

Dit is duidelik dat nie alle saamgestelde hoofwerkwoorde inkorting ondergaan as die IS *moet* sonder agens in 'n gegewe sin voorkom nie. Twee kategorieë, inkortbaar en nie-inkortbaar, word ondersoek aan die hand van saamgestelde werkwoorde met *-maak* as tweede lid, uit Combrink en Dodds 1988:

A1. Inkortbaar

Vir die werkwoorde in hierdie groep kan kontekste bedink (sien weer aantekening 1) word waarbinne *moet* as inkortingsneller kan optree:

aan die kant maak ("die huis *moet aan die kant*" < "die huis *moet aan die kant gemaak word*"); *blinkmaak*; *bymekaar-* ("die speelgoed *moet bymekaar*" < "die speelgoed *moet bymekaar gemaak word*"); *deurmekaar-* (bv. van bestanddele vir 'n gereg); *dood-* ("die jakkalse *moet —*"); *droog-* ("die klere *moet —*"); *droogskoon-* ("die klere *moet —* — kyk opmerking (b) hier onder); *fyn-* ("die aartappels *moet —*"); *gaar-* ("die kos *moet —*"); *gelyk-* ("die grond *moet —*"); *heel-* ("die kouse *moet —*"); *klaar-*; en op soortgelyke wyse ook *kleinmaak* (bv. van klippe); *leeg-* (bv. van asblikke); *los-*; *nat-* (bv. van blomme); *oop-* (bv. van vensters); *op-* (bv. van beddens); *reg-* (bv. van motors); *skoon-*; *toe-*; *vankant-* (bv. van ondiere); *vas-* (bv. van veters, toue); *vol-*; *vrymaak* (bv. van slawe).

B1. Nie-inkortbaar

Die lede van hierdie groep werkwoorde met die tweede lid *-maak* is skynbaar nie onderhewig aan die inwerking van die IS *moet* nie:

bangmaak (?die kinders *moet bang*); *buit-*; *dof-*; *goed-*; *groot-*; *lig-*; *mee-*; *oor-*; *pas-*; *saam-*; *seer-*; *skrik-* ("die kinders *moet skrik*" is nie noodwendig die inkortingsresultaat van 'n werkwoord *skrikmaak* nie, maar kan dieselfde soort sin wees as "die kinders *moet loop / sit*" ens. waar *skrik*, *loop*, *sit* in dié vorm as hoofwerkwoord optree); *staan-*; *staat-*; *uit-*; *vuil-*; *weg-*; *wysmaak*.

Hieroor kan die volgende opgemerk word:

- (a) By die werkwoorde van groep A1 kan die tweede lid *-maak* as semanties "onspesifiek" bestempel word omdat dit geredelik deur 'n ander tweede lid of ander werkwoord binne die sinskonteks vervang kan word sonder verlies aan of wysiging van die betekenis, bv.

- (34) die huis moet aan die kant (*kom / gemaak word*)
 (35) die glase moet blink (*-gemaak / -gevyf / -gepoleer word*)
 (36) die kouse moet heel (*-gemaak / -gestop / -gelap word*)
 (37) die veters moet los (*-gemaak / -geknoop / -gewoel word*).

- (b) *Droogskoon-* as inkortingsresultaat van *droogskoonmaak* is 'n interessante geval: die derde lid van die hoofwerkwoord, *-maak*, kan teoreties vervang word deur die semanties teenstrydige *-was*:

(38) die klere moet droogskoon (-*gemaak* / *gewas word*).

Moontlik staan die inkorting ook onder die invloed van *skoonmaak* — maar dit dui in elk geval op die hegtheid van die verbinding *droogskoon-* in *droogskoonmaak*: die infinitief is *om te droogskoonmaak*, nie **om droogskoon te maak* nie.

(c) Hierteenoor is die tweede lid *-maak* van groep B1 glad nie of moeilik vervangbaar:

aanmaak naas **aanroer*, **aanmeng*; *meemaak* naas *mee-?*; *pasmaak* naas **pas-?*; *saammaak* naas *saam-?*; *staatmaak* naas *staat-?*, ensovoorts.

(d) As die inkortingstuk van die hoofwerkwoorde in groep A1 'n spesifieke betekenis moet oordra, moet dit óf nie ingekort word nie, óf die spesifieke konteks moet aan die betrokkenes by die situasie bekend wees, bv.:

(39) die aartappels moet *fyngedruk* / *fyngesny* word

kan semanties aan die ingekorte sin

(40) die aartappels moet *fyn*

gelyk wees net as die betrokkenes in die situasie op die besondere handeling van *fyndruk* / *fynsny* besluit het.

(e) Hierdie semantiese beperking op die optrede van die IS *moet* kan ook getoets word aan 'n ander groep saamgestelde werkwoorde, nl. dié met *-trek* as tweede lid:

A2. Inkortbaar

aftrek (naas *afskeur*, *afskuur*, bv. van plakpapier); *intrek* (naas *inhaal*, *inbring*, *inkatrol* van bv. net); *lostrek* (naas *losskeur*, *lostoring*); *ooptrek* (naas *oopmaak* van laai, deur); *optrek* (naas *ophys*, *opkatrol*, *opwen* e.d.); *saamtrek* (naas *saamkom*, *saam vergader*); *terugtrek* (naas *terugkeer*, *terugroep*, *terugval* van bv. troepe); *toetrek* (naas *toemaak*, *toeskuif* van bv. vensters); *vastrek* (naas *vasbind*, *vasmaak*, *vaswoel* van bv. tou), en op dieselfde wyse *wegtrek* (naas *weggaan* van bv. bure).

(a) Let op: 'n Inkorting soos *toe-* (< *toemaak* / *toeskuif* / *toetrek*) kan weer eens nie staan vir 'n hoofwerkwoord met 'n spesifieke betekenis soos *toespyker* nie — tensy die betrokkenes by die uitingssituasie vooraf daarvan bewus is dat met 'n uiting soos

(41) die vensters moet *toe*

...bedoel word dat die vensters toegespyker moet word.

(b) Ander lukraak gekose gevalle bevestig ook dié onderskeid:

(42) die ligte moet aan (-*geskakel* / -*geslaan word*)

(43) die vuur moet aan (-*geslaan* / -*gesteek word*)

(44) Eduard moet die grens oor (-*steek* / -*gaan*)

(45) die tente moet op (-*geslaan* / -*gerig word*).

B2. Nie-inkortbaar

bak-, deur-, dwars-, krom-, na-, neer-, oor-, plat-, stuip-, tou-, vol-, voor-

C1. Inkortbaar op grond van utingsituasie

Sekere saamgestelde werkwoorde kan ingekort word na *moet* as IS binne 'n gegewe utingsituasie wat aan die betrokkenes daarby bekend is — al is die tweede lid nie semanties "onspesifiek" nie. In die sin

(46) Pietertjie se klere moet uit

is dit nie duidelik of *uit* die inkorting is van *uittrek* / *uithaal* / *uitkom* / *uithang* (bv. uit 'n kas of tas) nie. In hierdie geval slaag die inkorting slegs as die bedoe-ling van die spreker in die besondere gespreksituasie ondubbelsinnig is.

4. Beperking op inkorting vanweë onwaarskynlikheid

Daar is reeds melding gemaak van 'n gespreksituasie waarin 'n inkorting nie sal slaag nie omdat 'n uiting onwaarskynlik is, nieteenstaande die moontlikheid van sinonimiese vervanging van die tweede deel van die hoofwoord:

(47) Die glase / skoene moet **vuil* (< *vuilgemaak* / *vuilgesmeer word*)

Dié soort inkorting sou slegs moontlik wees as daar werklik 'n situasie bestaan waar dit wenslik of verpligtend is om glase of skoene vuil te maak. Die afwesigheid van data om hierdie besondere inkorting te staaf, kan egter nie as 'n *argumentum e silentio* geld nie.

5. Leksikografiese getuienis vir *moet* as inkortingsneller

Dit is nie binne die bestek van hierdie opstel moontlik om die volle kaartkorpus en rekenaardata van die WAT na te gaan vir getuienis wat al die inkortingsmoontlikhede van die hoofwerkwoord na *moet* sou staaf of afwys nie. Die logiese aanvangspunt is die lemma *moet*. Van die stawende data is die volgende, byeengebring onder die lemma waar dit in die WAT se kaartkorpus aangetref is:

5A. *moet*

skoonmaak:

- (48) ... *skoene wat moet skoon* (E. Kotze in *Sarie*, 5 Jun. 1974, 54)

doodmaak:

- (49) *Hy moet dood! Hy moet vrekemaak word* (C.M. v.d. Heever: *Laat Vrugte*, 1939, 191)

oorgaan / oorsteek (met agens — kyk (11a):

- (50) ... 'n groot grens waar hy moet oor (C.M. v.d. Heever: *Laat Vrugte*, 1939, 41)

uitskree / uitlap / uitkom:

- (51) *Dit moet uit, dit moet vertel word* (C.M. v.d. Heever: *Laat Vrugte*, 1939, 268)

weggaan:

- (52) *Ek is jammer dat ek so gou moet weg* (C.M. v.d. Heever: *Laat Vrugte*, 1939, 317)

klaarkom / klaarmaak:

- (53) *Die dam moes klaar* (*Die Huisgenoot*, 26 Nov. 1948, 3)

uitkom / uitsê:

- (54) *Die vraag moet uit: is jy kwaad?* (G.H. Franz: *Dillo*, 1956, 38)

Sintaktiese inkorting: verder gaan, loop, reis:

- (55) *Ek moet verder* (A.J.C. Volschenk in *De Goede Hoop*, 15 Sept. 1922, 104)

5B. *uit*

uitgooi:

- (56) *Die klein vlaggies [op die landsvlag] moet uit* (*Die Burger*, 11 Mei 1966, 1)

uittrek / uitgaan:

- (57) *Die witman moet uit* (*Die Burger*, 15 Febr. 1963, 12)

5C. *saam*

saamgaan/saamloop:

- (58) *Waarom die perde nou eintlik moes saam, het ek nooit geweet nie* (A.A.J. v. Niekerk: *Kelkiewyn en Kogelaar*, 1969, 35)

(59) *Hy moet saam met ons, Patel* (Sarie, 24 Jun. 1964, 94)

5D. oop

oop-en-toe jaag:

(60) *Ek weet van 'n hele paar jagters wat al moes oop en toe voor sy haelgeweer uit* (Chris Barnard in *Die Huisgenoot*, 13 Sept. 1974, 60)

5E. oor

oorgaan / oorsteek:

(61) *Daar is 'n vertraging by al die ponte waarmee jy moet oor* (Elsa Joubert: *Suid van die Wind*, 1962, 197)

5F. terug

teruggaan / terugloop:

(62) *Ek moet terug* (W.A. de Klerk: *Die Laer*, 1964, 228)

5G. Ander voorbeelde

Ander voorbeelde wat onder bostaande lemmas in die kaartkorpus voorgekom het, is:

(63) *Hy foeter altyd om te wil saam* (Chris Barnard in *Die Huisgenoot*, 12 Mei 1972, 70)

(64) *Dis die blik wat nie wou oop nie* (A. Rudolph in *Die Taalgenoot*, Jul. 1967, 16).

6. Die inkortingsneller en die woordeboek

Een opgawe by die WAT onder klaar verskyn in die kaartversameling, maar is nie opgeneem in die finale artikel in deel VI van die WAT nie:

(65) *Daar was 'n stuk wat moes klaar* (Linda Joubert: *Roosmaryn en Wynruit*, s.j., 17).

Uit hierdie voorbeeld en die artikels in ander Afrikaanse woordeboeke is dit duidelik dat ingekorte saamgestelde hoofwerkwoorde nie as leksikografies spesiale vorme beskou word nie. Daar is bv. geen opgawes by *sal* waarin *sal* as inkortingsneller optree in die kaartkorpus van die WAT nie.

In die konsepartikel van die lemma *moet* (bedoel vir WAT deel X) word die aandag gevestig op die optrede van dié hulpwerkwoord sonder die infinitief, bv.:

(66) dit was gouer donker as wat dit *moes* (< moes wees).

As hierdie tipe inkorting wel in die WAT weerspieël word, behoort die inkorting van saamgestelde hoofwerkwoorde na *moet* ook daar aandag te geniet — nie slegs volledigheidshalwe nie, maar ook omdat 'n sin soos

(4) die skoene *moet skoon*

in dié vorm afwyk van die grammatikale reëls vir die gebruik van die hulp- en die hoofwerkwoord in Afrikaans. Die opname van dié vorme sal ook die aanleerder van Afrikaans ter wille wees.

Erkenning en opdrag

Die stof vir hierdie opstel is grotendeels verkry uit die data van die Buro van die Woordeboek van die Afrikaanse Taal op Stellenbosch. My dank aan die Hoofredakteur vir toestemming om die kaartkorpus en rekenaardata te raadpleeg. Ek dink met groot dankbaarheid ook terug aan die samewerking wat ek oor 'n aantal jare van wyle my oudkollega, Pieter Harteveld, geniet het. Sy sin vir orde, sy vreugde in die korrekte manier, die netjiese metode, bly almal van ons by wat met hom in aanraking gekom het. Ek bring dus graag met hierdie opstel hulde aan sy nagedagtenis.

Aantekeninge

1. Vanweë die gebrek aan skriftelike stawende materiaal moet daar noodwendig van poëme gebruik gemaak word — maar uit spraaksituasies wat hopelik aan alle moedertaalsprekers oorbekend is.
2. Die invloed van Engels blyk duideliker wanneer dié oortreffende trap aan die begin van 'n sin voorkom. In Engels: "Most people believe..."; in Afrikaans: "Die meeste mense glo...". Dat hierdie onderskeid in die omgangstaal aan die verdwyn is, ly geen twyfel nie, maar tot onlangs is dit as die enigste standaardvorm in Afrikaans aanvaar.

Bibliografie

- Combrink, J.G.H. en R. McD. Dodds. 1988². *Retrograde woordeboek van Afrikaans*. Stellenbosch: Universiteit van Stellenbosch.
- Snijman, F.J. (Hoofred.). 1976. *Woordeboek van die Afrikaanse Taal*. Deel VI. Pretoria: Staatsdrukker.
- Van Schalkwyk, D.J. (Hoofred.). 1994. *Woordeboek van die Afrikaanse Taal*. Deel IX. Stellenbosch: Buro van die WAT.

Using Learner Corpora for L2 Lexicography: Information on Collocational Errors for EFL learners

Yukio Tono, *Department of English Education, Tokyo Gakuei University,
Tokyo, Japan*

Abstract: In this paper, we describe an on-going project of the corpus of EFL (English as a Foreign Language) learners in Japan and its application for pedagogical dictionary compilation. We especially focus on the learners' errors in verb collocation patterns and describe how the learner's dictionary can benefit from the learners' error information based upon the learner corpora.

Keywords: CORPUS, LEARNER CORPORA, CORPUS LINGUISTICS, PEDAGOGICAL DICTIONARY, L2 LEXICOGRAPHY, LEARNER'S DICTIONARY, APPLIED LINGUISTICS, COLLOCATIONS, ERROR ANALYSIS, COLLOCATION ERRORS, VERB PATTERNS, SECOND LANGUAGE ACQUISITION RESEARCH, DICTIONARY USER INFORMATION, USAGE COLUMN, NEGATIVE INFORMATION

Opsomming: Gebruik van aanleederskorpuse vir L2-leksikografie: Inligting oor kollokasionele foute vir EVT-aanleeders. In hierdie artikel beskryf ons 'n voortgaande projek oor die korpus van EVT- (Engels as 'n vreemde taal-) aanleeders in Japan en die toepassing daarvan vir die samestelling van pedagogiese woordeboeke. Ons fokus veral op die aanleedersfoute gemaak ten opsigte van werkwoordkollokasionele patrone, en beskryf hoe die aanleederswoordeboek voordeel kan trek uit die aanleeders se foutinligting gebaseer op die aanleederskorpuse.

Sleutelwoorde: KORPUS, AANLEEDERSKORPUSSE, KORPUSLINGUISTIEK, PEDAGOGIESE WOORDEBOEK, L2-LEKSIKOGRAFIE, AANLEEDERSWOORDEBOEK, TOEGEPASTE LINGUISTIEK, KOLLOKASIES, FOUTANALISE, KOLLOKASIONELE FOUTE, WERKWOORDELIKE PATRONE, NAVORSING OOR TWEETAALVERWERWING, WOORDEBOEKGEBRUIKERSINLIGTING, GEBRUIKSKOLOM, NEGATIEWE INLIGTING

1. Introduction

Recent development of corpus linguistics and actual corpora has been remarkable. Many dictionaries published recently all enjoyed in some way or another the use of large corpora; for example, the *COBUILD English Dictionary* used the Bank of English, the corpus of 20 million words in contemporary English,

developed at the Birmingham University. *Longman Dictionary of Contemporary English* and *Oxford Advanced Learners' Dictionary of Current English* used the British National Corpus, produced by an academic and industrial consortium consisting of Oxford University Press, Longman, Chambers Harrap, Oxford University Computing Services, Lancaster University's Unit for Computer Research on the English Language and the British Library. Older corpora for research were all gathered by the institution ICAME (International Computer Archive of Modern English), which is an international organization of linguists and information scientists working with English machine-readable texts. The aim of the organization is to collect and distribute information about English language material available for computer processing, and about linguistic research on this material, completed or in progress, in order to compile an archive of English text corpora in machine-readable form, and to make material available to research institutions.

Even though many different kinds of corpora have been available internationally, very few researchers have yet built up a corpus of the language learner. There may be a couple of reasons for this; first, the data collected from language learners is in most cases erroneous. The primary interest of the corpus builders at present is to describe the status quo of native speakers' language, so they are basically not interested in collecting learner language data. Secondly, and related to the first, most of the researchers in applied linguistics or TESL (Teaching English as a Second Language) / TEFL (Teaching English as a Foreign Language) are not sufficiently informed about the expertise of corpus linguistics. They are either more or less classroom-oriented researchers or theoreticians like UG-based SLA researchers (Universal Grammar-based Second Language Acquisition researchers), who stress the intuition of the native speaker, rather than the collection of a large text.

More and more attention has been paid, however, to building a corpus of the language learner. To date, International Corpus of Learner English (ICLE) has been one of the largest and the most systematic corpus development projects in the world (Granger 1994). Longman has been developing Learners' Corpus for its dictionary project¹. John Milton of Hong Kong University of Science and Technology has already collected about 8 million words of the writing of Chinese students of English². In Japan, Asao and others held a symposium on EFL learner corpora and SLA (Asao et al. 1995). Our project is also one of the few attempts to develop learner corpora.

In this paper, we will first describe our learner corpus project and then show the application of learner corpora data to English pedagogical dictionary-making as an example and examine the potential which learner corpora have for future L2 lexicography.

2. TGU Learner Corpus Project

Tokyo Gakugei University, a national teacher training college in Tokyo, Japan, launched the project on EFL writing instruction in 1988. The primary research

interest was focussed on the effect of teacher feedback on EFL writing quantity and quality (see Hatori et al. 1990; Kanatani et al. 1993; Tono and Kanatani 1995 for more details). Throughout the data-collection procedures, we collected the free composition data in English from subjects of different academic backgrounds (eighth grade through twelfth grade) and accumulated the data in a machine-readable form.

Table 1 shows the framework of the corpus. The data collection procedure has been largely dependent on the research design of the original writing project. Therefore, the learner profile does not seem to be entirely systematic. For instance, the data for third-year senior high school students was obtained for the first project in 1989 and the number of the subjects was 280. But the following project in 1993 only allowed for 120 subjects for each grade. The data for SH1 was not obtained at the time of the second project in 1993, because the primary focus in our original project was to see if teacher feedback on writing did make a difference as the academic grades increased. This is also the reason why we did not obtain the data for SH1 which would enable us to see the differences between JH groups and SH groups clearly. We will have to fill the gap by collecting the data for SH1 in the near future. We should also note that the data available for the present study was limited to the subcorpora except for SH3. Because of learner profile database management problems, we could not use SH3 data for the analysis of collocation errors.

The size of the whole corpus is about 0.7 million words. As can be seen in Table 1, the size becomes larger for upper-grade groups because more advanced students wrote longer essays. We will have to collect more data for lower-grade samples to create a balance in size among the different academic grade sub-corpora. This corpus is one of the largest learner corpora available in Japan and probably one of the first attempts in the world to collect the interlanguage data from different developmental stages.

3. Collocation errors of English basic verbs

In order to see how the learner corpora contribute to L2 dictionary-making, let us look at the actual data taken from the corpora and discuss its application for lexicographical description. Since it was impossible to examine all the lexical items in the corpus, we chose basic verbs and their collocations for analysis. Table 2 shows the list of the verbs used for the study. As Sinclair says, in order to study the behavior of words in texts, we need to have available quite a large number of occurrences. "About half of the vocabulary of a text — even a very long text — consists of words that have occurred once only in that text." (Sinclair 1991: 18)

In order to get statistically meaningful results, we have to obtain enough observations for each lexical item. In this sense, it was difficult to deal with lexical items whose frequencies were relatively low. Since the size of our

learner corpus was around 0.7 million words, it was almost impossible to have enough occurrences of each of the basic verbs listed in Table 2.

Another alternative was to choose high frequency words such as *the, of, and, to, a, in, that, I, it*, and so forth. They are the forms which occur so frequently that there is no problem to apply statistical procedure to those items. However, as can be seen, most of them are so-called function words and the behavior of these words is rather fixed. We thought that it would be more interesting and of more central importance to include basic verbs in our scope. This does not necessarily mean that the study of those functional words is unimportant. We would like to deal with those items in future research.

4. Corpus analysis procedure

4.1 The basic procedure of text processing

The basic procedure of text processing is shown below:

- 1) Data Input
- 2) Preprocessing
 - i) SGML removal
 - ii) Dehyphenation
 - iii) Tokenisation
 - iv) Morphological Analysis
- 3) Tagging (if necessary)
- 4) Low-level Parsing (if necessary)³
(Grefenstette 1995)

In our analysis, we did not have to preprocess the data as such because they were not taken from electronic sources or OCR. The problem, however, is that we transcribed the composition in a Japanese word-processed format, so all the Japanese characters were typed in Japanese. This made it difficult to compare our data with the data from the Bank of English by PC-DOS programs such as LEXA. All the Japanese characters were just unrecognizable on the program.

Tagging and low-level parsing are necessary steps that must be taken in order to investigate the syntactic behaviors of the words in depth, but in this instance we could not use these procedures. The main reason was that the normal tagger or parser did not work correctly on erroneous texts. Therefore, if you are serious about tagging the learner data, you have to do it manually, or first run the automatic tagger and then correct the text manually. This will be one of the biggest obstacles for further research in this area⁴. Very few studies have been done on how to systematically tag erroneous texts (see, however, Meunier forthcoming). We believe that we will have to overcome this problem in order to fully appreciate the benefit of learners' corpora.

We obtained three different types of statistics for each verb lemma. Here we have to clarify the use of the terms. A lemma is what we normally mean by a 'word.' Many words in English have several actual word-forms — so that, for example, the verb *to give* has the forms *give, gives, given, gave, giving, and to give*. In this text, the composite set of word-forms is called the lemma. This definition is based upon Sinclair (1991: 173). The three statistics are frequency score, MI-score and T-score. Let us take a closer look at each of these.

4.2 Frequency scores

The simplest way to look at the corpus data is to get a frequency list, i.e. how often each different word-form occurs in the text. There are a couple of ways to arrange the list. Sinclair (1991: 30-31) has described three ways: first, turning the text into a list of the word-forms in the order of their first occurrence, noting the frequency of each; second, sorting it in an alphabetical order; third, sorting in a frequency order. In either case, it is very easy to compare relative occurrences of each word.

Let us look at an example. Table 3 shows a part of the frequency list of the verb *bring* in our learner corpus. This data simply tells us that the most frequently occurring words with the lemma *bring* are *my* (69 times), *out* (62 times), *a* (26 times), *the* (12 times) and so forth. It indicates that the learners use this verb with noun phrases and phrasal verbs such as *bring out*.

4.3 Mutual Information Statistic

The mutual information statistic was first introduced for corpus analysis by Church and Hanks (1990). It basically works as a tool for identifying interesting associations among words in a corpus. Suppose that we saw the sequence "bring a " showing up a number of times in the concordances to BRING and wanted to know if there might be a linguistically interesting pattern. Some sequences in the concordances are interesting (e.g. *bring out*), but others such as *bring a* are not, even though they may be quite frequent. Mutual information can help distinguish the more interesting sequences from the less interesting ones by comparing the joint probability of the sequences with chance. Pairs of words with high mutual information scores are likely to be interesting to a researcher. (For more details, see Church et al. 1991; 1994) Table 4 shows the mutual information statistic for *bring*.

4.4 t-scores

The t-scores compare probabilities that "a third word co-occurs with either of two-words" (Grefenstette 1995: 61). For example, we are interested in which is

more common to say, *powerful tea* or *strong tea*. The t-scores will statistically examine which words are significantly more likely to appear after *strong* than after *powerful* (Church et al. 1991: 125). Table 5 shows the t-scores of the verb *bring*:

4.5 Procedures for collocation data analysis

After choosing the basic verbs, we first obtained the frequency lists of each kind of verb form. Next we picked up collocation errors from the lists. Since we had not tagged all the texts yet, we could not pick up errors according to the parts of speech information. Instead, we identified the errors by looking at the first words which immediately followed the node words (in this case, verb lemmas).

5. The results of verb collocation analysis

Table 6 indicates the relative frequency of the basic verbs selected for our study. The verb selection was made according to the frequency data of an English learner's dictionary. The frequency list indicates that even though we chose 70 different verbs, more than a half of them could not actually be used for our study⁵. For example, the verb *carry* occurred only 15 times in the whole learner corpus data. It is very unlikely that any interesting error pattern would appear in such small samples. If we try to generalize any particular pattern by statistically judging its probability, then we need at least more than 10 expected frequencies in each cell⁶. In our case in Table 6, only a small number of verbs such as *become, bring, come, go, get, have, make, play, see, take, think, and want* meet this condition.

Table 7 shows the list of verb collocation errors. The number of the learner errors obtained from the individual composition tasks was rather limited. The main reason for this is that for our free composition tasks, we did not use a multiple-draft design in which the subjects were asked to rewrite the same drafts again and again. Instead, we used different topics for each writing task. Therefore, it was more difficult to collect the data of the same error patterns or corrected forms of the same verbs in different compositions. In spite of the difficulties in data collection, it still indicates some interesting error patterns of the basic verb collocations. We will discuss the results and implications for L2 dictionary making.

6. Integrating the error information into lexicographical description

The learner corpus data shows that the learners fixed error patterns in their use of verb collocations. There are many possible sources of errors such as interlin-

gual errors (overgeneralisation from L1 structures or semantic or lexical structures) or intralingual errors (overapplication of L2 rules, etc.). Whatever the sources, it would be useful for the learners of English to find the information on frequently occurring error patterns. Let us look at some of the common error patterns for EFL learners in Japan and how we could integrate such error information into the dictionary design.

6.1 Errors of verb meanings

The results show that the learners had a tendency to use wrong verbs which were quite similar in meaning. **Become to do*, for instance, is a literal translation of the Japanese phrase "suru youni naru". Learners usually learn the meaning of *become* as "naru" and *come* as "kuru". For Japanese learners of English, the word *become* is more strongly associated with the phrase "suru youni naru" (*come to do*) than the word *come* itself. This kind of error is caused by L1 transfer of verb meanings. The same type of error was observed in the phrase such as **look a dream* (in Japanese, is used the verb *miru* (look; watch) for "have a dream") or **take concert* (which means "have a concert").

In L2 lexicography, therefore, it is very important to provide usage notes on frequently occurring errors such as **become to do* under the entry *come* or *become*. Such learner errors have been ignored in describing a lexical entry, but if it is designed for language learners, the dictionary should contain such information in problematical areas for learners.

6.2 Errors of verb patterns and collocations

The data also shows that learners make quite systematic errors in the use of prepositions or particles after verbs. For instance, many subjects dropped the prepositions in phrases such as "come to ...", "come back to ...", "go to ...", "look at ...", "think about ..." and so on. In Japanese, no prepositions are needed for these verb expressions, so this might be another case of L1 interference. It is also quite confusing for Japanese learners of English that some of these verbs could be used without the prepositions if the following elements are adverbials (e.g. *Come here. Come back here. Go home.*). Therefore, knowing which prepositions or particles should follow the verbs is also another problematical area for Japanese learners of English.

Another common error is to use wrong verb patterns. For example, **go to shopping* instead of saying *go shopping*, or **want do* for *want to do*. These grammatical patterns are very complicated for Japanese learners and they have to learn the behaviors of each verb one by one. Currently most bilingual dictionaries in Japan and monolingual learners' dictionaries such as LDOCE or COBUILD all provide useful grammatical codes for these verb patterns. The

information on the most difficult verb patterns for certain groups of learners, however, has not been fully investigated and described in a dictionary.

In pedagogical dictionaries, more and more information on these collocation errors of "verb + preposition / particle" or other verb patterns should be systematically provided. Especially, the learners should be warned of not only possible errors but also frequent errors by collecting more data on learner English. For advanced learners, the collocation information for the verbs or nouns at 5000 to 7000 word levels is very important, but not many dictionaries offer useful information in a systematic way for this level of lexical items.

7. Conclusion

So far we have seen how the learner corpus can contribute to the systematic analysis of learner errors and how those errors should be dealt with in dictionaries. The effect of negative evidence (i.e. the information on 'something is not possible') in dictionaries is to be empirically tested, but it is worth noting that the information on L1-related errors or the most frequently occurring errors can provide the L2 dictionary users with useful guidelines for correct usage.

Some bilingual (English-Japanese) dictionaries in Japan contain this kind of negative information, but there are still many editors and lexicographers who have reservations about providing "incorrect" usage in a dictionary. This question, however, is worth investigating empirically and more attempts should be made to improve the design of pedagogical dictionaries in order to best suit the needs of language learners.

Notes

1. Longman is said to have about 8 million words in its learner's corpus. (P. Scholfield, personal communication)
2. Milton (personal communication).
3. Some researchers classify tagging and parsing as one of the preprocessing stages (for instance, Church et al. 1991). This stage was based upon the lecture given by Gregory Grefenstette in the Seminar on Computational Lexicography at Kossuth Lajos University in Debrecen, Hungary, from Nov. 27 to Dec. 1, 1995.
4. I would like to thank Fanny Meunier for her helpful comments on the problem of error tagging.
5. These frequency scores were based upon the data from JH-2, JH-3, and SH-2 because of the technical problems we had at the time of data analysis in 1995. This is why the frequency scores were rather small in size.
6. Brown (1988: 190). Usually the occurrences of certain lexical patterns are regarded as frequencies of certain categorical variables. Therefore, non-parametric analysis such as chi-square or phi-coefficient is suitable.

References

- Asao, K., K. Nozawa, S. Ozeki, M. Sugiura and K. Harada. 1995. Learners' Corpus and Second Language Acquisition Research. Paper presented at the symposium at 34th Annual JACET Convention, Seijo University, Tokyo.
- Brown, J.D. 1988. *Understanding Research in Second Language Learning*. Cambridge University Press.
- Church, K. and P. Hanks. 1990. Word Association Norms, Mutual Information and Lexicography. *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*: 76-83.
- Church, K., W. Gale, P. Hanks and D. Hindle. 1991. Using Statistics in Lexical Analysis. U. Zernik (Ed.). *Lexical Acquisition*: 115-164. New Jersey: Lawrence Erlbaum Associates, Inc.
- Church, K., W. Gale, P. Hanks, D. Hindle and R. Moon. 1994. Lexical Substitutability. B.T.S. Atkins and A. Zampolli (Eds.). *Computational Approaches to the Lexicon*: 153-177. Oxford: Clarendon Press.
- Granger, S. 1994. The Learner Corpus: a Revolution in Applied Linguistics. *English Today* 39, vol.10 (3): 25-29.
- Grefenstette, G. 1995. *Computational Lexicography*. Paper presented at the Seminar on Computational Lexicography held at Kossuth Lajos University, Debrecen, Hungary, November 27 - December 1, 1995.
- Hatori, H. et al. 1990. *Effectiveness and Limitations of Instructional Intervention by the Teacher. Writing Tasks in EFL*. Tokyo: The Ministry of Education.
- Kanatani, K., K. Itoh, T. Noda, Y. Tono and N. Katayama. 1993. *The Role of Teacher Feedback in EFL Writing Instruction*. A research project report, Grant-in-Aid for Scientific Research (B). Tokyo: The Ministry of Education, Science and Culture.
- Meunier, F. Forthcoming. Tagging and Parsing Interlanguage. L. Beheydt (Ed.). *Linguistique appliquée dans les années 90*: 21 - 29.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Tono, Y. and K. Kanatani. 1995. EFL Learners' Proficiency and Roles of Feedback: Towards the Most Appropriate Feedback for EFL Writing. *Annual Review of English Language Education in Japan* 6: 1-11. The Federation of English Language Education Societies in Japan.

Table 1: Framework of TGU Learner Corpora Project**Learner Profile:**

Academic year		Number of Subjects
1) 8 th grade	(JH 2) ¹	120
2) 9 th grade	(JH 3)	120
3) 10 th grade	(SH 1) ²	— ³
4) 11 th grade	(SH 2)	120
5) 12 th grade	(SH 3)	280

Learner Profile:

All the writing tasks were free compositions on the following topics:

- 1) Which would you prefer, rice or bread for breakfast?
- 2) What would you bring out if fire broke out?
- 3) Write the story of "After Urashima Taro"⁴
- 4) Tell us about your most horrible dream.
- 5) Tell us about your school festival.
- 6) Tell us about your summer holidays.

Corpus size:

1) 8 th grade	(JH 2) ⁵	96,696 words
2) 9 th grade	(JH 3)	110,037 words
3) 10 th grade	(SH 1) ⁶	—
4) 11 th grade	(SH 2)	143,942 words
5) 12 th grade	(SH 3)	348,340 words

699,015 words

1 JH 2 = Junior High School 2nd year

2 SH 1 = Senior High School 1st year

3 This project was originally conducted for comparing the junior high school groups and the senior high school groups in term of the effect of teacher feedback on composition drafts. The data for SH 1 has not been taken because of the technical problems of our original research design.

4 *Urashima Taro* is a traditional Japanese folktale. Urashima saved a turtle and it took him to Sea Paradise and there he received a beautiful gift box. When he was back on shore, he opened the box and became an old man. The students were asked to write what happened to Urashima after that.

5 JH 2 = Junior High School 2nd year

6 SH 1 = Senior High School 1st year

Table 2: The list of the verbs used in the study

ask	knock	set
break	know	settle
bring	lay	shake
call	leave	show
carry	live	sit
change	look	speak
come	make	stand
count	mind	start
cut	move	stop
do	open	strike
draw	pass	take
drive	pay	talk
drop	pick	tell
fall	play	think
feel	pull	throw
find	push	touch
get	put	turn
give	rise	want
go	roll	wash
hang	run	wear
help	say	wish
hold	see	work
keep	send	write

Table 3: Frequency lists of the verb *bring*

	-3	-2	-1	0	1	2	3				
i	17	i	178	will	119	bring my	69	my	25	and	25
so	16	want	9	would	46	bring out	62	money	14	i	19
happened	4	so	5	to	31	bring a	26	bankbook	9	but	9
be	4	will	5	i'll	20	bring the	12	a	7	of	9
and	3	if	4	i	8	bring some	6	because	6	first.	9
earthquake	3	able	4	should	3	bring all	6	of	5	my	8
if	2	and	3	won't	3	bring it	5	bank	5	because	5
fire	2	forget	2	must	3	bring money.	4	bankbook.	5	if	5
mustn't	2	time	2	can't	2	bring money	4	cds	4	book	4
time	2	like	2	not	2	bring to	4	but	4	is	4
a	2	way	1	be	2	bring nothing	4	it	3	money	4
probably	2	heavy	1	help	2	bring them	4	out	3	from	3
too	2	well.	1	he	2	bring much	3	nothing.	3	a	3
secondly	1	earth-quakei	1	i'd	2	bring something	3	lot	3	card	3
火事	1	things	1	don't	1	bring anything	3	first	3	it	3
expensive	1	have	1	also	1	bring nothing	3	i	3	first	2
the	1	food	1	may	1	bring out.	2	money.	3	in	2
thing	1	decided	1	never	1	bring mathematical	1	books	3	too.	2
get	1	anything	1	ie	1	bring mu	1	guitar.	2	the	2
occured	1	doraemon.	1	been	1	bring stamp	1	anything	2	new	2
january	1	happened	1	it	1	bring moneys.	1	food	2	soccer	2
sleep	1	memo.	1	could	1	bring file.	1	old	2	old	2
	300	heavèy	1	usually	1	bring things	1	when	2	photo	2
i'd	1	house	1			bring precious	1	and	2	too	2

Table 4: List of MI-scores of the verb *bring*

3		2		1		0	1		2		3	
wure	15.581	earth-quakei	15.581	will	12.994	bring	mathematical	15.581	cassettes	15.581	teddy	15.581
expensive	15.581	moment	15.581	would	12.986	bring	handbooks	14.581	cardboards	15.581	kit	15.581
kickin	15.581	heavy	14.581	i'll	12.835	bring	mu	13.976	personal	15.581	panic	15.581
occured	15.581	ready	12.754	i'd	11.976	bring	coco's	13.239	outdoor	15.581	player's	15.581
ofcourse	15.581	afford	12.239	won't	11.898	bring	stamp	13.239	important	15.581	zack	15.581
mustn't	14.239	able	11.452	should	11.473	bring	everywhere	12.976	moneybank	15.581	instrument	15.581
havea	13.976	forget	11.169	help	10.584	bring	bird	12.754	mental	14.581	elder	14.581
gets	13.581	happened	10.806	must	10.135	bring	out	12.529	walk-man	14.581	file	13.976
secondly	13.581	heavy	10.807	may	9.754	bring	precious	12.239	platterlots	13.976	remind	13.976
probably	13.239	try	10.313	also	9.169	bring	personal	10.86	attack	13.976	uniform	13.976
ufo	13.239	decided	10.239	been	9.118	bring	my	10.809	electric	13.976	photo	13.754
happened	12.806	minutes	10.169	to	8.994	bring	anything	10.736	bento	13.581	pass	13.581
there're	12.581	i	9.942	could	8.754	bring	album	10.703	keyboard	13.581	violin	13.581
january	12.239	want	9.495	never	8.842	bring	nothing	10.678	key	13.581	card	13.581
perhaps	12.101	way	9.169	can't	8.308	bring	others	10.607	lighter	13.581	pet	13.581
mine	11.581	anything	9.151	be	7.98	bring	something	10.203	shrine	13.581	mountainering	13.581
i'd	10.976	need	9.085	not	7.854	bring	some	10.022	note	13.239	palace	13.239
changed	10.806	if	8.712	usually	6.86	bring	all	9.991	compact	13.239	base	12.754
earthquake	10.754	food	8.572	don't	5.995	bring	them	9.672	photos	13.239	discs	12.581
once	10.432	will	8.421	he	5.911	bring	two	9.186	coins	13.239	6	12.391

Table 5: List of t-scores of the verb *bring*

	-3	-2	-1	0	1	2	3				
i	4.079	i	13.328	will	10.907	bring my	8.302	my	4.992	and	4.988
so	3.992	want	2.996	would	6.781	bring out	7.873	money	3.738	i	4.317
happened	2	will	2.23	to	5.557	bring a	5.09	bankbook	2.999	of	2.99
be	1.996	so	2.221	i'll	4.471	bring the	3.444	a	2.628	but	2.988
earthquake	1.731	able	1.999	i	2.764	bring some	2.447	because	2.444	my	2.815
and	1.692	if	1.995	won't	1.732	bring all	2.447	bank	2.235	if	2.232
mustn't	1.414	and	1.692	should	1.731	bring it	2.221	of	2.223	because	2.23
probably	1.414	forget	1.414	must	1.731	bring nothing	1.999	cds	1.999	book	1.999
fire	1.413	time	1.409	i'd	1.414	bring them	1.998	but	1.982	money	1.993
too	1.411	like	1.405	help	1.413	bring money	1.993	books	1.73	is	1.98
time	1.409	earth-quake	1	can't	1.41	bring to	1.97	lot	1.73	card	1.732
if	1.407	moment	1	be	1.409	bring anything	1.731	first	1.73	from	1.73
a	1.381	heavy	1	not	1.408	bring something	1.731	out	1.726	it	1.712
occured	1	ready	1	he	1.391	bring much	1.728	it	1.712	a	1.705
kichen	1	afford	1	may	0.999	bring mathematical	1	i	1.628	photo	1.414
expensive	1	happened	0.999	also	0.998	bring bandbooks	1	lighter	1.414	soccer	1.414
wure	1	heavy	0.999	been	0.998	bring mu	1	shrine	1.414	bank	1.413
ofcourse	1	try	0.999	could	0.998	bring coco's	1	cat	1.414	new	1.413
havea	1	decided	0.999	never	0.997	bring stamp	1	pillow	1.414	first	1.411
secondly	1	minutes	0.999	usually	0.991	bring everywhere	1	purse	1.414	too	1.411
gets	1	way	0.998	don't	0.984	bring bird	1	album	1.414	old	1.41
ufo	1	anything	0.998	it	0.966	bring precious	1	tennis	1.413	or	1.41
there're	1	need	0.998	is	0.96	bring personal	0.999	anything	1.413	can't	1.41
january	1	food	0.997			bring album	0.999	what	1.412	when	1.405

Table 6: The frequency list of the verbs used in the study

ask	25	knock	2	set	16
break	75	know	128	settle	0
bring	254	lay	3	shake	9
call	61	leave	63	show	42
carry	15	live	159	sit	15
change	38	look	115	speak	14
come	256	make	275	stand	27
count	2	mind	22	start	33
cut	9	move	33	stop	20
do	272	open	65	strike	0
draw	1	pass	10	take	459
drive	4	pay	19	talk	37
drop	1	pick	0	tell	51
fall	64	play	280	think	508
feel	146	pull	4	throw	12
find	113	push	5	touch	2
get	411	put	27	turn	22
give	120	rise	9	want	558
go	480	roll	2	wash	6
hang	0	run	125	wear	13
help	80	say	154	wish	21
hold	41	see	219	work	60
keep	43	send	10	write	75

Table 7: The list of collocational errors of the basic verbs

Errors	n	Examples (sic)
<i>ask for non-NP</i>	3	I might call Dora and ask for *take me to 22th Century.
<i>ask sb V-ing</i>	1	I will ask my brother *helping me.
<i>become to do</i>	10	If there are a lot of money, I only become *to be able to buy a lot of things.
<i>bring + Ø</i>	1	I won't be able to bring *because of panick.
<i>bring to NP</i>	4	I want to bring *to it.
<i>change NP</i>	6	It became bigger and bigger and changed *Doraemon JP.
<i>be changed NP</i>	2	He was changed *his seikaku JP.
<i>come (no prep) NP</i>	9	Why don't you come *my school festival?
<i>come back NP</i>	8	So he was glad to come *back his house.
<i>come there</i>	7	Many people came *there and very crowded.
<i>do NP</i>	30	I couldn't do *breath / hayaben JP / jikken JP.
<i>fall NP</i>	6	I often a dream that I fell *a hall
<i>feel S</i>	2	I don't feel *don't have breakfast every morning.
<i>feel to do</i>	1	I don't feel *to have had a breakfast
<i>get oldness</i>	1	Taro's getting *oldness
<i>get VP-ing</i>	1	I am getting *sleeping sleepier and sleepier
<i>get use to NP</i>	1	we got *use to it when we were going ...
<i>go (no prep) NP</i>	13	I went *cocert this group with my friend.
<i>go to ADV</i>	3	I go *to somewhere to play
<i>go to V-ing</i>	6	When he went *to fishing,
<i>have V-ing</i>	1	I don't have *bringing things very much
<i>have ADJ</i>	1	I had *interested in this year.
<i>help NP</i>	6	I helped *a lot of work.
<i>help V-ing</i>	1	In the band, I help *making posters.
<i>hold (be held)</i>	7	The festival *held for three days.
<i>keep me alone</i>	1	Please keep *me alone.
<i>keep to do</i>	1	If you were to keep *to take no breakfast.
<i>leave at NP</i>	3	I leave *at the home at 6:50
<i>leave to NP</i>	1	Everyone left to school.
<i>live NP</i>	2	One day I was living *a wonderous room.
<i>look a dream</i>	1	I don't like to look *a dream

Table 7 (continued)

<i>look</i> (no prep) NP	22	One day I looked *her face,
<i>make</i> NP to do	1	Who made. *people to eat three meals?
<i>play</i> (no prep) NP	2	I used OTOSHIDAMA for buy the clothes and play *city.
<i>be run by</i> NP	3	I was being run *by a bad man.
<i>say</i> NP + NP	1	I want to say *my classmate "Thank you".
<i>show</i> to NP	1	I want show *to everyone as many as I could.
<i>sit</i> (no prep) NP	4	I can't sit *the seat in train.
<i>speak</i> (no prep) NP	1	He spoke *many people about his adventure.
<i>stand</i> (no prep) NP	1	I hope that I will stand *the stage.
<i>stop</i> (no prep) NP	1	That trains don't stop *station that I was there.
<i>take</i> (wrong) NP	1	A American rock band come to Japan and take *concert.
<i>take</i> NP (no prep) NP	3	he takes him *Ryugujō again.
<i>talk</i> NP	2	He wouldn't talk *anyone else about his story
<i>tell</i> NP NP	1	I must not tell . *others it.
<i>think</i> (no prep) NP	15	I can't think *all things
<i>think</i> VP	8	I thought *buy new ward-processor
<i>turn back</i> + Adj	1	he tried to turn *back young.
<i>turn</i> NP	3	After he turned *old man,
<i>want</i> VP	15	I want *be tall
<i>wish</i> if ...	1	I wish *if there were some.
<i>wish</i> do	2	he wish *return to Ryugujō.
<i>be written</i> NP	2	A paper on my desk is written *"look behind you."

Note: The column n indicates the number of subjects who made these errors.

Textual Condensation in Printed Dictionaries. A Theoretical Draft*

Herbert Ernst Wiegand, *Germanistisches Seminar
der Universität Heidelberg, 69117 Heidelberg, Germany*

Abstract: This article presents an excerpt from a theory of lexicographic texts which deals particularly with dictionary articles. Almost all characteristics of dictionary articles considered as typically lexicographic may be regarded as results of textual condensation processes. A theory of textual condensation in lexicography thus makes it possible to analyse the lexicographic textual condensation from a uniform perspective; it allows us to calculate exactly the degree of textual condensation, it contributes to the construction of a scientific predicate *user-friendlier as*, and gives us insights which will make the formulation of dictionary articles more teachable.

The theory consists of two parts: a partial theory of inner and a partial theory of outer textual condensation. The first part explains in which way textual condensation may be understood as a process leading from a full text (a text showing complete cohesion and explicit syntax) to a condensed article text (with addressing as syntax substitute and dependence on a metatext). With respect to a monosemous lemma sign, the textual condensation process goes through only one stage, to be exemplified here. With respect to a polysemous lemma sign, there is a second stage, including shifts to the left. Certain article types with certain microstructures such as the annexed microstructures may go through a third stage of textual condensation, accompanied by shifts to the right.

The second part of the theory deals with the condensation processes which regard the carriers of the guiding element and which, inter alia, lead to reference articles and lemma clusters.

Finally, attention is drawn to a different form of textual condensation regarding the proportion of printed characters in relation to the total printed matter.

Keywords: LEXICOGRAPHY, TEXTUAL CONDENSATION, LEXICOGRAPHIC THEORY, DICTIONARY ARTICLE, LEXICOGRAPHIC TEXTUALIZATION, LEXICOGRAPHIC TERMINOLOGY, LEXICOGRAPHIC MICROSTRUCTURE, LEXICOGRAPHIC MACROSTRUCTURE, INNER TEXTUAL CONDENSATION, OUTER TEXTUAL CONDENSATION

Opsomming: *Tekstuele verdigting in gedrukte woordeboeke. 'n Teoretiese konsep.* Hierdie artikel bied 'n uittreksel van 'n teorie van leksikografiese tekste wat spesifiek handel oor woordeboekartikels. Feitlik alle kenmerke van woordeboekartikels wat as tipes

* This paper was presented as keynote lecture at the First International Conference of the African Association for Lexicography, held at the Rand Afrikaans University, Johannesburg, 1st - 2nd July, 1996. The wording of the published version is the same as that of the lecture, with the exception that the handout and the bibliography have been integrated into the text and the footnotes have been added.

leksikografies beskou word, kan gesien word as die resultaat van tekstuele verdigtingsprosesse. 'n Teorie van tekstuele verdigting in die leksikografie maak dit dus moontlik om leksikografiese tekstuele verdigting vanuit 'n uniforme perspektief te analiseer; dit stel ons in staat om die graad van tekstuele verdigting presies te bepaal, en dra by tot die daarstelling van 'n wetenskaplike predikaat *meer gebruikersvriendelik as*, en lewer insigte wat die formulering van woordeboekartikels meer onderrigbaar maak.

Die teorie bestaan uit twee dele: 'n deelteorie van binnetekstuele verdigting en 'n deelteorie van buitetekstuele verdigting. Die eerste gedeelte verduidelik hoe tekstuele verdigting verstaan kan word as 'n proses wat verloop van 'n volle teks ('n teks wat volledige samehang en eksplisiete sintaksis vertoon) na 'n verkorte artikelteks (met adressering as sintaktiese plaasvervanger en die afhanklikheid van 'n metateks). In die geval van 'n monosemiese lemmamerker verloop die tekstuele verdigtingsproses slegs deur een stadium en dit word hier toegelig. By polisemiese lemmamerkers is daar 'n tweede stadium wat verskuiwings na links insluit. Sommige artikeltipes met sekere mikrostrukture soos die aangehegte mikrostrukture mag deur 'n derde stadium van tekstuele verdigting gaan, met gepaardgaande verskuiwings na regs.

Die tweede deel van die teorie handel oor die verkortingsprosesse wat betrekking het op die draers van die gidselement en wat onder andere lei tot verwysingsartikels en lemmagroepes.

Ter afsluiting word die aandag gevestig op 'n ander vorm van tekstuele verdigting, nl. die verhouding van gedrukte karakters tot die totale gedrukte materiaal.

Sleutelwoorde: LEKSIKOGRAFIE, TEKSTUELE VERDIGTING, LEKSIKOGRAFIESE TEORIE, WOORDEBOEKARTIKEL, LEKSIKOGRAFIESE TEKSTUALISERING, LEKSIKOGRAFIESE TERMINOLOGIE, LEKSIKOGRAFIESE MIKROSTRUKTUUR, LEKSIKOGRAFIESE MAKROSTRUKTUUR, BINNETEKSTUELE VERDIGTING, BUITETEKSTUELE VERDIGTING

Content

1. Preliminary remark
2. On the relevance of a theoretical draft of textual condensation
3. Outline of the theoretical draft
4. Basic terminology of a theory of textual condensation in printed dictionaries
5. Inner textual condensation
 - 5.1 The first stage of inner textual condensation
 - 5.2 The second stage of inner textual condensation
 - 5.3 The third stage of inner textual condensation
6. Outer textual condensation
7. Conclusion
8. Acknowledgements
9. Abbreviations of the names for the item classes
10. Bibliography

1. Preliminary remark

The aim of this article is to present a small excerpt from a theory of lexicographic texts. Terms which originate in this theory are used and are not all explicitly introduced here. I do hope, though, that this paper will present a coherent picture of this theory.¹

So far, a complete theoretical draft of textual condensation in lexicography does not exist. My proposal for such a draft is relatively complex (cf. Wiegand 1996b). Please keep in mind that in this article the theoretical draft is portrayed in extremely simplified terms. Without simplification it would not be possible to portray the complete theoretical draft within the framework of an article. In addition, the risk of not being understood *hic et nunc* would be relatively high.

2. On the relevance of a theoretical draft of textual condensation

Almost everybody knows that in dictionaries it is particularly the texts of the dictionary articles which have specific features, with the result that one can speak of genuine lexicographic textualization. What is not common knowledge is the fact that almost all features of article texts which may be regarded as being genuinely lexicographic may be explained as the results of processes of textual condensation. The development of a complete theoretical draft of textual condensation will, therefore, allow us to explain lexicographic textualization proceeding from a single approach and seen from a uniform perspective. There are several reasons why this is of interest to dictionary research.

Firstly, it will make dictionaries belonging to the same dictionary type exactly comparable with respect to specific textual features, i.e. in such a way that the respective degree of textual condensation may be calculated in a strict mathematical sense of the word.

Secondly, since the respective degree of textual condensation has a strong influence on the text processing by dictionary users, the methodological possibility of being able to calculate the degree of textual condensation represents an essential step in developing a manageable formula by which the user-friendliness of dictionaries may be calculated (cf. Wiegand 1996b).

Thirdly, a complete theoretical draft of lexicographic textual condensation is, last but not least, useful in order to make the writing of dictionary articles teachable in such a way that further correlations in the formulation of article texts will be recognized.

1 Important parts of this theory are already published in German and English (cf. Hausmann and Wiegand 1989; Konerding and Wiegand 1995; Pan Zaiping and Wiegand 1995; Wiegand 1983a; 1983b; 1984; 1987; 1988; 1988a; 1988b; 1989; 1989a; 1989b; 1990; 1995; 1996; 1996a; 1996b; 1996c; 1996e); a complete version will be dealt with in the second volume of my book "Dictionary Research" which is in preparation.

3. Outline of the theoretical draft

The theory of textual condensation in lexicography consists of the following two parts (cf. fig. 1):

- a partial theory of inner textual condensation; and
- a partial theory of outer textual condensation.

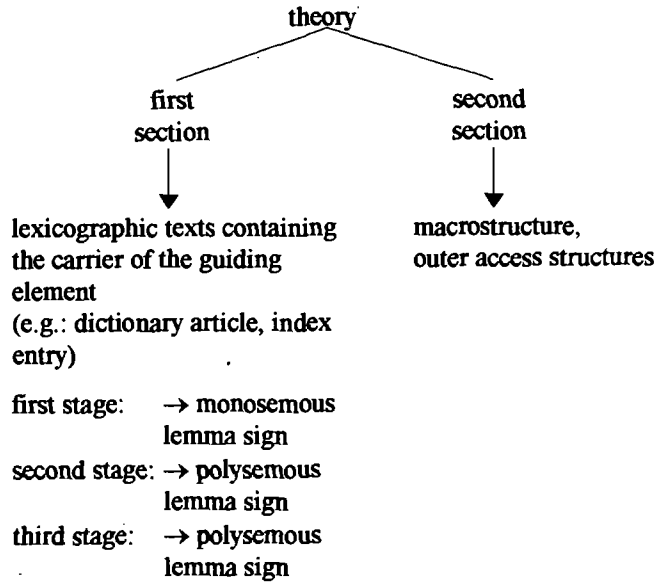


Fig. 1: Two parts of a theory of textual condensation in printed dictionaries

The inner textual condensation concerns all those lexicographic partial texts containing a carrier of the guiding element, in particular dictionary articles and index entries. Regarding dictionary articles in alphabetical order, three different stages of inner textual condensation can be distinguished.

With respect to articles of monosemous lemma signs, the process of textual condensation only goes through the first stage and the procedures of textual condensation applied become effective only in regard to the elements belonging to the domain of the concrete microstructure (sensu Wiegand 1989a; cf. Wiegand 1990[91]).

With respect to articles of polysemous lemma signs the process of textual condensation initially goes through the first stage and then through a second stage, and for specific types of microstructures, a third stage. During the

second and third stage of inner textual condensation with respect to articles appearing in alphabetical dictionaries the procedures of textual condensation applied also become effective in regard to the structures which are formed by the textual constituents, particularly the microstructures.

Outer textual condensation, to be discussed in the second part of the theory, concerns the macrostructure as well as all outer access structures. It becomes effective in regard to the elements belonging to the domains of concrete macro- and access structures.

I did not intend to develop a specific part of the theory which would deal with that type of textual condensation which relates to the mediostructure (sensu Wiegand 1996). The reason for this is the following: with respect to the mediostructure, textual condensation processes can only relate to the elements and not to the relations. Most of the elements belonging to the domain of the mediostructure, however, also belong to either the macro- or the microstructure. This is why the textual condensation which relates to the elements regarding the mediostructure will be discussed within the framework of the two partial theories mentioned above.

The small number of elements belonging to the mediostructure, which at the same time represent elements belonging to the structure of the front and back matter, do not have to be explicitly dealt with in this theory, just like the other results of the textual condensation appearing in the structure of the front and back matter. The reason for this is that these results are not results of the textual condensation which are specific for lexicographic texts.

4. Basic terminology of a theory of textual condensation in printed dictionaries

Every standardized dictionary article, i.e. every article which is formulated in accordance with a set of lexicographic standards relating also to the textual condensation (cf. fig. 2), may be regarded as a two-part condensate created by applying procedures of textual condensation in relation to a respective full text. The full text consists of a title and its cotext. By means of textual condensation the title becomes the lemma (or: headword) and the cotext turns into a linearly structured set of items relating to that lemma sign. Thus, every dictionary article may be regarded as a text which consists of two successive partial condensates (cf. Wiegand 1988a; 1996e).

The full text is a necessary construct of the theory. For instance, if one speaks of a dictionary article being condensed, then a procedure of text processing is presupposed in which a text T_1 is condensed into another text T_2 . The presupposed text T_1 is one of the possible full texts. There are a number of full texts which may be reconstructed for every dictionary article, proceeding from the condensed dictionary article. The following is valid for all full texts of a dictionary article:

- (a) They have to be the same with respect to the elementary propositional contents, with regard to either the form of the lemma sign or the form of other lexicographically processed terms, and they have to be the same with respect to the number of these propositional contents as well as their order.
- (b) As far as the selection of the missing relational expression between the lemma and the item giving the meaning paraphrase is concerned, they may be different.
- (c) They may be different on the surface of the text.

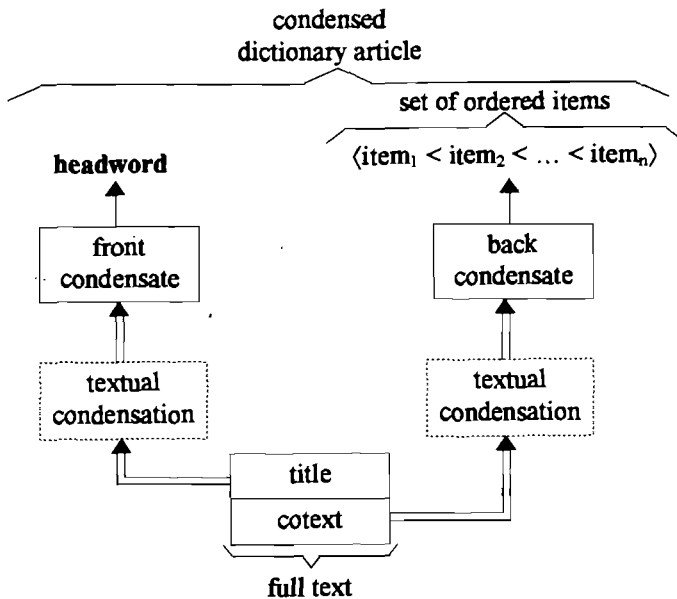


Fig. 2: Visualisation of the condensation process: from a full text to a condensed dictionary article (for further details see Wiegand (1988a; 1996b)); $\text{item}_1 < \text{item}_2 = \text{item}_1 \text{ precedes } \text{item}_2$

At this point, I would like to present the following short dictionary article (= da_1), taken from the *Collins Concise Dictionary of the English Language* (CCDE):

da_1 : **date palm** *n.* a tall feather palm grown in tropical regions for its sweet edible fruit

Fig. 3: Dictionary article from CCDE

In regard to this dictionary article one may reconstruct two full texts, which contain the following title: Dictionary article on *date palm*

This title could have the following cotexts, for example:

- (1) *Date palm* is a noun. A date palm is a tall feather palm grown in tropical regions for its sweet edible fruit.
- (2) *Date palm* is a noun, and its meaning can be given by the paraphrase "a tall feather palm grown in tropical regions for its sweet edible fruit".

The following becomes obvious: the propositional content *that date palm is a noun* is the same in (1) and (2). The presentation of the missing relational expression between the lemma and the item giving the meaning paraphrase, however, is different.

It is possible to lay down standards for the reconstruction of the full texts. Reconstructing the full texts by means of standards becomes necessary in that case — as we will see in more detail — in which one intends to calculate the degree of textual condensation of dictionary articles and compare it with others.

A full text which is subjected to at least one (but not more than a finite number of) textual condensing operation is related to the condensate, i.e. a condensed text as a result of subjecting the full text to textual condensing operations, in several different relations. The following four relations are of special importance:

- x is a standardized full text of the condensate y
- y is a condensate of a standardized full text x
- y is denser in propositional terms than x
- x is less dense in propositional terms than y

The latter two relations may also hold for different condensates of the same full text.

The textual condensing operations mentioned, which may also be called methods of textual condensation, are the following: shortening, abbreviating, omitting, shifting, substituting, summarizing, and embedding.

The respective application of the different methods of textual condensation is laid down in the standardization instructions for textual condensation, which are included in the instruction manual.

Texts do not have cohesion and coherence in the simple positivistic sense. Rather, these two different types of textual connection are always the result of interpretations.

We can speak of cohesion if it is possible to establish a relation at the syntactic, propositional, and referential level (cf. Hellwig 1984a). Subjecting a full text to methods of textual condensation results in denser cohesion; this means that on the way from the full text to the condensate many parts of the full text

which give cohesion instructions are lost.

Despite the fact that there are common traditions of textual condensation, which correspond to the users reading habits within the framework of their general practice of dictionary look-up, the following is generally valid: the results of textual condensation have to be explained in one of the metatexts of the dictionary.

Informally, we now have at our disposal a selected number of essential terms of the theory. In the following, the theory of textual condensation in lexicography will be portrayed by means of examples.

5. Inner textual condensation

As we have seen in section 3, inner textual condensation pertains to lexicographic partial texts containing a carrier of the guiding element. We will now take a look at the first stage of textual condensation by means of a dictionary article on a lemma sign which the lexicographer interpreted as being monosemous.

5.1 The first stage of the inner textual condensation

Let us take a look at da_2 , a dictionary article taken from the *Handwörterbuch der deutschen Gegenwartssprache* (= Concise Dictionary of Contemporary German; abbreviated: HWDG):

da_2 : **B**öschung, die; -, -en meist durch *Mauerwerk, Steine, Pflanzen befestigte, schräg abfallende Fläche im Gelände, bes. die eines Damms, Walls*: die B. bepflanzen, befestigen

Translation of da_2 :

da_2 : **bank**, the; -, -s, the sloping side of an area of raised ground, particularly that of a dam, an embankment, often reinforced by a thick wall, mound of stones or plants; to plant, reinforce the bank

Fig. 4: Dictionary article da_2 from HWDG and its translation

In the following, we will at first take a look at one of the possible full texts of da_2 . The numbers appearing in the square brackets do not belong to the full text (ft_1); they merely serve as a means of allowing easier reference to the text.

$ft_1(da_2)$: [1] Dictionary article on *Böschung* [bank]. [2] The proper spelling of *Böschung* is |*Böschung*|. [3] The form of the nominative singular is *Böschung*. [4] *Böschung* is stressed on the first short syllable. [5] *Böschung* is a noun; [6] its gender is female. [7] The form of the genitive singular is *Böschung*. [8] The form of the nominative plural is *Böschungen*. [9] *Böschung* means any of: *meist durch Mauerwerk, Steine, Pflanzen befestigte, schräg abfallende Fläche im Gelände, bes. die eines Damms, Walls*. [10] Examples of the use of *Böschung* are: *die Böschung bepflanzen, die Böschung befestigen*.

We will first look at [1], the title of the full text which represents a descriptive title. Descriptive titles may be used in one of the following three ways (cf. Hellwig 1984a; Wiegand 1988a):

- (1) They are headings of a cotext.
- (2) They are used within a running text in order to make a reference to another text (i.e. their cotext).
- (3) They are used in catalogs, bibliographies etc. in order to facilitate the search for specific texts.

Particularly when used as headings of a cotext, descriptive titles have the following characteristics, which do not have to occur in combination in every title:

- They are short metatexts of the corresponding cotext.
- They contain hints on the pragmatic purpose of the cotext.
- They imply what might be questionable in the cotext and thus contain hints on the text theme.
- They may aid in interpreting the cotext.
- They confine the cotext with respect to its outer appearance.
- In their function as chapter title, title link, margin title etc. they structure texts.

If one carries out the text shortening operation on the title of the full text, i.e. "[1] Dictionary article on *Böschung*", i.e. in such a way that the text segment "Dictionary article on" is deleted, one obtains the lemma *Böschung* as the front condensate.

The question is now which of the six characteristics ascribed to the title of the full text appears in the condensate of the title, that is the lemma? Obviously, when answering the six respective questions, one has to consider that a lemma does not appear separately, but *exclusively* as part of a dictionary. The first question to be answered is the following: Can lemmata be regarded as

short metatexts of the corresponding dictionary article? The answer has to be: "Yes, they can". The reasoning leading to this answer has to be stated as follows: The title of the full text contains the noun *Böschung*. After the part of the title "Dictionary article on" has been deleted, a mentioned expression remains as lemma. It is by this lemma that the lemma sign *Böschung* is mentioned as lexical unit, the lemma sign standing for the lemma sign paradigm (cf. Wiegand 1983). This means that a user gets to know something about the corresponding dictionary article by means of the lemma, i.e. which part of the dictionary subject is processed within the dictionary article. The mentioning of linguistic signs represents the lexicographic procedure applied in order to introduce the lemma sign as referent for language-reflexive predications in the dictionary article.

The second question, i.e. whether or not lemmata contain hints on the pragmatic purpose of the dictionary article, has to be denied. Such hints are merely conveyed by the dictionary title.

The third question, i.e. whether or not the lemma contains a hint on the text theme, has to be affirmed. Here, it has to be taken into consideration that a lemma always represents part of a dictionary belonging to a specific dictionary type. In our example, *Böschung* is a lemma appearing in a general-purpose monolingual dictionary. This lemma then conveys that specific features of the lemma sign, particularly its forms of inflection and its meaning, represent the theme of the corresponding dictionary article. If we assume the lemma *Böschung* to be part of a dictionary of synonymy in which the synonyms are presented in a cumulative way, then it apparently hints at a different theme of the corresponding dictionary article. The dictionary type therefore contributes to the determination of themes.

The fourth question, too, i.e. whether or not a lemma aids in interpreting the dictionary article, has to be affirmed. For instance, *da*₂ contains the condensed item giving a competence example "die B. bepflanzen". The user may infer from the lemma what the abbreviation B. stands for.

Finally, the fifth and sixth question has to be affirmed as well, because it is obvious without further explanation that lemmata confine dictionary articles with respect to their outer appearance and that they structure the word list in their function as title links.

In the following, we will turn to the full text, of which we have discussed the title so far. Let us first take a look at how the sentences [2] through [8] of the full text are condensed into the comment on form (cf. fig. 5).

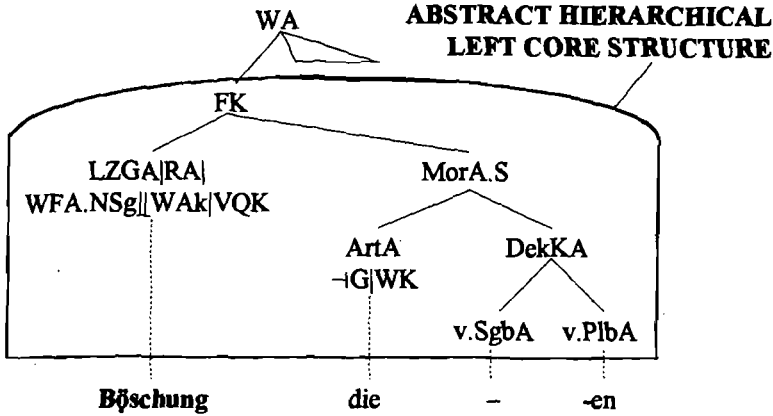


Fig. 5: Annotated structure graph for an excerpt from the (concrete and abstract) hierarchical microstructure which can be assigned to the comment on form of *da₂*; *abbreviations:* WA = dictionary article; FK = comment on form; LZGA = item giving the form of the lemma sign; RA = orthography item; WFA.NSg = wordform item for the nominative singular; Wak|VQK = marking giving the word accent and the vowel quantity; "||" = below enlarged by; MorA.S = morphology item for nouns; ArtA = item giving the definite article; "-|" = what follows, is inferable; G = gender; WK = word class; DekKA = item giving the declension class; v.SgbA = condensed item giving the singular formation; v.PlbA = condensed item giving the plural formation.

I will now describe the process of condensation of the comment on form step by step.

Step 1: Sentence [2] of the full text is omitted. Instead, a metatext of the corresponding dictionary will tell the user that the spelling of the lemmata is in accordance with the official rules of orthography. The propositional content p_2 belonging to [2] reads as follows when put into the canonic *that*-formulation: p_2 = *that the proper spelling of "Böschung" is |Böschung|*. The user familiar with dictionary look-up and the metatext assigns p_2 to the lemma **Böschung**. The item giving the form of the lemma sign is at the same time an orthography item.

Step 2: Sentence [3] of the full text is also omitted. Instead, a metatext of the corresponding dictionary will tell the user that the canonic form for lemmata by which nouns are mentioned is the nominative singular. The propositional content belonging to sentence [3] reads as follows: p_3 = *that the form of the nominative singular is "Böschung"*. The user with a good knowledge of dictionary look-up can assign p_3 to the lemma as well. The item giving the form of the lemma sign is therefore at the same time a wordform item for the nominative singular.

Step 3: Sentence [4] of the full text is substituted by a dot which is placed below the accented syllable. This results in the creation of an item giving the form of the lemma sign below enlarged by a marking giving the word accent and the vowel quantity. Since the function of the dot is explained in one of the

dictionary's metatexts, the user familiar with dictionary look-up is able to refer this dot to its address, the item giving the accented syllable, and assign to the ordered pair of dot and item giving the accented syllable those propositional contents which sentence [4] of the full text contains.

Step 4: Sentences [5] and [6] of the full text are substituted by the definite article *die* which follows the lemma. The metatext of the dictionary gives the explanation that the gender and the word class may be inferred from the item giving the article. The user with a good knowledge of dictionary look-up may therefore infer from the item giving the article *die*, which is addressed to the item giving the form of the lemma sign, that *Böschung* is a noun and that it is of the female gender, when he is referring the item given under the article referring to *Böschung*. In other words: the user with a good knowledge of dictionary look-up may assign to the ordered pair of *der* and *Böschung* those two propositional contents which sentences [5] and [6] of the full text contain.

Here it can be seen that it is necessary to clearly distinguish the items given in the dictionary article from lexicographic information given to a user. Pieces of lexicographic information are cognitive entities. One cannot infer them merely from a single item. One rather has to relate the item to its address, which always represents a different item. In other words: the user always obtains a piece of lexicographic information from an ordered pair of items (cf. Wiegand 1996e).

Step 5: When transforming the full text into the dictionary article, sentence [7] is condensed in such a way that only *Böschung* remains. Next, *Böschung* is substituted by the hyphen "-", resulting in the creation of the condensed item giving the singular formation. Within the framework of a real lexicographic process, such items are given, for instance, by applying a rule of substitution which reads "Within the framework of the item giving the declension class word stems are to be substituted by the hyphen".

For the user, the hyphen as the item giving the singular formation represents a free-standing repetition symbol. He may be able to infer a piece of lexicographic information only if he reverses, so to speak, the operation of substituting. He can do this only if he knows what the hyphen stands for at this position and if he is at the same time familiar with the linguistic form which he has to insert. Here, this form is represented by the mentioned lemma sign. This means that the repetition symbol is an item with cohesion instruction. The user has to establish an anaphoric relation from the repetition symbol to its address, the item giving the form of the lemma sign, and substitute the hyphen by the word stem which may be inferred from this item. It is only after having carried out this operation that a user will receive the information which may be stated as follows: the form of the genitive singular of *Böschung* is *Böschung*.

Step 6: In this step, sentence [8] of the full text, which reads "The form of the nominative plural is *Böschungen*", is first condensed in such a way that only *Böschungen* remains. Next, the word stem is substituted by the hyphen, resulting in the condensed item giving the plural formation. Here, the repetition

symbol is not free-standing, but joined to the word ending. With respect to the operations the user has to carry out in order to receive the right piece of information, the same is valid, as explained in step 5.

After having completed step 6 within the first stage of textual condensation, one may apply the code of the dictionary for the typographic and nontypographic structural indicators to the results obtained by the process of textual condensation. Then one will obtain the comment on form of da_2 .

In the following, we will take a look at how the sentences [9] and [10] of the full text are condensed into the comment on semantics of da_2 . You will find a structural graph depicting this part of the dictionary article in fig. 6.

It is a common phenomenon, about which much has been written, too (cf. Wiegand 1985 and 1989c), that in dictionary articles of modern dictionaries the two-place relational terms which establish the relation between the lemma sign and the item giving the meaning paraphrase are frequently left out. This is also true for the HWDG and thus also in regard to da_2 . Incidentally, there are at least 80 possible relational terms (R) in the German language, if one includes the various forms of technical language. Examples are:

- R₁: an x is a y
- R₂: x means as much as y
- R₃: x means "y"
- R₄: x is used in order to refer to y

In these relational terms, "x" is a variable for lemma signs and "y" serves as a variable for items giving the meaning paraphrase. In the full text on da_2 , I have chosen "x means as much as y" in order to formulate sentence [9].

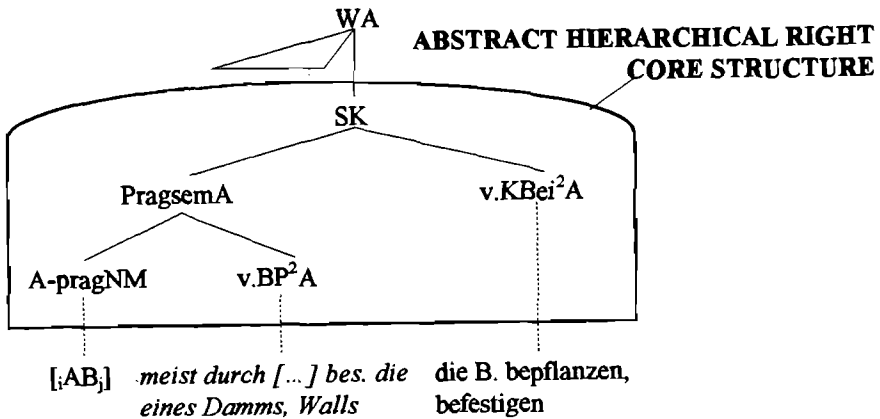


Fig. 6: Annotated structure graph for an excerpt from the (concrete and abstract) hierarchical microstructure which can be assigned to the comment on semantics of da_2 ; abbreviations: SK = comment on semantics; PragsemA = pragmatic-semantic item; v.KBei²A = condensed item giving competence examples, which allows ascertaining two competence examples; A-pragNM = item concerning the pragmatic zero marking; v.BP²A = condensed item giving meaning paraphrases, which allows to ascertain two meaning paraphrases

Step 7: In order to get from sentence [9] to the item giving the meaning paraphrase of da_2 the following textual condensing operations have to be carried out: in sentence [9], the mentioned lemma sign *Böschung* and the relational term "means as much as" have to be deleted. If this were the only condensing operation, then the item giving the meaning paraphrase of da_2 would not be condensed. However, further condensing operations have to be carried out. The word *besonders* has to be abbreviated and the word *oder* has to be substituted by a comma. If the operations of substitution just mentioned are in fact carried out, then this will result in a non-natural textual condensation, because we will obtain a section of a text, i.e. "*die eines Damms, Walls*", regarded as ungrammatical in the language which is being lexicographically processed. The non-natural textual condensation may appear in numerous variants. This is bad lexicographic practice, particularly for users who are not native speakers.

Step 8: In a last step, sentence [10] of the full text is condensed firstly so that only the two examples *die Böschung bepflanzen* and *die Böschung befestigen* remain. These two examples are the same with respect to the first two word forms. In the second example, *die* and *Böschung* are therefore deleted, leading to a comprimation which, in turn, results in the creation of a condensed item yielding competence examples. This allows us to determine two competence examples, i.e. *die Böschung bepflanzen, befestigen*. In this item, *Böschung* is abbreviated as B., and we obtain the item appearing in da_2 . In this case, too, the result of the textual condensation is ungrammatical so that the textual condensation is non-natural.

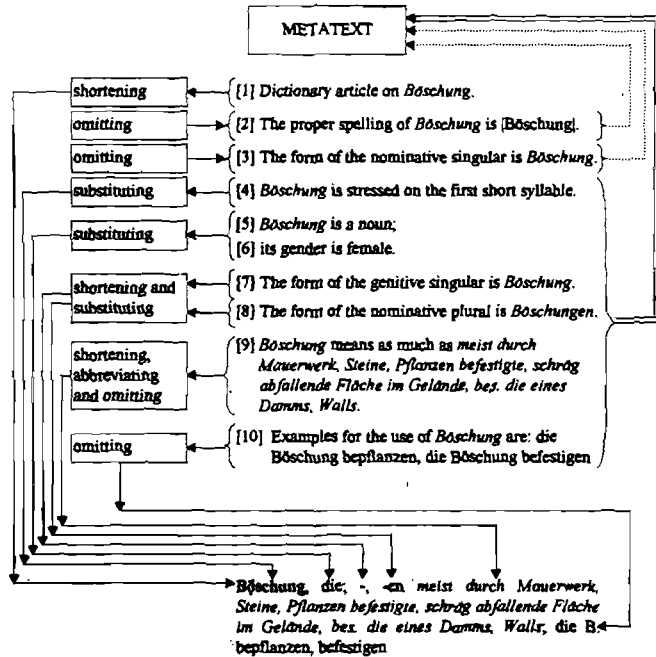


Fig. 7: Visualisation of the condensing operations to get from a full text to a dictionary article

In fig. 7, you will find a survey of the textual condensing operations treated.

At this point, the following can be said: all textual condensing operations which have to be carried out at least once in order to get from a full text to a dictionary article on a monosemous lemma sign may be assigned to a set. The n -th carrying out of the operations belonging to this set which is necessary in order to obtain the condensate represents the first stage of textual condensation.

All standardized dictionary articles on monosemous lemma signs may therefore be understood as the result of textual condensation processes of the first stage. Their degree of textual condensation may be different. The propositional density of the respective dictionary articles may be calculated if one proceeds as outlined in the following.

For each item type which includes elementary items it is possible to formulate a standardized sentence scheme for a sentence of a full text. This sentence scheme always contains either a variable "x" for the lemma sign or, in the case of nonlemmatic addressing, a variable for other items. If necessary, it may contain further variables.

As item type, let us take, for example, the condensed item giving the plural formation. It is then possible to formulate the following as the corresponding standardized sentence scheme for a sentence of a full text: "The nominative plural of x is y."

Here, "x" is the variable for the lemma sign and "y" the one for the plural form. With respect to da_2 , the corresponding sentence of a full text reads as follows: *The nominative plural of "Böschung" is "Böschungen"*.

By means of the sentence scheme it is now possible to create full text sentences for all items giving the plural formation in a single or in a number of dictionaries. These full text sentences consist of the same number of wordforms and, if one does not consider the inserted wordforms, of the same number of letters. If one proceeds like this with each and every item type, one obtains an ordered set of sentence schemes for every dictionary article, by means of which it is possible to create a full text which, in a certain sense, has the same length. This, in turn, will make it possible to calculate different quotients and compare dictionary articles with respect to their degree of textual condensation (cf. Wiegand 1996b). The quotients refer to the number of printed characters used in relation to the number of the elementary propositional contents.

5.2 The second stage of inner textual condensation

With respect to dictionary articles on polysemous lemma signs the operations of the shifting to the left have additionally to be taken into consideration. They represent the second stage of inner textual condensation. Shifting operations of this kind result in textual structures which probably only exist in lexicographic texts.

For all descriptive, factual-referential texts which are not dictionary arti-

cles, and thus for full texts, too, it is valid that statements given within a text are valid until they are revised by subsequent statements given in the same text or, respectively, until interferences from subsequent text sections are suggested which revise the validity of the preceding statement. The statements of a text therefore have fields within the text in which they are valid and they are related to each other by relations which regulate their validity, such as, for example, *exactly when*, *then-relations* or *either-or-relations* etc. The items given in a dictionary article are also valid and they have fields within the article in which they are valid, which becomes particularly comprehensible if one understands them as the result of statements of a full text which have been subjected to textual condensing operations. The fact that a statement is valid means, *inter alia*, the following: with respect to an ordered pair (consisting of this item and its address) there exists at least one valid statement, from which the same propositional content is inferable as from the pair. The items given in the dictionary article are also related to each other by relations which regulate their validity, and if these do not correspond to those of the natural textual constitution they have to be standardizingly regulated, something which has to be explained in the metatext. The occurrence of non-natural relations which regulate validity in articles on polysemous lemma signs may be understood as conditioned by the operations of the shifting to the left, which I will now explain in a few words.

When doing this, I have to rely on extreme simplification. At first, we will assume the most simple case, i.e. the shifting to the left of the pragmatic items.

Holzerei, die; -, -en

1. Fußball umg. abwertend *das Holzen*:
das Spiel artete in eine üble H. aus

2. landsch. umg. abwertend *Prügelei*:
so war denn bald eine böse Holzerei im Gange
WELK Grambauer 334

Translation of da_3 :

(The polysemous word "Holzerei" is rendered in the English language by the monosemous words "rough game" and "roughhouse".)

- $da_{3,1}$: rough game, the; -, -s
soccer colloquial, pejorative *clogging*: the match
turned into a rough game
- $da_{3,2}$: roughhouse, the; -, -s
geographically, colloquial, pejorative *roughhouse*: they
soon engaged in rough boisterous play

Fig. 8: Dictionary article da_3 from HWDG and its translation

This dictionary article is taken from the WDG. Here, the shifting to the left which is prescribed in this dictionary and which has been carried out in article da_4 , for example (cf. fig. 9), has been neglected.

Generator, der; -s, Generatoren (lat.) Techn.
 1. Maschine, die mechanische Energie in elektrische umwandelt, Stromerzeuger, Dynamo
 2. Schachtofen zur Erzeugung von Heiz- und Treibgas
 zu 1 Atom-, Dampf-, Drehstrom-, Gas-, Turbo-, Wechselstromgenerator

Translation of da_4 :

da_4 : generator, the; generators <lat.> technology
 1. a machine by which mechanical energy is changed into electrical energy, electricity producing device, dynamo
 2. shaft furnace kiln in which gas or vapor is formed to 1 nuclear-, vapor-, three-phase current-, gas-, turbo-, alternating current generator

Fig. 9: Dictionary article da_4 from WDG and its translation

The two items in da_3 whose shifting to the left has been neglected are the pragmatic item on style, which reads "umgangssprachlich", and the item on the pejorative meaning, which reads "abwertend". They each appear twice, i.e. once in the first and once in the second sub-comment on semantics. This is why they should have been shifted to the left, i.e. in that article position in which in da_4 there is the item giving the subject field, which reads "Technik".

If we do the neglected shifting operation now, the article da_3 has the form which can be seen in fig. 10.

da_3' : **Holzerei**, die; -, -en umg. abwertend
 1. Fußball *das Holzen*: das Spiel artete in eine üble H. aus.
 2. landsch. *Prügelei*: so war denn bald eine böse Holzerei im Gange WELK Grambauer 334

Translation of da_3' :

$da_{3.1}'$: rough game, the; -, -s colloquial pejorative
 soccer *clogging*: the match turned into a rough game
 $da_{3.2}'$: roughhouse, the; -, -s colloquial pejorative
 geographically *roughhouse*: they soon engaged in rough boisterous play

Fig. 10: Dictionary article da_3' and its translation

In fig. 11 the annotated structural graph of the comment on semantics of the incorrect article *da₃* will be found. Below, fig. 12 will be found the structural graph of the comment on semantics after the shifting to the left has been carried out within the comment on semantics.

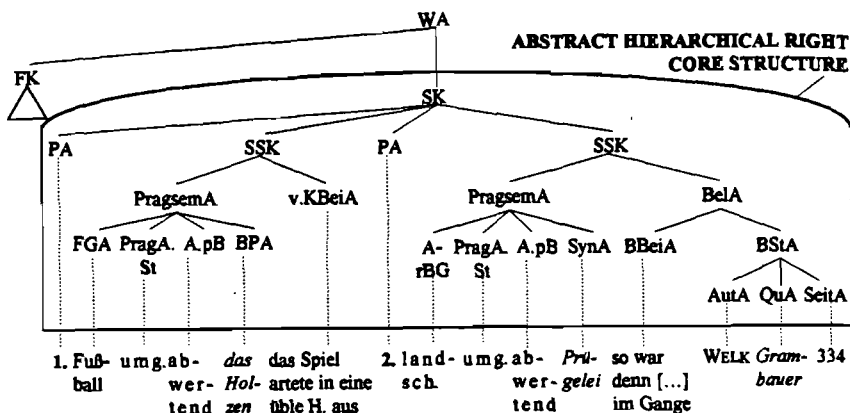


Fig. 11: Annotated structure graph for an excerpt from the (concrete and abstract) hierarchical microstructure which can be assigned to the comment on semantics of *da₃*; *abbreviations:* PA = item giving polysemy; SSK = sub-comment on semantics; PragsemA = pragmatic-semantic item; v.KBeiA = condensed item giving a competence example; BelA = quotation item; FGA = item giving the subject field; PragA.St = pragmatic item on style; A.pB = item of the pejorative meaning; BPA = item giving the meaning paraphrase; A-rBG = item concerning regional restriction of use; SynA = item giving the synonym; BBeiA = item giving a quoted example; BStA = item giving the place where the quoted example is found; AutA = item giving the author's name; QuA = item giving the source; SeitA = item giving the page.

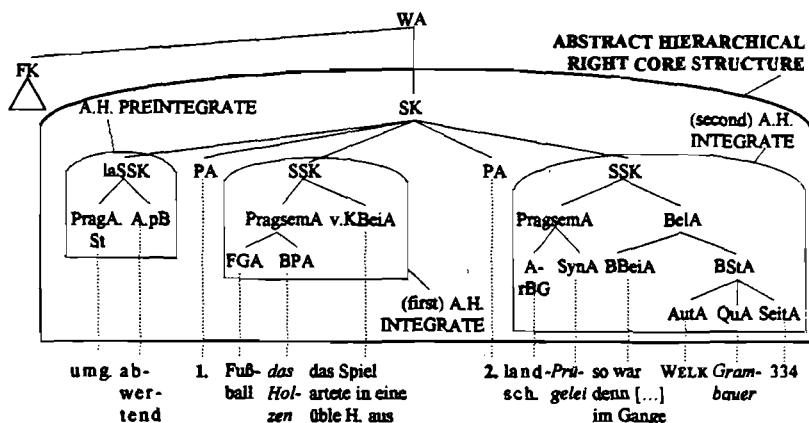


Fig. 12: Annotated structure graph for an excerpt from the (concrete and abstract) hierarchical microstructure which can be assigned to the comment on semantics of *da₃*; *abbreviations:* laSSK = sub-comment on semantics, shifted to the left; A = ABSTRACT; H = HIERARCHICAL.

The comparison of the two structural graphs shows that by means of the shifting to the left the comment on semantics obtains a new direct textual constituent, i.e. the sub-comment on semantics shifted to the left (laSSK). This sub-comment directly precedes the first item giving polysemy. The partial structure of the microstructure which belongs to the sub-comment is called *pre-integrate*. By means of the shifting to the left the complete microstructure of the dictionary article is changed. A simple integrated microstructure turns into a simple partially integrated microstructure with a pre-integrate. Thus the propositional density increases because, on the whole, two pragmatic items, which appear twice in da_3 , are deleted. First and foremost, however, the scope of the shifted items has to be explicitly determined. One of the metatexts has to provide the user with the information that the shifted items "umgangssprachlich" and "abwertend" are valid for the use of the lemma sign with respect to both the first and the second meaning. This means that the textual scope of the two items given in the sub-comment on semantics shifted to the left extends to both of the subsequent sub-comments on semantics. That this is the case cannot be inferred from the shifted items themselves, nor from their textual position. The reason is that after the shifting to the left has been carried out, there are textual connections within the article which do not appear in non-lexicographic texts.

There are further forms of shifting to the left, which I cannot discuss at this point for lack of space (cf. Wiegand 1996b). I would only like to point out that in articles on polysemous lemma signs the complete comment on form, too, may be understood as a textual constituent shifted to the left.

Whereas in the first stage of textual condensation the textual condensing operations become effective only with respect to the elements belonging to the domains of the microstructures, the operations of the shifting to the left, which represent the second stage of textual condensation, also change the relations of the structure.

This is also the case in regard to the third stage of inner textual condensation, which I will tackle now.

5.3 The third stage of inner textual condensation

We speak of the third stage of inner textual condensation in the case of such operations as the shifting to the right which are accompanied by textual condensation. This is only rarely the case. The operation of shifting to the right serves first and foremost as a means of loosening the text. It would be wrong to assume that textual loosening is exactly the opposite of textual condensation. The reason is that if an operation of shifting to the right is carried out the propositional density does not decrease. What happens instead is that the distribution of data within the article is merely changed in such a way that the inner rapid access structure is differentiated, resulting in smaller fields of search with respect to certain search questions (cf. Wiegand 1996a).

One example for an article showing the results of a shifting-to-the-right operation is da_4 in fig. 9. Here, the items giving the compounds are shifted from the first sub-comment on semantics to the right into an annex. In case of a shifting to the right of items given in a sub-comment on semantics the affiliation to the scope of the item giving the meaning paraphrase is lost at the level of the dictionary form. To retrieve it, the item giving the semantic affiliation "Zu 1" has to precede the annex.

Let us now take a look at the article da_5 , given in fig. 13. Here we see two annexes as a result of shifting-to-the-right operations.

- da_5 : **Benennung, die; -, -en**
1. *Name, Bezeichnung*: einer neuen Ware eine B. geben; die alte B. ändern; bekannt unter der B. ...; In diesen Benennungen [der Häuser] war die Sitte der Jahrhunderte deutlich zu erkennen G. KELLER 6,310 (*Kleider machen Leute*)
 2. *das Benennen, die Namengebung*: die B. der neuen Straße nach einem bekannten Dichter
 3. *papierdt. das Namhaftmachen*: die B. von Zeugen; Ein solches Vorgehen wurde im Grundsatz beschlossen und sodann zur Benennung der einzelnen Opfer geschritten G. KELLER 6,575 (*Verlor. Lachen*)
- zu 1 u. 2 Orts-, Warenbenennung
zu 2 Umbenennung

Fig. 13: Dictionary article da_5 from WDG

If there is an annex in a dictionary article, as in the case of da_5 , which is preceded by an item giving the semantic affiliation, by which the annex is assigned to two or more sub-comments on semantics, then the carrying out of the shifting also contributes to loosening the text, but at the same time it is connected with a textual condensing operation of omitting. In da_5 , the condensed item giving the compound word, which allows to ascertain two compound words, i.e. *Ortsbenennung* and *Warenbenennung*, appeared once in the first and once in the second sub-comment on semantics before the shifting to the right had been carried out. However, it has not been shifted to the right twice into two different annexes, but only once, i.e. once it has been deleted. This is why the item giving the semantic affiliation has to read "Zu 1 und 2" and the annex is a condensed annex.

Just like shifting to the left, shifting to the right, too, affects the relations of the microstructures, resulting in the creation of partially integrated microstructures with a postintegrate.

We have now outlined the inner textual condensation and will now briefly tackle the outer one.

6. Outer textual condensation

Outer textual condensation pertains to the carriers of the guiding element such as lemmata and sublemmata, for example. A simple example will be found in fig. 14.

da₆₋₉: **Pflaume** *f* bot олхўри
 ~ baum *m* олхўри дарахти
 ~ kern *m* олўри данаги
 ~ kompost *n* олхўри компоти

Fig. 14: Four dictionary articles from Benjaminow et al. (1985)

These are four short articles taken from Benjaminow et al. (1985). The condensation can be found insofar as, with respect to three lemma signs, which are compounds, the respective first constituent has been substituted by a tilde. This results in partial lemmata (sensu Wiegand 1983). During the condensing operation, however, a mistake has been made, because in each case it has been forgotten to insert the linking letter *n* as the first letter of the partial lemma.

With respect to condensed cross-reference articles, we will find a different form of outer textual condensation. An example, taken from the FWB, is given in fig. 15.

da₁₀: **ablösungsbrief, -geld, -hilfe, -pfund-**
geld, -zeit, s. *ablösung* 2.

Fig. 15: Condensed cross-reference article from FWB

This condensed cross-reference article, from which 5 cross-references can be inferred, may be understood as if it were the result of the 5 cross-reference articles after they have been subjected to condensing operations, which follow each other in alphabetical order and which are given in fig. 16.

ablösungsbrief, s. *ablösung* 2.
 ablösungsgeld, s. *ablösung* 2.
 ablösungshilfe, s. *ablösung* 2.
 ablösungspfundgeld, s. *ablösung* 2.
 ablösungszeit, s. *ablösung* 2.

Fig. 16: Five cross-reference articles

During the transformation of the five cross-reference articles into one condensed cross-reference article, the first thing to be done, for example, is to delete all those items giving the cross-reference which read "s. ablösung", except for those appearing in the last cross-reference article. Next, with respect to all lemma signs except the first one, i.e. *ablösungsbrief*, the first constituent and the linking letter *s* are substituted by a tilde. Finally, all parts of the text which are left on this line are moved up. The result is the condensed cross-reference article given in fig. 15. Expressing it in terms of the macrostructure we have a lemma cluster, whereas in terms of the microstructure there is an item giving the form of the lemma sign within the condensed cross-reference article, from which five lemma signs are inferable.

Even though there are a number of other cases, we have now gained a modest insight into outer textual condensation and I would like to come to a concluding remark.

7. Conclusion

In conclusion, I would like to draw your attention to a phenomenon which I deliberately did not discuss, but which also belongs in the context of textual condensation. Textual condensation serves inter alia as a means to save printing space. To save printing space means that a certain amount of printable matter is filled with as many propositional contents brought into written form as possible. This may also be achieved by selecting, for example, a smaller type size, by reducing the line spacing, and by making minimal use of different means of outer structural formatting such as paragraph formation etc. This is then a type of textual condensation which does not change the proportion of the number of printed characters used in relation to the number of elementary propositions, but the proportion of the number of printed characters used in relation to the printable matter. A discussion on this different type of textual condensation will be found in Wiegand 1996b.

8. Acknowledgements

I would like to thank Antje Immken for translating the German text, as well as Matthias Kammerer for producing the typescript.

9. Abbreviations of the names for the item classes

A.pB	Angabe der pejorativen Bedeutung	<i>item of the pejorative meaning</i>
A-pragNM	Angabe zur pragmatischen Nullmarkierung	<i>item concerning the pragmatic zero marking</i>
A-rBG	Angabe zur regionalen Beschränkung des Gebrauchs	<i>item concerning regional restriction of use</i>
ArtA	Artikelangabe	<i>item giving the definite article</i>
ArtA...G WK	Artikelangabe, aus der das Geschlecht und die Wortklasse erschließbar ist	<i>item giving the definite article, which allows to ascertain the gender and the word class</i>
AutA	Autorenangabe	<i>item giving the author's name</i>
BBeiA	Belegbeispielangabe	<i>item giving a quoted example</i>
BelA	Belegangabe	<i>quotation item</i>
BPA	Bedeutungsparaphrasenangabe	<i>item giving the meaning paraphrase</i>
BStA	Belegstellenangabe	<i>item giving the place where the quoted example is found</i>
DekKA	Deklinationenklassenangabe	<i>item giving the declension class</i>
FGA	Fachgebietsangabe	<i>item giving the subject field</i>
FK	Formkommentar	<i>comment on form</i>
GA	Genusangabe	<i>item giving the gender</i>
laSSK	links ausgelagerter Subkommentar zur Semantik	<i>sub-comment on semantics, shifted to the left</i>
LZGA	Lemmazeichengestaltangabe	<i>item giving the form of the lemma sign</i>
MorA.S	Morphologieangabe bei Substantiven	<i>morphology item for nouns</i>
PA	Polysemieangabe	<i>item giving polysemy</i>
PragA.St	pragmatische Angabe zum Stil	<i>pragmatic item on style</i>
PragsemA	pragmatisch-semantische Angabe	<i>pragmatic-semantic item</i>
QuA	Quellenangabe	<i>item giving the source</i>
RA	Rechtschreibangabe	<i>orthography item</i>
SeitA	Seitenangabe	<i>item giving the page</i>
SK	semantischer Kommentar	<i>comment on semantics</i>
SSK	semantischer Subkommentar	<i>sub-comment on semantics</i>
SynA	Synonymangabe	<i>item giving the synonym</i>
v.BP ² A	verdichtete Bedeutungsparaphrasenangabe, aus der zwei Bedeutungsparaphrasen erschließbar sind	<i>condensed item giving meaning paraphrases, which allows to ascertain two meaning paraphrases</i>
v.KBeiA	verdichtete Kompetenzbeispielangabe	<i>condensed item giving a competence example</i>
v.KBei ² A	verdichtete Kompetenzbeispielangabe, aus der zwei Beispiele erschließbar sind	<i>condensed item giving competence examples, which allows to ascertain two competence examples</i>
v.PlbA	verdichtete Pluralbildungsangabe	<i>condensed item giving the plural formation</i>
v.SgbA	verdichtete Singularbildungsangabe	<i>condensed item giving the singular formation</i>
WA	Wörterbuchartikel	<i>dictionary article</i>
WAK VQK	Wortakzentangabe und Vokalquantitätskennzeichnung	<i>marking giving the word accent and the vowel quantity</i>
WFA.NSg	Wortformenangabe für Nominativ Singular	<i>wordform item for the nominative singular</i>

10. Bibliography

- Benjaminow, J.R., T.S. Mirsatow, S.S. Saidow and M.S. Rustambekowa.** 1985. *Deutsch-Uzbekisches Wörterbuch*. Taschkent.
- Frühneuhochdeutsches Wörterbuch.** 1989. Hrsg. v. Anderson, Robert R., Ulrich Goebel, Oskar Reichmann. Bd. 1: Einführung. A—äpfelkern. Bearb. v. Oskar Reichmann. Berlin. New York. [= FWB].
- Handwörterbuch der deutschen Gegenwartssprache.** 1984. In zwei Bänden. Von einem Autorenkollektiv unter der Leitung von Günter Kempcke [...]. Bd. 1: A—K; Bd. 2: L—Z. Berlin. [=HWDG].
- Hausmann, Franz Josef and Herbert Ernst Wiegand.** 1989. Component parts and structures of general monolingual dictionaries: a survey. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 328-360.
- Hellwig, Peter.** 1984. Titulus oder über den Zusammenhang von Titeln und Texten. Titel sind ein Schlüssel zur Textkonstitution. *Zeitschrift für germanistische Linguistik* 12: 1-20.
- Hellwig, Peter.** 1984a. Grundzüge einer Theorie des Textzusammenhangs. *Text — Textsorten — Semantik. Linguistische Modelle und maschinelle Verfahren*. Hrsg. v. Anneli Rothkegel und Barbara Sandig. Hamburg (Papiere zur Textlinguistik 52): 51-79.
- Ising, Gerhard.** 1978. Textverdichtung und Redundanz in der Lexikographie. *Wissenschaftliche Zeitschrift der Wilhelm-Pieck-Universität Rostock* 27. *Gesellschafts- und sprachwissenschaftliche Reihe*. H. 12: 9-13.
- Konerding, Klaus-Peter and Herbert Ernst Wiegand.** 1994. Framebasierte Wörterbuchartikel. Zur Systematisierung der lexikographischen Präsentation des Bedeutungswissens zu Substantiven. *Lexicographica* 10 [published 1995]: 100-170.
- Pan Zaiping and Herbert Ernst Wiegand.** 1995. Über die Musterartikel für das Große Deutsch-Chinesische Wörterbuch. Zugleich ein Beitrag zu einer Theorie zweisprachiger lexikographischer Texte. *Studien zur zweisprachigen Lexikographie mit Deutsch II*. Hrsg. v. Herbert Ernst Wiegand. Hildesheim. New York (Germanistische Linguistik 127-128): 63-190.
- The New Collins Concise Dictionary of the English Language.** 1982. Managing editor: William T. McLeod. Consultant editor: Patrick Hanks. London and Glasgow. [= CCDE].
- Wiegand, Herbert Ernst.** 1983. Was ist eigentlich ein Lemma? Ein Beitrag zur Theorie der lexikographischen Sprachbeschreibung. *Studien zur neuhochdeutschen Lexikographie III*. Hrsg. v. Herbert Ernst Wiegand. Hildesheim. Zürich. New York (Germanistische Linguistik 1-4/82): 401-474.
- Wiegand, Herbert Ernst.** 1983a. Ansätze zu einer allgemeinen Theorie der Lexikographie. *Die Lexikographie von heute und das Wörterbuch von morgen*. Hrsg. v. Joachim Schildt und Dieter Viehweger. Berlin (Ost) (Linguistische Studien, Reihe A. Arbeitsberichte 109): 92-127.
- Wiegand, Herbert Ernst.** 1983b. Überlegungen zu einer Theorie der lexikographischen Sprachbeschreibung. *Symposium zur Lexikographie. Symposium on Lexicography. Proceedings of the Symposium on Lexicography, September, 1-2 1982, at the University of Copenhagen*. Ed. by Karl Hyldgaard-Jensen and Arne Zettersten. Hildesheim. Zürich. New York (Germanistische Linguistik 5-6/82): 35-72.

- Wiegand, Herbert Ernst. 1984. On the Structure and Contents of a General Theory of Lexicography. *LEXeter Proceedings. Papers from the International Conference on Lexicography at Exeter 9-12 September 1983*. Ed. by R.R.K. Hartmann. Tübingen (Lexicographica. Series Maior 1): 13-30.
- Wiegand, Herbert Ernst. 1984a. Prinzipien und Methoden historischer Lexikographie. *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Hrsg. v. Werner Besch, Oskar Reichmann, Stefan Sonderegger. 1. Halbbd. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 2.1): 557-620.
- Wiegand, Herbert Ernst. 1985. Eine neue Auffassung der sog. lexikographischen Definition. *Symposium on Lexicography II. Proceedings of the Second International Symposium on Lexicography May 16-17, 1984 at the University of Copenhagen*. Ed. by Karl Hyldgaard-Jensen and Arne Zettersten. Tübingen (Lexicographica. Series Maior 5): 15-100.
- Wiegand, Herbert Ernst. 1987. Zur handlungstheoretischen Grundlegung der Wörterbuchbenutzungsforschung. *Lexicographica* 3: 178-227.
- Wiegand, Herbert Ernst. 1988. Vorüberlegungen zur Wörterbuchtypologie. Teil I. *Symposium on Lexicography III. Proceedings of the Third International Symposium on Lexicography. May, 14-16th 1986 at the University of Copenhagen*. Ed. by Karl Hyldgaard-Jensen and Arne Zettersten. Tübingen (Lexicographica. Series Maior 19): 3-105.
- Wiegand, Herbert Ernst. 1988a. Wörterbuchartikel als Text. *Wörterbücher. Artikel und Verweisstrukturen. Jahrbuch des Instituts für deutsche Sprache 1987*. Hrsg. v. Gisela Harras. Düsseldorf (Sprache der Gegenwart LXXIV): 30-120.
- Wiegand, Herbert Ernst. 1988b. "Schanghai bei Nacht." Auszüge aus einem metalexikographischen Tagebuch zur Arbeit beim Großen Deutsch-Chinesisches Wörterbuch. *Studien zur neuhochdeutschen Lexikographie VI*, 2. Teilbd. Mit einem Namen- und Sachregister zu den Bänden I-VI sowie einer Bibliographie zur Wörterbuchforschung. Hrsg. v. Herbert Ernst Wiegand. Hildesheim. Zürich. New York (Germanistische Linguistik 87-90): 522-626.
- Wiegand, Herbert Ernst. 1989. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: Alphabetische Anordnungsformen und ihre Probleme. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 371-409.
- Wiegand, Herbert Ernst. 1989a. Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 409-462.
- Wiegand, Herbert Ernst. 1989b. Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*, 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 462-501.
- Wiegand, Herbert Ernst. 1989c. Die lexikographische Definition im allgemeinen einsprachigen Wörterbuch. *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 530-588.

- Wiegand, Herbert Ernst.** 1990. Printed Dictionaries and Their Parts as Texts. An Overview of More Recent Research as an Introduction. *Lexicographica* 6. [published 1991]: 1-126.
- Wiegand, Herbert Ernst.** 1994. Deutsch-Türkmenisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Türkmenischen Staatlichen Magtymguly-Universität in Aschghabat. *Lexicographica* 10. [published 1995]: 249-300.
- Wiegand, Herbert Ernst.** 1995. Lexikographische Texte in einsprachigen Lernerwörterbüchern. Kritische Überlegungen anlässlich des Erscheinens von Langenscheidts Großwörterbuch Deutsch als Fremdsprache. *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag.* Hrsg. v. Heidrun Popp. München, 463-499.
- Wiegand, Herbert Ernst.** 1996. Über die Mediostrukturen bei gedruckten Wörterbüchern. *Symposium on Lexicography VII. Proceedings of the Seventh Symposium on Lexicography May 5-6, 1994 at the University of Copenhagen.* Ed. by Arne Zettersten, Viggo Hj. mager. Tübingen (*Lexicographica. Series Maior* 76): 11-43.
- Wiegand, Herbert Ernst.** 1996a. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger Lexikographischen Kolloquium.* Hrsg. v. Herbert Ernst Wiegand. Tübingen (*Lexicographica. Series Maior* 70): 1-82.
- Wiegand, Herbert Ernst.** 1996b. Textverdichtung in der Lexikographie. Ein wichtiger Ausschnitt aus einer Theorie lexikographischer Texte. *Symposium on Lexicography VIII. Proceedings of the Eighth Symposium on Lexicography May 2-3, 1996 at the University of Copenhagen.* Ed. by Arne Zettersten, Viggo Hj. mager. Tübingen [forthcoming].
- Wiegand, Herbert Ernst.** 1996c. *A theory of lexicographic texts. An overview.* [Keynote lecture, given at the Annual Congress of the Linguistic Society of Southern Africa 1996. Forthcoming].
- Wiegand, Herbert Ernst.** 1996d. Deutsch-USbekisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Staatlichen Usbekischen Weltsprachen-Universität in Taschkent. *Lexicographica* 12. [Forthcoming].
- Wiegand, Herbert Ernst.** 1996e. *Printed Language Dictionaries and their Standardization: Notes on the Progress Toward a General Theory of Lexicography.* [Forthcoming].
- Wolski, Werner.** 1989. Formen der Textverdichtung im allgemeinen einsprachigen Wörterbuch. *Wörterbücher. Ein internationales Handbuch zur Lexikographie.* 1. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand. Berlin. New York (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1): 956-967.
- Wolski, Werner.** 1991. Formen der Textverdichtung im zweisprachigen Wörterbuch. *Wörterbücher. Ein internationales Handbuch zur Lexikographie.* 3. Teilbd. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand. Berlin. New York 1991 (Handbücher zur Sprach- und Kommunikationswissenschaft 5.3): 2837—2854.
- Wörterbuch der deutschen Gegenwartssprache.* 1974⁷. Hrsg. v. Ruth Klappenbach u. Wolfgang Steinitz. 1. Bd.: A—deutsch; 2. Bd.: Deutsch—Glauben; 3. Bd.: glauben—Lyzeum. Hrsg. v. R. Klappenbach u. W. Steinitz. [...] 7. bearb. Aufl. Berlin.

Die WAT en etimologie: Word die kringloop voltooi?

Willem Botha, *Buro van die Woordeboek van die Afrikaanse Taal,
Stellenbosch, Suid-Afrika*

Abstract: The WAT and etymology: Will the cycle be completed? Afrikaans historical lexicography is a neglected discipline and to a large extent metalexicographers hold the overall-descriptive Afrikaans dictionary, *Die Woordeboek van die Afrikaanse Taal*, responsible for this state of affairs. The WAT, however, describes itself as a synchronic dictionary aimed at describing the present usage of Afrikaans. Limited etymological information is included. Ironically it was the ideal of the first editor of the WAT that it should contain concise etymological information on the whole of the Afrikaans lexicon. This task however, proved to be well beyond his reach and future editors did not believe a historical description of Afrikaans to be the aim of the synchronic WAT. Practical proposals are made for the satisfactory presentation of etymological information on the Afrikaans lexicon.

Keywords: ETYMOLOGY, ETYMOLOGICAL DICTIONARY, HISTORICAL LEXICOGRAPHY, OUTSIDE-AUTHORS, SATELLITE PROJECTS, SECONDARY ETYMOLOGICAL SOURCES, SYNCHRONIC DICTIONARY

Opsomming: Die Afrikaanse historiese leksikografie is 'n verwaarloosde dissipline en in die metaleksikografie word die vinger na die *Woordeboek van die Afrikaanse Taal* se wegstroom van etimologiese inligting gewys as een van die hoofoorse van hierdie toestand. Die WAT beskryf homself egter as 'n omvattende sinchroniese woordeboek wat hoofsaaklik gerig is op die hedendaagse taalgebruik. Tog was dit die ideaal van die eerste hoofredakteur om bondige etimologiese inligting by elke lemma te verskaf — 'n taak wat geblyk het ver buite sy en toekomstige redaksies se bereik te wees. Praktiese voorstelle word gemaak vir die bevredigende aanbieding van etimologiese inligting oor die Afrikaanse leksikon.

Sleutelwoorde: BUIITE-OUTEURS, ETIMOLOGIE, ETIMOLOGIESE WOORDEBOEK, HISTORIESE LEKSIKOGRAFIE, SATELLIETPROJEKTE, SEKONDÊRE ETIMOLOGIESE BRONNE, SINCHRONIESE WOORDEBOEK

1. Inleidend

Dit ly geen twyfel nie dat die historiese leksikografie die mees verwaarloosde dissipline van die Afrikaanse taalkunde is en inderdaad nog in sy kinderskoene staan (Gouws en Ponelis 1992). In die metaleksikografie word die vinger na die *Woordeboek van die Afrikaanse Taal* se afskeep van etimologiese inligting

gewys as een van die grootste bydraende faktore tot hierdie toestand. Die WAT beskryf homself egter as 'n omvattende sinchroniese woordeboek "wat hoofsaaklik gerig is op die hedendaagse taalgebruik" (Van Schalkwyk 1994: i). Betekenisonderskeidings word nie in historiese volgorde gegee nie, maar word van die hoogste na die laagste gebruiksfrekwensie, tydens bewerking, gerangskik.

Die situasie rondom etimologiese inligting in die WAT was egter nog nooit staties nie. Dit het geëvolueer van 'n gestelde ideaal van bondige etimologieë by elke lemma na geen etimologieë nie; dáárna na die selektiewe aangawe van etimologieë, en laastens na die insluiting van beperkte etimologiese inligting op 'n selektiewe basis.

In hierdie artikel word ondersoek ingestel na oplossings vir die bevredigende aanbieding van etimologiese inligting oor die Afrikaanse leksikon.

2. Historiese verloop van historiese inligting in die WAT

2.1 'n Ideaal word nagestreef: Die WAT as eerste bron van etimologiese inligting oor die Afrikaanse leksikon — 1926-1945

Prof. J.J. Smith, verbonde aan die destydse Departement Hollands-Afrikaans aan die Universiteit van Stellenbosch, is in 1926 as eerste hoofredakteur van die WAT aangestel. Die vraagstuk oor die ontstaan van Afrikaans was in dié jare op die voorgrond en prof. Smith het sy eerste en laaste openbare lesings aan dié onderwerp gewy. Smith se standpunt was dat Afrikaans spontaan uit 17de-eeuse Nederlands ontwikkel het en dit is nie onwaarskynlik nie dat hy sy opvatting oortuigend wou bewys deur die aanbieding van gegewens van die taal self in die WAT (Gericke 1991: 75-76). Aangesien byna geen etimologiese navorsing in dié tyd bestaan het nie, het Smith hom in die voorbereidingswerk vir die WAT, al hoe meer in etimologiese navorsing verdiep.

'n Paar "proefdefinisies" is in 1936 in die *Stellenbossche Oudstudent* gepubliseer in die hoop dat belangstellendes welwillende kritiek aan die hoofredakteur sal stuur. Dit is insiggewend dat etimologiese gegewens kortliks aan die einde van elke artikel tussen vierkanthakies gegee is (Gericke 1991: 55). Dit was in ooreenstemming met Smith se uiteensetting van die samestelling van die WAT in *Die Suidsterm* van 13 Febr. 1937. Daarin verduidelik hy dat die herkoms van elke grondwoord, onduidelike samestelling of afleiding aan die einde van die artikel gegee sal word. Dit sou baie bondig wees: "net genoeg om aan die nie-vakkundige leser te sê of die woord uit Nederlands geërf of aan 'n ander taal ontleen is" (Smith 1962: 101).

Die aangawe van etimologie was bewys van Smith se ideaal om gestalte te gee aan 'n wetenskaplike verklaring van die herkoms van die Afrikaanse leksikon. Die studie van allerlei plaaslike taalreste en vervorminge kon 'n verrassende lig werp op die vraagstuk van herkoms. Dit moes egter verbind word

met die studie van ou bronnemateriaal en daar moes in Nederlands, Frans, Duits, Portugees, die ander inheemse tale en veral in die Nederlandse dialekte gesoek word na gegewens met betrekking tot moontlike vertrekpunte en ooreenkomste (Gericke 1991: 76).

2.2 Die realisme seëvier: Geen etimologie in 'n oorwegend sinchroniese woordeboek nie — Deel I tot IV

Die selfopgelegde taak van etimologiese navorsing het egter só baie van Smith se tyd in beslag geneem dat dit as een van die oorsake geïdentifiseer is waarom daar in 19 jaar niks gepubliseer is nie. Ironies genoeg het baie min konkrete bewys van dié navorsing behoue gebly. 'n Deel daarvan lê verskuil tussen die vier miljoen optekeninge in kaartvorm in die databank van die WAT. Die grootste deel daarvan was waarskynlik deel van sy persoonlike wete en is nêrens opgeteken nie. 'n Enkele publikasie waarin van sy navorsing vervat is, het in 1962 verskyn, nl. *Op ons taalakker*. Hierin is die artikels byeengebring wat hy vir die gelyknamige taalrubriek in *Die Suiderstem* geskryf het.

Die tydrowendheid van prof. Smith se etimologiese navorsing het daartoe aanleiding gegee dat daar uiteindelik besluit is op 'n oorwegend sinchroniese WAT. Daar is verder geoordeel dat etimologieë, hoe interessant ook al, nie noodsaaklik is vir betekenisbeskrywing in 'n sinchroniese woordeboek van Afrikaans nie (Hauptfleisch 1989: 12). Met die aanstelling in 1946 van dr. P.C. Schoonees, hoofredakteur van dele I-IV, het die Raad van Beheer gekonstateer dat afgesien word van die opname van etimologiese inligting.

Daar is egter wel in 'n mate etimologiese inligting verstrek in dele I tot V aangesien die herkomstaal by leenwoorde en leenwoordgroepe aangegee is, bv. *aperto libro* (L.), *bric-à-brac* (Fr.), *Cockney* (Eng.), *Führer* (D.), *juvenilia* (L.), *kalamaika* (Hongaars), *kalambak* (<Mal.).

2.3 Etimologie op 'n selektiewe skaal: Deel VI tot VIII

Na die verskyning van Boshoff en Nienaber se *Afrikaanse Etimologieë* in 1967 het die redaksie besluit om op 'n selektiewe skaal "etimologiese aanduidings" te gee, in die besonder by Afrikaanse woorde wat inheems ontstaan het, omdat sulke etimologieë nie geredelik in ander bronne beskikbaar is nie. Die besluit het vervolgens in WAT VI tot VIII neerslag gevind.

Die aangawe van etimologieë het daarop neergekom dat etimologie in beginsel wel gegee is by inheemse Afrikaanse woorde sowel as by leenwoorde en leenvertalings uit nie-Suid-Afrikaanse tale.

Dit is egter nie gegee by Afrikaanse woorde van direkte Nederlandse herkoms nie. Die ratio hieragter was dat die herkoms van hierdie gedeelte van die Afrikaanse woordeskat in Nederlandse etimologiese en verklarende woorde-

boeke soos veral die WNT nageslaan kon word. 'n Uitsondering is gemaak in die geval van Ndl. leenwoorde uit ander tale.

Die etimologie is een of twee stappe terug gevoer tot deursigtigheid bereik is, bv. **krieket** (<E. cricket), **heemkunde** (via N. <D. Heimatkunde, "kennis v.d. aarde, geografie en folklore v.d. eie land"). Die woord in die herkomstaal is nie in sy dele verdeel en die dele afsonderlik verklaar nie. By gevalle van onsekere etimologie is die etimologie gekwalifiseer deur *moontl./waarsk.* daarvoor te plaas. Die etimologiese gegewens is volgens 'n vaste stelsel by verskillende tipes woorde gegee, bv. leenwoorde (genre (F., "soort")) en basterleenwoorde (**hokkie** (<E. hockey)).

In Deel VII het onder andere die volgende voorbeelde volgens die beleid van selektiewe etimologiese inligting verskyn:

komers', komberse; kombersie. (<N. kombaer(t)s, kombe(e)rs, kombaars, "skeeps- en soms soldatekomers", *waarsk.* 'n verbastering van Port. coberta, "bedekking vir 'n bed")

Soos in baie ander gevalle is bg. 'n verbetering en verheldering van Boshoff en Nienaber se inskrywing.

kondens'melk. (kondens: <E. condensed).

kon'trabande. Ook soms *kontraband*. (<It. contrabando, F. contrabande).

ko'nka, konkas; konkatjie. Ook soms *ghonka, katonka, kongha, tjonka, tonka*, (<X. inkonkxa, "blik", (i)t(h)unga (ilit(h)unga), "(melkemmer)"; <Z i(li)t(h)unga, "(melk)emmer").

kombuis', kombuise; kombuisie. Ook soms i.d. spreektaal *kambuis*. (via N. <Mnl. cabuse, "voorraadkamer; kookruimte op 'n skip").

kong'si, kongsi's. (via N. <Sjinees).

komboloi, komboloia (G.) of kombolois.

Soms word die etimologiese gegewens nie in die gleuf daarvoor aangegee nie, maar op 'n indirekte wyse in die sitate. By 'n woord soos **koerant** wat via Ndl. uit die Frans ontleen is en nie meer as 'n leenwoord aangevoel word nie, word uitgebreide etimologiese inligting uit Johan Combrink se *Waar kom daardie woord vandaan?* in 'n sitaat aangegee. 'n Soortgelyke prosedure word gevolg by o.a. *koejawel, kanferfoelie* en *kraakporselein*.

3. Etimologie in 'n sinchroniese woordeboek: 'n Metaleksikografiese perspektief

In die metaleksikografie is daar periodiek uitsprake gemaak oor die opname van etimologiese inligting in 'n sinchroniese woordeboek en, meer spesifiek, in die WAT.

Scholtz (1972: 107) spreek hom sterk uit teen die opname van etimologieë in die WAT. Hy sê onder meer: "Die WAT is nie veronderstel om etimologieë te gee nie ... Dit lyk my wenslik dat die WAT hom tot sy eintlike taak bepaal, om ons so 'n volledig moontlike inventaris en beskrywing van die hedendaagse Afrikaanse woordeskat te gee. Die tyd het nog nie gekom dat ons 'n woordeboek van Afrikaans op historiese grondslag kan aanpak nie. Daar sal nog enorm baie materiaal versamel en voorstudies geskrywe moet word."

Alhoewel hierdie uitspraak reeds 24 jaar gelede gemaak is, dui die stand van navorsing daarop dat die situasie nie wesenlik verander het nie.

L.C. Eksteen (1976) deel nie Scholtz se siening nie en is van mening dat daar genoeg materiaal beskikbaar is sodat dit moontlik en wenslik is om 'n Afrikaanse woordeboek op historiese grondslag te begin. Hy meld ook dat daar in die hedendaagse leksikografie 'n vermenging van diachroniese en sinchroniese praktyke voorkom. Sinchroniese woordeboeke gee dikwels ook historiese inligting, soos kort etimologiese verwysings.

Volgens Gouws (1989: 64) maak Scholtz in sy siening oor die aangawe van etimologiese inligting nie voorsiening vir die inligtingsoorvleueling wat in verskillende woordeboektypes voorkom nie. Talle sinchroniese woordeboeke sluit ook etimologiese inligting as 'n mikrostrukturelement in. Selfs historiese inligting soos die dagtekening van die eerste optekening van 'n woord se gebruik, word soms gegee. Dit is volgens hom ook die geval met die WAT wat baie inligting bevat wat binne die historiese woordeboek se terrein val.

Gouws en Ponelis (1992) spreek die mening uit dat die eise van omvattende historiese leksikografie ver buite die bereik van selfs twee sulke uitsonderlik skerp en ywerige intellekte soos dié van Boshoff en Nienaber val en dat bittermin etimologiese navorsing reeds gedoen is: "Though valuable work has been done, Afrikaans historical lexicography is still in its infancy. The discipline faces a challenging and daunting task well beyond the most passionate dedication of individuals."

Hulle voer verder aan dat die term "omvattende sinchroniese woordeboek" 'n *contradictio in terminis* is, aangesien 'n omvattende woordeboek se bestek nie beperk kan word tot slegs 'n sinchroniese beskrywing van die taal nie. Die taak van die WAT is volgens hulle om meer as sinchroniese inligting te gee.

In resensies van die WAT-dele, wat vir lank die enigste vorm was waarin die metaleksikografie in Afrikaans beoefen is (Gouws (moet nog verskyn)), word dikwels kritiek gelewer op die hantering van etimologiese gegewens.

In haar resensie van die agste deel van die WAT sonder Edith H. Raidt (1993) die wegstroom van etimologieë en die afwesigheid van historiese beginsels uit as tekortkominge in dié deel. Haar kritiek is teen die "hibridiese opset van die WAT as 'n verklarende woordeboek wat slegs nou en dan etimologies, en slegs nou en dan histories te werk gaan." Volgens haar weerspieël dié benadering die leemtes in die leksikografie van Afrikaans, nl. "ons besit geen historiese woordeboek van Afrikaans nie, en die etimologieë van woorde is maar

net gedeeltelik beskryf." Sy gaan dan egter voort deur te sê: "Daar is beslis reeds baie navorsing gedoen en talle etimologieë is goed nagevors; bestaande studies soos die van Boshoff en Nienaber en Scholtz behoort deeglik benut te word."

Indien die WAT hierdie bronne "deeglik" sou benut, sou hy nog steeds 'n hibriediese karakter hê, omdat die WAT soveel meer leksikale items opneem as wat die genoemde bronne behandel. Die etimologiese navorsing in die betrokke bronne is dikwels van 'n bespiegelende aard, wat die opname daarvan in die WAT, waarvan die inhoud as gesaghebbend beskou word, problematies maak. Verdere navorsing word dikwels vereis wat baie tydrowend kan wees.

Op 25 November 1989 begin 'n polemiekk in *Die Burger* oor etimologiese inligting in die WAT wat voortgevloei het uit 'n verdoemende artikel gemik op die WAT op 18 November deur die uitgewer Koos Human. Die bekendste onder die polemici was Wilhelm Grütter, rubriekskrywer van *Die Burger* wat aanvanklik die aantyging maak dat die WAT geen etimologieë verskaf nie terwyl dit van hom verwag kan word; daarna dat etimologieë op inkonsekwente wyse opgeneem word. Na reaksie van Botha (1989) en Hauptfleisch (1989) is dit veral Harteveld (1989 en 1990) wat oënskynlik daarin slaag om bevredigende antwoorde te verskaf aan die ontevrede gebruikers. Die kern van Harteveld se verweer was dat die bestaande etimologiese navorsing ontoereikend is, dat die WAT nie oor die personeel, tyd en geld beskik om etimologiese navorsing te onderneem nie, dat etimologiese inligting in 'n sinchroniese woordeboek soos die WAT, soos ander historiese inligting, as 'n bonus beskou moet word en dat die tydgenootlike beskrywing van woordbetekenisse in die WAT die belangrikste is.

4. Etimologie op 'n selektiewe en beperkter skaal: Deel IX

Na die voltooiing van Deel VIII, en die letter K, was die beplanning van die WAT primêr op verkorting en versnelling gerig. As deel van die versnelling van die werk aan die WAT is daar in 1989 besluit om die aangawe van etimologiese inligting vanaf Deel IX te verminder. Intensiewe navorsing oor die herkoms van woorde is as te tydrowend beskou omdat daar nie 'n enkele volledige etimologiese bron in Afrikaans bestaan nie en omdat die redakteurs 'n gebrekkige kennis van die historiese taalkunde het. Die ingewikkelde hanteeringswyse van etimologiese inligting was 'n bydraende tydrowende faktor.

Die redaksie het besluit om minder etimologiese inligting te gee, alhoewel hulle bewus was van:

- * die sterk aandrang van die gebruiker op meer en uitgebreide etimologieë,
- * die feit dat Gouws (1989: 64) en Eksteen (1976: 67-75) aangetoon het dat sinchroniese woordeboeke dikwels etimologiese inligting insluit,

- * dat daar benewens die etimologieë in Boshoff en Nienaber se *Afrikaanse Etimologieë* en die *WNT* ook enkele ander etimologiebronne bestaan
- * en dat dit sou lei tot 'n vergroting van die hibriediese karakter van die WAT.

Daar is onder andere aangevoer dat alhoewel daar enkele etimologiese bronne bestaan, etimologiese inligting selde onveranderd uit 'n ander bron oorgeneem kan word (vgl. die behandeling van *konka* in die WAT en in Boshoff en Nienaber) en dat oënskynlik eenvoudige gevalle verskillend behandel word in verskillende bronne (vgl. die behandeling van *camoufla*ge).

Daar is tot die slotsom gekom dat die beplanning van die WAT primêr op verkorting en versnelling gerig is en dat etimologie op die toe heersende of op 'n uitgebreide skaal slegs verskaf kan word indien 'n bykomende redaksielid met 'n taalhistoriese opleiding aangestel word.

In Deel IX van die WAT is etimologiese inligting in die volgende gevalle aangebied:

- * Leenwoorde en leenwoordgroepe wat **onveranderd** in Afrikaans opgeneem is en waarvan die ortografie as vreemd aangevoel word, ongeag daarvan of dit via 'n ander taal opgeneem is en of dit na betekenis van die herkomswoord verskil, bv. *lady* s.nw. (*Eng.*), *lazaretto* s.nw. (*It.*), *kunstablyment* s.nw. (establishment (*Eng.*)).
- * Voorbeelde van volksetimologie, bv. *katjiejering* s.nw. (*volksetim.* <Mal. katjapiring, "glasbakkie"), *hanepoot* s.nw. (poot: *wsk. volksetim.* <Ndl. kloot, "teelbal"), *lewensliksens* s.nw. (liksens: *volksetim.* <essens).
- * Simplekse, samestellings, uitdrukkings en ander woordgroepe wat van 'n eiennaam afgelei is of waarin 'n eiennaam voorkom, bv.: *leotard* s.nw. (*na J. Leotard, Fr. sweefstokarties*), *limerick* s.nw. (*na Limerick, stad en graafskap in Ierland*) *liebigkondensator* s.nw. (*na J. von Liebig, D. chemikus*), *lesbiër* s.nw. (*na Lesbos, tans Mutilene, 'n eiland i.d. Egeïese see*).
- * Waar iets wat eg Afrikaans is, deur 'n vorm uit 'n ander taal verdring word, bv. *lighuis* s.nw. (*n.d. Eng. lighthouse*), *lanskorporaal* s.nw. (*lans: n.d. Eng. lance*), *die laaste strooi* of *die laaste strooi wat die kameel se rug breek* (*n.d. Eng. the last straw of the last straw which breaks the camel's back*).
- * Afrikaanse letterwoorde en verkortings, bv. *vigs* s.nw. (<*verworwe immunitetsgebreksindroom*).

5. Kritiek op die hantering van etimologiese inligting in Deel IX van die WAT

Na voltooiing van Deel IX in 1994 in die rekordtyd van twee en 'n half jaar het die hoeveelheid beskikbare etimologiese inligting wat as gevolg van die nuwe

stelsel verlore gaan, redaksieledede opnuut laat nadink oor die aangawe van etimologiese inligting. Veranderings in die aanbod van etimologiese inligting is aanvanklik vir Deel X, wat die letter M omvat, oorweeg, maar is laat vaar ten einde nie die werk aan die deel te vertraag nie.

Feinauer (1996) is in haar resensie van Deel IX van die WAT van mening dat die WAT dit moet oorweeg om etimologiese inligting uit te brei om ook "bepaalde" uitdrukkings in te sluit: "Uitdrukkings in Deel IX wat hulle leen tot etimologiese toeligting is bv.: die lakens uitdeel, onder die loep neem, op die lappe kom, op jou louere rus." Sy meld nie watter maatstawwe aangelê moet word om te ontkom aan die kritiek van inkonsekwente aangawe van etimologiese inligting nie as bv. van die navorsing van D.F. Malherbe (1925) en J. du P. Scholtz (1985) gebruik gemaak sou word.

6. Moontlike opsies vir die bevredigende verskaffing van etimologiese inligting

6.1 Aanpassing van die huidige aanbod van etimologiese inligting in die WAT

'n Hinderlike aspek van die hantering van etimologiese inligting in Deel IX is dat daar by leenwoorde waarvan die skrifbeeld minimaal van die herkomstaal verskil, geen etimologiese inligting verstrek word nie, bv. lasaret (It. lazaretto), lasso (Sp. lazo), lechaim (Hebr. lechayim), letieng (So. leting). So gaan 'n aansienlike hoeveelheid beskikbare etimologiese inligting verlore.

Steeds uitgaande van 'n primêre ingesteldheid op versnelling en verkorting kan daar deur 'n geringe aanpassing in die stelsel voorsiening gemaak word vir bogenoemde gevalle. By leenitems waarvan die ortografie as vreemd aangevoel word en die Afrikaanse skrifbeeld verskil van dié van die leentaal, kan die verskil aangedui word, bv. lasaret (<It. lazaretto), moretla (<So. moretlwa), nabob (<Port. nababo). In die Toeligting kan gesê word dat die aanduiding (<herkomstaal) nie noodwendig beteken dat dit 'n direkte ontleding is nie. Dieselfde sou ook gesê kon word van die aanduiding van die herkomstaal in gevalle soos largo (It.) en restaurant (Fr.).

Indien daar 'n verskil in betekenis in Afrikaans ingetree het, moet die betekenis in die herkomstaal aangedui word, bv. mpata (<Z. umphata, "iemand wat nie goed begryp nie"), nausea (<Gr. nausia, "seesiekte").

By samestellings waarvan die eerste of laaste komponent 'n leenwoord is, kan dieselfde hantering gevolg word, bv. nabadingetjie (naba: Nama, "om te blink" of mntl. <Nama gabba, "beker, skottel"), narkohipnose (narko: <Gr. narke, "verlamming"), nabank (na: hou mntl. verband met Nama !na, benaming vir 'n groot skilpadsoort of met Nama !noa, "grys").

Dit is egter te betwyfel of dié geringe aanpassing 'n bevredigende oplossing is vir die leemtes rondom etimologiese inligting in Deel IX van die WAT.

6.2 Etimologiese inligting by elke lemma in die WAT deur gebruikmaking van sekondêre bronne

'n Radikale oplossing sou wees om etimologiese inligting by elke lemma te oorweeg. Dit is haalbaar indien slegs gebruik gemaak word van alle beskikbare sekondêre bronne en geen oorspronklike etimologiese navorsing gedoen word nie. Die strewe moet nie wees na 'n historiese oorsig van die morfologiese of fonologiese ontwikkeling van leksikale items nie, maar die klem moet val op **betekenis**. Dit gaan dus nie in die eerste plek oor die oorspronklike vorm en die veranderings wat in die loop van die geskiedenis daarin plaasgevind het nie, maar oor die **ontwikkeling in die betekenis en die verheldering wat dit bring ten opsigte van die huidige betekenis van die item**.

Indien daar by leengoed nie enige bewys is van interessante betekenisontwikkeling nie, word slegs die taal van herkoms aangedui. Dit beteken nie noodwendig dat dit 'n direkte ontlening is nie, bv. **nasi goreng** (*Mal.*), **lasaret** (<*It.* lazaretto). Daar word nie daarop ingegaan of die woord *via* 'n ander taal ontleen is nie, omdat dit soms tydrowend of selfs onmoontlik is om presies vas te stel langs watter weg 'n item 'n bep. leksikon bereik het. So is Nederlandse etimoloë onseker of **narkotiseur** aan Frans of Duits ontleen is en is dit raadsaam om slegs die verband met Latyn en Grieks aan te dui: **narkotiseur** (<*Ndl.* narcotiseur <*Lat.* narcoticus <*Gr.* narkotikos <narkoun "verlam" <narke "verlamming").

By erfgoed moet daar sprake wees van interessante betekenisontwikkeling in die historiese verloop van die *Ndl.* woord, of die *Afr.* woord moes in vergelyking met sy *Ndl.* ekwivalent interessante betekenisontwikkeling ondergaan het of deur iets eiesoortig gekenmerk word, anders word slegs aangedui dat die woord *Ndl.* van herkoms is, bv. **naaf** (*Ndl.*). In Van Dale se *Etimologiese Woordenboek* (1991) word bv. die volgende inligting by **naaf** verskaf:

naaf (middenstuk waardoor de as gaat) **middelnl.** *nave*, *na(e)f*, **oudsaksisch** *naba*, **oudhd.** *naba*, **oudeng.** *nafu*, **oudnoors** *nof*; buiten het germ. **oudpruisisch** *nabis*, **oudindisch** *nabhya* - [naaf].

Hier is dus geen sprake van inligting oor betekenisontwikkeling nie, maar slegs van 'n morfologiese oorsig binne en buite die Germaanse taalfamilie.

Versigtigheid moet aan die dag gelê word om nie sonder meer te sê dat een woord "uit 'n ander kom nie" en in navolging van Boshoff en Nienaber en ander bronne kan gesê word dat 'n woord "verband hou" of "korrespondeer" met, of volgens die *Ndl.* bronne "beantwoord aan", 'n ander. Daar hoef selde na die, vir ons, onbekende en veraf lede van die taalfamilie soos Noors, Oudfries, Oudhoogduits, Goties en Sanskrit verwys te word, maar soms kan na die voorkoms van 'n woord in die ander Germaanse tale verwys word, of aangedui word dat 'n woord verband hou met Latyn of Grieks of Engels omdat daar 'n mate van bekendheid daarmee by gebruikers bestaan. Dit kan gesin-

jaleer word deur die leestekengebruik wat dit voorafgaan, bv. 'n kommapunt:

naak b.nw. Ook *nakend* en, verouderend, *naakt*. (<Ndl. naakt <Mnl. naket, naect, nakent; Nhd. nackt, Eng. naked; *hou verband met Eng. nude en Lat. nudus "kaal"*)

By wisselvorme kan slegs etimologiese inligting gegee word by die vorm waar die definisie verskaf word:

nartjie s.nw., nartjies. Selde ook *naartjie*. (*wsk. <Arab. naranj of mntl. <Tam. nartei*)

Lemmas word nie in komponente opgebreek ter verduideliking van hul vorming nie, maar waar slegs een komponent etimologiese inligting vereis, word slegs inligting oor daardie komponent verstrekk, bv.

naberou: **naberou is galberou** idiomatiese uitdrukking (galberou: <Ndl. galg(e)berouw "berou uit doodsangs aan die voet van die galg" of "vals, onegte berou uit vrees vir straf en te laat getoon". *Uitdr. onbekend in Ndl.*)

nagmerrie s.nw. (merrie: Ndl. <Mnl. merrië volksetim. <mare "nagspook, kwelgees"; *korrespondeer met Eng. (night)mare*)

naboots ww., het nageboots. (<Ndl. nabootsen; bootsen: <O.fra. bocer "beeldhouwerk in reliëf maak" <boce "bult, verhewenheid")

Die skrywer het volgens bogenoemde riglyne etimologieë saamgestel vir alle items in die Buro se databank vanaf na tot Nazi. Daar is weer eens ondervind dat etimologiese inligting selde onveranderd uit 'n ander bron oorgeneem kan word, veral omdat bronne van mekaar verskil. Dit is dan dikwels nodig om verdere navorsing te doen. Bronne wat met vrug gebruik is, is verklarende woordeboeke, nl. *Oxford English Dictionary*, *Woordenboek der Nederlandsche Taal*, *Webster's* (verskillende uitgawes), *American Heritage Dictionary* (verskillende uitgawes), *A Latin Dictionary* van Lewis en Short, *Nouveau Petit Larousse*; tweetalige woordeboeke, veral Engels / Nederlands, Nederlands / Engels; etimologiese woordeboeke, nl. De Vries (1971), De Vries / De Tollenaere (1991), *Van Dale Etimologisch Woordenboek* (1991), Boshoff en Nienaber (1967), Onions se *The Oxford Dictionary of English Etymology* (1969), *Dictionary of South African English on Historical Principles* (1991) en ander bronne, veral Scholtz (1965, 1972, 1985) en Smith (1962). Konsultasie met kenners van vreemde tale het ook gereeld plaasgevind waar onsekerheid bestaan het.

Daar is heelwat leenwoorde uit Nama wat met *na-* begin en wat eiesoortige probleme vir die leksikograaf meebring. Baie van die woorde word met een van vier tipes suigklanke uitgespreek en benewens die *na*-spelling kom 'n *gh-*, *ka-*, *ha-* en *tna*-spelling dikwels ook voor. Waardevolle mededelings deur prof. J.A. Engelbrecht van Pretoria ten opsigte van uitspraak en herkoms is in

die databank opgeneem. Die dentale (/), alveolêre (=), palato-alveolêre (!) en laterale (//) suigklanke word in die skrifbeeld in die etimologieveld aangedui omdat daar geen fonetiese tekens daarvoor bestaan nie:

na [na:] (*met aanvangsuigklank*) b.n.w. (<Nama !na) Lekker of baie lekker.

Versnelling en die bereiking van doelwitte word 'n steeds dringender aspek vir die voortbestaan van die WAT. Alhoewel die redaksie besef dat die toepassing van bogenoemde beginsels waarskynlik die gebruiker se verwagtings sal bevredig, maak die vereistes van versnelling en die bereiking van doelwitte dit moeilik uitvoerbaar.

6.3 Vennootskap in 'n Etimologiewoordeboek

Die Buro het in 1992 begin om met die medewerking van buite-outeurs satelietpublikasies uit te gee. *Woordkeusegids: 'n Kerntesourus van Afrikaans* het in 1992 verskyn en *Afrikaanse Idiome en ander Vaste Uitdrukkings* in 1994.

Dit wil voorkom asof die enigste wyse waarop die Buro 'n bydrae kan lewer om die opvallende leemte in die Afrikaanse leksikografie te vul, die medewerking aan 'n etimologiewoordeboek is. In so 'n geval sal die woordeboek deur buite-outeurs saamgestel word, terwyl die Buro betrokke sal wees by die beplanning, administratiewe begeleiding en setwerk.

Sodanige woordeboek sou binne drie tot vyf jaar voltooi kan word indien 'n woordeboek van ongeveer 20 000 inskrywings aangepak word. Dit sou ongeveer vier maal die omvang van Boshoff en Nienaber se *Afrikaanse Etimologieë* wees. Die groot voordeel van so 'n projek bo deurlopende etimologie in die WAT is dat gebruikers binne vyf jaar oor etimologiese inligting van A tot Z in Afrikaans sal beskik in plaas daarvan om te wag totdat die WAT tot by Z voltooi is. Die inligting vervat in die etimologiese woordeboek sou ook in toekomstige WAT-dele gebruik kon word.

7. Ten slotte

Die huidige redaksie van die WAT is nie onwillig om deurlopend etimologieë aan te bied nie. Die moontlikheid is inderdaad ondersoek deur middel van 'n proefloop toe etimologieë, met gebruikmaking van sekondêre bronne, saamgestel is vir die deel van die Afrikaanse leksikon wat strek vanaf na tot Nazi.

Versnelling en die bereiking van doelwitte het egter so 'n belangrike deel van die WAT se beplanning geword dat dit eenvoudig nie haalbaar is nie. Deur 'n etimologiewoordeboek as 'n satellietprojek tot stand te bring wil die Buro van die WAT egter 'n bydrae lewer tot die vul van 'n leemte in die Afrikaanse leksikografie en ook iets tereg laat kom van die ideaal van sy eerste hoofredakteur.

Verwysings

- Botha, W.F.** 1989. Brief. *Die Burger*, 6 Desember 1989: 12.
- Combrink, Johan.** s.j. *Waar kom daardie woord vandaan?* S.p.: Suid-Afrikaanse Uitsaaikorporasie.
- De Klerk, W.J. en F.A. Ponelis.** 1976. *Gedenkbundel H.J.J.M. van der Merwe*. Pretoria: Van Schaik.
- Eksteen, L.C.** 1976. 'n Woordeboek van Afrikaans op historiese grondslag. De Klerk, W.J. en F.A. Ponelis (Reds.). 1976: 67-75.
- Feinauer, I.** 1996. Die negende deel van die Woordeboek van die Afrikaanse Taal. *Lexikos* 6: 233-271.
- Gouws, R.H.** 1989. *Leksikografie*. Kaapstad: Academica.
- Gouws, R.H. en F.A. Ponelis.** 1992. The development of Afrikaans and the Lexicographical Tradition. Zgusta, L. (Red.). 1992: 77-104.
- Gouws, R.H.** (moet nog verskyn) *Linguistiese Teorie, lexicografiese Praxis und das Woordeboek van die Afrikaanse Taal*. *Festschrift H.E. Wiegand*.
- Grütter, W.** 1989. Briewe. *Die Burger*, 25 November 1989: 14; 13 Desember 1989: 12; 29 Desember 1989: 10.
- Hartevelt, P.** 1989. Brief. *Die Burger*, 22 Desember 1989: 8.
- Hartevelt, P.** 1990. Brief. *Die Burger*, 16 Januarie 1990: 10.
- Hauptfleisch, D.C.** 1989. Brief. *Die Burger*, 22 Desember 1989: 8.
- Raidt, E.H.** 1993. Die agste deel van die *Woordeboek van die Afrikaanse Taal*. *Lexikos* 3: 215-226.
- Malherbe, D.F.** s.j. (1925). *Afrikaanse spreekwoorde en verwante vorme*. Bloemfontein: Nasionale Pers.
- Scholtz, J. du P.** 1972. *Afrikaans-Hollands in die agtiende eeu*. Kaapstad: Nasou.
- Scholtz, J. du P.** 1985. Afrikaanse woorde en uitdrukkings '— eiegoed of erfgoed?'. *Tydskrif vir Geesteswetenskappe* 25(4): 235-290.
- Smith, J.J.** 1962. *Op ons taalakker*. Pretoria: J.L. van Schaik.
- Van Schalkwyk, D.J.** (Red.). 1994. *Woordeboek van die Afrikaanse Taal*. Deel IX. Stellenbosch: Buro van die WAT.
- Zgusta, L.** (Red.). 1992. *History, Languages and Lexicographers*. Tübingen: Max Niemeyer.

Leksikografiese behoeftevervulling*

Rufus H. Gouws, *Departement Afrikaans en Nederlands,
Universiteit van Stellenbosch, Suid-Afrika*

Abstract: Meeting Lexicographic Needs. A multilingual and multicultural South Africa has its own lexicographic needs and the fulfilment of these needs demands thorough planning and the employment of a sound and effective organisational theory. This paper focuses on a few priorities to meet the South African lexicographic needs. The contents of a dictionary as well as the presentation of the information should be directed at a well-defined target user. Not only the linguistic needs of this target user but also his research skills should have an influence on the planning and compilation of any specific dictionary. Using a metalexicographic model this paper discusses different aspects of interest in determining priorities for the fulfilment of lexicographic needs. Lexicographers have to be familiar with components like dictionary use, dictionary criticism, the history of dictionaries and a general theory of lexicography. The importance of planning and a sound training is emphasised. Training should not only be aimed at lexicographers but also at dictionary users. It is easier to achieve the meeting of lexicographic needs in an environment where a well-established dictionary culture exists.

Keywords: DICTIONARY, DICTIONARY CRITICISM, DICTIONARY CULTURE, DICTIONARY TYPOLOGY, DICTIONARY USE, FULFILMENT OF NEEDS, GENERAL THEORY OF LEXICOGRAPHY, HISTORY OF DICTIONARIES, LEXICOGRAPHY, METALEXICOGRAPHY, USER-FRIENDLINESS

Opsomming: 'n Meertalige, multikulturele Suid-Afrika het eiesoortige leksikografiese behoeftes en die vervulling hiervan verg indringende beplanning asook die aanwending van 'n doeltreffende organisasieteorie. In hierdie artikel word aandag gegee aan enkele prioriteite vir die vervulling van die Suid-Afrikaanse leksikografiese behoeftes. Daar word van die standpunt uitgegaan dat die inhoud van 'n woordeboek sowel as die aanbieding van inligting gerig moet wees op 'n duidelik omskrewe teikengebruiker. Nie net die linguïstiese behoeftes van hierdie gebruiker nie maar ook sy naslaanvaardighede speel 'n rol by die beplanning van 'n spesifieke woordeboek. Aan die hand van 'n metaleksikografiese model word verskillende aspekte bespreek wat 'n invloed op die vasstelling van prioriteite in die vervulling van leksikografiese behoeftes kan hê. Leksikograwe en voornemende leksikograwe moet vertrouwd wees met woordeboekgebruik, woordeboekkritiek, die geskiedenis van woordeboeke sowel as 'n algemene leksikografieteorie. Die waarde van beplanning en opleiding word deurgaans sterk beklemtoon. Opleiding moet nie net op woordeboekopstellers gerig wees nie maar ook op woordeboekgebruikers. Leksikografiese behoeftevervulling word makliker bereik in 'n omgewing waar 'n woordeboekkultuur reeds gevestig is.

Hierdie referaat is gelewer by die seminaar leksikografie as 'n finansiële bate in 'n veeltalige Suid-Afrika, gehou by die Buro van die Woordeboek van die Afrikaanse Taal, Stellenbosch, 12 April 1996.

Lexikos 6 (AFRILEX-reeks/series 6: 1996): 171-183

Slutelwoorde: ALGEMENE LEKSIKOGRAFIETEORIE, BEHOEFTEVERVULLING, GEBRUIKERSVRIENDELIKHEID, GESKIEDENIS VAN WOORDEBOEKE, LEKSIKOGRAFIE, METALEKSIKOGRAFIE, WOORDEBOEK, WOORDEBOEKGEBRUIK, WOORDEBOEKKRITIEK, WOORDEBOEKKULTUUR, WOORDEBOEKTIPOLOGIE

1. Inleidend

Die leksikografie het sy ontstaan in 'n preteoretiese omgewing gehad. Die samestelling van die eerste woordeboeke is nie aan die hand van teoretiese modelle uitgevoer nie maar is gedoen ter bevrediging van die praktiese behoeftes van werklike gebruikers. Een van die belangrikste kenmerke van die leksikografiese ontwikkeling van die afgelope twee dekades is 'n leksikografiese perspektiefverandering wat gelei het tot die vestiging van 'n sterk teoretiese komponent in die leksikografie en tot 'n aktiewe wisselwerking tussen teorie en praktyk. In hierdie "Goue Eeu van die leksikografie", vgl. Ilson (1985), is die tradisionele leksikografiese praktyk gehandhaaf maar daarnaas en aanvullend daartoe het die metaleksikografie sy plek as volwaardige dissipline ingeneem. 'n Ondersoek na 'n volgorde vir leksikografiese behoeftevervulling moet gevolglik aan sowel die teoretiese as die praktiese aspekte aandag gee.

Oor die status van die leksikografie is daar al baie en baie uiteenlopende uitsprake gemaak. Of die leksikografie as 'n wetenskap, 'n vaardigheid, 'n spesiale tegniek al dan as 'n kunsvorm beskou moet word, is vir baie mense geen onproblematiese besluit nie, vgl. Wiegand (1989: 246 e.v.). Vir Wiegand (1989: 248) is die leksikografie 'n wetenskaplike praktyk want die werklike leksikografiese bedrywighede is nie daarop gerig, soos in die geval van 'n wetenskap, om teorieë te formuleer of te toets nie, maar die een gemeenskaplike kenmerk van alle leksikografiese bedrywighede is hulle doelstelling dat woordeboeke as gebruiksvoorwerpe tot stand gebring moet word (Wiegand 1989: 251). Leksikografie is 'n praktyk wat daarop gerig is om woordeboeke tot stand te bring sodat 'n ander praktyk, naamlik die kulturele praktyk van woordeboekgebruik, moontlik gemaak kan word. Die suksesvolle uitvoering van die leksikografiese praktyk word medebepaal deur invloede uit die teoretiese taalkunde. Taalteorieë is nie spesifiek op die leksikografiese praktyk gerig nie en oor jare heen moes leksikograwe bestaande taalteorieë analiseer en interpreteer met die oog op die toepasbaarheid daarvan vir die leksikografie. Algaande het die behoefte ontstaan aan 'n algemene leksikografieteorie wat aan praktiserende leksikograwe die nodige teoretiese leiding kon bied.

Die werk wat oor veral die afgelope twee dekades in hierdie verband gedoen is, het gelei tot die ontwikkeling van die metaleksikografie as 'n volwaardige komponent binne die breër leksikografiese terrein. Volgens Wiegand (1984: 14 e.v.) moet 'n algemene leksikografiese teorie op 'n sistematiese manier 'n verwerking en verduideliking bied ter motivering van die kennis wat nodig is om leksikograwe toe te rus om hulle werk op 'n gepaste manier en so goed as

moontlik te doen. In die leksikografiese praktyk moet 'n onderskeid gemaak word tussen linguistiese en saakleksikografie. Tot die laasgenoemde kategorie behoort ensiklopedieë en vakwoordeboeke wat nie primêr gerig is op 'n weergawe van taalkundige inligting nie. Die eerste kategorie bevat gewone verklarende en vertalende asook ander woordeboeke wat die taalkundige bewerking van lemmas as hoofdoel het. Om verwarring te voorkom wanneer daar na die verskillende komponente van die leksikografie verwys word, maak Wiegand (1989: 253) 'n onderskeid tussen linguistiese leksikografie, dit is die praktiese komponent, en woordeboeknavorsing, dit is die teoretiese komponent. Hartmann (1996: 238) maak 'n onderskeid tussen woordeboek skryf en woordeboeknavorsing.

Bestaande metaleksikografiese modelle is primêr daarop gerig om 'n verhoging in die gehalte van woordeboeke tot gevolg te hê. Alhoewel woordeboekprojekte vandag nog soms met 'n redelike mate van sukses in 'n omgewing geïsoleerd van teoretiese invloed tot stand kan kom, kan dit algemeen aanvaar word dat 'n deeglike teoretiese begronding een van die eerste prioriteite van enige leksikograaf of voornemende leksikograaf moet wees. Hier speel metaleksikografiese bevoegdheid 'n wesenlike rol. Om vertrouwd te raak met die eise van die metaleksikografie, kan as 'n eerste stap in die volgorde van leksikografiese behoeftevervulling gesien word.

2. Die Suid-Afrikaanse leksikografiese toneel

Die meertalige en multikulturele Suid-Afrikaanse samelewing bring 'n eiesoortige stel leksikografiese eise mee. Dié belangrikste aspek van toekomstige leksikografiese behoeftevervulling in die Suid-Afrikaanse konteks is deeglike beplanning en die inwerkingstelling van 'n sogenaamde organisasieteorie. Ongelukkig was dit tot dusver geen opvallende kenmerk van die meeste Suid-Afrikaanse leksikografiese bedrywighede nie.

Die beplanning van leksikografiese projekte in 'n meertalige en multikulturele samelewing kan vanuit verskillende vertrekpunte gedoen word. 'n Pan-organisatoriese benadering kyk oorkoepelend na die leksikografiese prentjie en doen voorstelle aan die hand. 'n Taalspesifieke benadering daarenteen vereis dat daar afsonderlike leksikografiese beplanning vir elke individuele taal gedoen word. In die eerste geval sal algemene leksikografiese neigings die opdragte domineer terwyl taalspesifieke behoeftes in die laasgenoemde geval die sterkste aan die orde kom. Oorvleueling is 'n probleem wat baie maklik kan ontstaan in die geval van die taalspesifieke model. Die bergpredikasiesindroom, dit wil sê waar die linkerhand nie weet wat die regterhand doen nie, lei maklik daartoe dat taal X byvoorbeeld besluit om 'n vertalende woordeboek met die taalpaar X en Y te maak, terwyl taal Y terselfdertyd ook op so 'n woordeboek besluit. Die geskeie werkswyses sou dan baie maklik leksikografiese teenproduktiwiteit in die hand kon werk. Daarom is dit belangrik dat die twee benaderings nie onderling uitsluitend moet wees nie.

'n Onderskeid wat gemaak moet word, is dié tussen sake van algemene belang en sake wat slegs taalspesifieke waarde het. In die beplanning rondom sowel die linguistiese as die saakleksikografie in 'n meertalige bestel behoort die swaartepunt op die taalspesifieke benadering te wees, maar die indiwiduele tale mag nie slegs in isolasie werk nie. Ruimte moet geskep word vir beplanning van gemeenskaplike projekte. Uit die navorsing van Wiegand wat tot die formulering van sy metaleksikografiese beginsels gelei het, blyk die algemene en nie-taalspesifieke toepassingswaarde van sy teorie baie duidelik. Woordeboeknavorsing is nie gerig op die taalspesifieke nie maar stel die fokus op 'n algemene leksikografieteorie. Dit is geen geheim dat leksikograwe 'n baie klein deelversameling van die spesie homo sapiens uitmaak en dat die hoop op inflasie binne hierdie subkultuur gering is. Daarom is dit des te belangriker dat die praktiese leksikografie van indiwiduele tale wel deur kundiges en deskundiges uit die betrokke taalgemeenskappe bedryf moet word, maar dat navorsing en kundigheid oor gedeelde waardes nie by wyse van oorvleueling deur lede van elke gemeenskap gedoen moet word nie. Dit geld veral ook die Suid-Afrikaanse leksikografiese toneel. Sentra moet bestaan waar taalspesifieke leksikografiese projekte uitgevoer word. Die praktiese werk wat in hierdie sentra gedoen word, moet steeds 'n wisselwerking vertoon met die werk wat op die metaleksikografiese terrein gedoen word. Praktiserende leksikograwe se dagtaak maak dit vir hulle onmoontlik om volledig vertrouwd te bly met die heersende metaleksikografiese strominge. Elke projek moet egter blootgestel word aan bevindinge uit die metaleksikografiese navorsing wat vir die spesifieke projek van waarde is. Gevolglik is die bestaan van 'n sentrum of sentra van metaleksikografiese kundigheid en die toegang van indiwiduele woordeboekprojekte tot so 'n sentrum een van die werklike prioriteite in die proses van leksikografiese behoeftevervulling in Suid-Afrika.

3. Tipologiese verskeidenheid versus gebruikersbehoeftes

Die taalverskeidenheid en die behoefte aan verskillende woordeboeksoorte het reeds baie bygedra tot die omvang van Suid-Afrikaanse leksikografiese projekte. Daar is sekere wanbalanse tussen tale, tussen woordeboeksoorte en tussen bestaande leksikografiese werkswyses. Riglyne is nodig sodat die strewe na 'n bevredigende woordeboekversameling so gou en so goed as moontlik verwesenlik kan word. Om dit moontlik te maak, moet daar vir elke Suid-Afrikaanse taal 'n indringende behoeftestudie gedoen word. Alhoewel so 'n studie die behoeftes van die leksikografiepraktyk moet aandui, val dié studie onmiddellik binne die oorvleuelingsterrein van teorie en praktyk.

Die vasstelling van leksikografiese behoeftes mag nie gedoen word met 'n berekening soos die volgende nie:

Dit is die volledige tipologiese verskeidenheid
 - *Dit het ons*
 = *Dit het ons nodig.*

Leksikografiese behoeftes is nie die verskil tussen die beskikbare woordeboeke van enige gegewe taal en dié van die volledige tipologiese spektrum nie. Die toepassing van so 'n benadering kan baie maklik lei tot 'n konflik tussen woordeboek en gebruiker omdat die gebruiker, ondanks die tipologiese uitgebreidheid van 'n gegewe taal se leksikografiese praktyk, nie 'n bevrediging van sy / haar werklike behoeftes vind nie. 'n Taal wat oor slegs enkele woordeboeke beskik, mag in 'n hoër mate die leksikografiese behoeftes van die spesifieke taalgemeenskap bevredig as wat die geval is met 'n taal wat oor 'n uitgebreide tipologiese verskeidenheid beskik. Leksikografiese behoeftevervulling mag nie primêr gerig wees op die produksie van woordeboeke ter vervollediging van die teikentaal se tipologiese spektrum nie; dit moet gerig wees op die produksie van woordeboeke ter bevrediging van die teikentaalgemeenskap se leksikografiese behoeftes. Dit mag byvoorbeeld meebring dat 'n eerste prioriteit vir 'n gegewe taal die produksie van 'n woordeboeksoort is wat reeds in die betrokke taal bestaan maar wat nie aan die nodige vereistes voldoen om die gebruikersbehoefte te bevredig nie. Die produksie van enige nuwe woordeboek asook die heruitgawe van bestaande woordeboeke moet uit 'n gebruikersgedrewe behoefte volg. Die tyd is verby vir die publikasie van woordeboeke ter wille van woordeboeke.

Dit is nie om dowe neute dat Wiegand in sy raamwerk vir die metaleksikografie navorsing oor woordeboekgebruik tot een van die hoofkomponente verhef nie, vgl. die bespreking wat volg oor die hoofkomponente van 'n metaleksikografiese model. Hierdie komponent is juis daarop gerig om dinge vir die gebruiker makliker te maak in die benutting van woordeboeke en die ontsluiting van die verskillende inligtingstipes. Volgens Kühn (1989: 112) moet die gebruik van 'n woordeboek as 'n kommunikatiewe handeling begryp word. So 'n siening het regstreekse implikasies vir die leksikograaf. Die sentrale funksie van woordeboeke is dat hulle as naslaanwerke 'n kulturele praktyk bevorder, vgl. Reichmann (1988: 395). Dit is 'n kultuurpedagogiese rol waardeur sowel linguistiese as saakwoordeboeke die werk van die skool voortsit en uitbou, of dikwels selfs vervang; selfs al bereik hulle nie hulle doel ten volle nie. Wat van bepalende belang is, is dat die naslaanhandeling suksesvol uitgevoer moet word. Dit is slegs moontlik as die leksikografie en die teikengemeenskap in harmonie met mekaar verkeer, vgl. Hausmann (1989: 13). Hierdie harmonie word dikwels bedreig deurdat die hoë inligtingsdigtheid en die komplekse aanbiedingstruktuur van woordeboeke die gewone gebruiker van die woordeboek vervreem. In 'n gebruikersgedrewe leksikografie kan hierdie konfliktsituasie ontlont word deur die vestiging van 'n benadering van gebruikersvriendelikheid en van 'n woordeboekkultuur. Volgens Hausmann (1989: 13) verwys gebruikersvriendelikheid in hierdie konteks na 'n aanpassing van die leksikografie by die gemeenskap, terwyl woordeboekkultuur te make het met die aan-

passing van die gemeenskap by die leksikografie. Hierdie twee begrippe moet as swaartepunte binne 'n gebruikersgedrewe leksikografie beskou word. Dit plaas 'n verantwoordelikheid op sowel die leksikograaf as die gemeenskap.

Ter wille van gebruikersvriendelikheid moet die leksikograaf nie net vertrouwd wees met die spesifieke taalkundige behoeftes van 'n vooraf bepaalde teikengebruikersgroep nie maar hy / sy moet ook rekening hou met dié groep se naslaanvaardighede. Om die sukses van 'n woordeboek te verseker, moet die leksikograaf die inligting op so 'n manier aanbied dat die verlangde inligtingsoordrag so maklik as moontlik kan plaasvind en dat die teikengebruiker maklike toegang tot die verstrekte inligting het.

Die basis van enige gesprek oor leksikografiese behoeftevervulling mag nie fokus op 'n standpuntinname oor die produksie van woordeboeke nie. Die vertrekpunt van so 'n gesprek moet binne 'n metaleksikografiese raamwerk wees. Dit gaan nóg oor tipologiese uitbreiding nóg oor 'n uitbouing van óf die makro- óf die mikrostruktuur van bestaande woordeboeke. Die vasstelling van leksikografiese behoeftes sowel as van 'n volgorde vir die vervulling daarvan moet volgens 'n gebruikersgedrewe model gedoen word en wel aan die hand van objektiewe kriteria wat volg uit primêre woordeboeknavorsing.

As daar vir die Suid-Afrikaanse situasie gekyk moet word na 'n volgorde van leksikografiese behoeftevervulling, is dit belangrik om aandag te gee aan sowel taalspesifieke as algemene aspekte. Riglyne hiervoor word vervolgens gebied.

4. Taalspesifieke aspekte

Alhoewel dit belangrik is dat daar vir elke Suid-Afrikaanse taal 'n behoorlike opname gemaak moet word van beskikbare woordeboeke, vergelyk in hierdie verband die werk wat reeds gedoen is in Alberts en Mtintsilana (1988), moet Hartmann (1989: 103) se uitspraak in gedagte gehou word, naamlik dat die ontwerp van enige woordeboek voorafgegaan moet word deur 'n analise van gebruikersbehoefte. Dit geld nie net die ontwerp van die enkele woordeboek nie maar ook die beplanning van enige breër leksikografiese netwerk. 'n Wesenlike vertrekpunt in die vasstelling van 'n volgorde van leksikografiese behoeftevervulling vir die Suid-Afrikaanse tale is dat daar vir elkeen van dié tale indringend aandag gegee moet word aan 'n empiriese ondersoek na die werklike leksikografiese behoeftes van die betrokke taalgemeenskap. Omdat hierdie ondersoek ten beste binne 'n teoretiese raamwerk kan geskied, is dit belangrik dat beplanners vertrouwd moet wees met die heersende strominge in die metaleksikografie. Vir hierdie doel is dit nodig dat aandag gegee moet word aan die kriteria wat gestel word in 'n algemene leksikografieteorie. In hierdie verband bied die teoretiese werk van H.E. Wiegand nuttige leiding aangesien die formulering van 'n algemene leksikografieteorie een van die hoofkomponente van sy metaleksikografiese beskouing is.

4.1 'n Metaleksikografiese raamwerk

In 'n gesprek oor die metaleksikografie moet daar geen twyfel bestaan oor die status en funksie van die teoretiese riglyne wat deur die metaleksikografiese raamwerk aan die praktiserende leksikograaf voorgeskryf word nie. Volgens Geeraerts (1989: 287) is die beginsels van leksikografie eerder vergelykbaar met dié van die ingenieurswese as met dié van die wette van die fisika. Die wette van die fisika spesifiseer die beginsels wat die struktuur en optrede van die fisiese werklikheid beheer. Die beginsels van die leksikografie is egter nie taalwette nie. Hulle spesifiseer nie die fundamentele reëls wat die linguïstiese struktuur en linguïstiese optrede beheer nie. Hulle spesifiseer die onderliggende aksiomas van sekondêre strukture en sekondêre optrede. Die struktuur van woordeboeke en die optrede van leksikograwe is slegs indirek verwant aan die struktuur van taal en die optrede van taalgebruikers; soortgelyk aan die manier waarop die beginsels van ingenieurswese verwant is aan die reëls van die fisika. Ingenieurswese is 'n doelgerigte menslike aktiwiteit wat aandag moet gee aan onder andere die wette van die materiële wêreld. Net so is leksikografie 'n doelgerigte menslike aktiwiteit waarvoor die beginsels van taal een van 'n reeks parameters is wat die uiteindelige vorm van 'n woordeboek bepaal.

Wiegand (1984: 14 e.v.) onderskei vier hoofafdelings waaruit die metaleksikografie bestaan, te wete navorsing oor woordeboekgebruik, navorsing oor die geskiedenis van leksikografie, navorsing oor woordeboekkritiek en navorsing oor 'n algemene leksikografiëteorie. Onder 'n algemene leksikografiëteorie sluit Wiegand aspekte in soos die volgende: die doel van 'n woordeboek, organisasieteorie en die beplanning van woordeboekprojekte, leksikografiese taalnavorsing wat fokus op die verkryging, toetsing en verwerking van data, asook leksikografiese taalbeskrywing wat woordeboektipologie en die teksteorie vir leksikografiese tekste aan die orde stel. Behoeftebepaling en die volgorde vir die vervulling van leksikografiese behoeftes van individuele tale kan met sukses binne hierdie raamwerk van Wiegand gedoen word.

4.1.1 Navorsing oor woordeboekgebruik

Oor hierdie aspek van die metaleksikografie word daar nie nou volledig gepraat nie. Vir enige leksikografiese projek maar ook vir die breër leksikografiese bedryf van enige individuele taal moet navorsing oor woordeboekgebruik as een van die belangrikste vertrekpunte dien. Dit is veral die posisie van die gebruiker, sy behoeftes en sy vaardighede wat hier ter sprake kom. Vergelyk Hartmann (1996: 242 e.v.) se siening dat die beslissende rede vir die bestaan van woordeboeke, die behoefte aan inligting is. Die lokus van hierdie naslaanbehoefte is die woordeboekgebruiker met sy bevoegdheids- en vooroordele. Dit moet nie nodig wees om te raai waarvoor 'n bepaalde werk deur 'n

anonieme koper gebruik moet kan word nie. Daar moet empiries vasgestel word vir watter funksies dit saamgestel is om óf 'n voordelige óf 'n onbevredigende gebruik te hê. Verskillende metodes kan gevolg word om hierdie vasstelling suksesvol af te handel.

Die manier waarop woordeboekontsluitingsvaardighede aangeleer moet word, staan sentraal in die navorsing oor woordeboekgebruik. Vir die Suid-Afrikaanse situasie is dit van die hoogste belang.

4.1.2 Navorsing oor die geskiedenis van woordeboeke

Die meeste woordeboeke het voorgangers en ook navolgers, vgl. Hartmann (1986: vii). Daarom is 'n goeie begrip van die oorsprong en ontwikkeling van woordeboeke belangrik vir die vooruitgang van die leksikografie. Dit moet aanvaar word dat min woordeboeke sonder enige steun op vroeëre woordeboeke tot stand kom. As dit binne perke gedoen word, is daar in beginsel geen fout hiermee te vind nie. Waarmee daar wel fout te vind is, is wanneer woordeboekgebruikers mislei word deurdat hulle onder 'n verkeerde indruk gebring is aangaande die aard, omvang, selfstandigheid en funksies van 'n spesifieke woordeboek. Hierdie soort woordeboekkriminaliteit, vgl. Hausmann (1989: 99), is dikwels daaraan te wyte dat leksikograwe nie erkenning aan hulle bronne of 'n duidelike uitspraak oor die spesifieke woordeboeke maak nie.

In die vasstelling van 'n volgorde vir leksikografiese behoeftevervulling in Suid-Afrika moet daar aandag gegee word aan die invloed wat bestaande woordeboeke steeds op nuwe woordeboekprojekte uitoefen. Dolezal (1986: 47, 54) wys daarop dat die geskiedenis van die Engelse leksikografie bewys dat die "Engelse woordeboek" gesien kan word as 'n enkele bewerkte en herbewerkte teks. In die Suid-Afrikaanse leksikografie het sekere woordeboeke en woordeboeksoorte, en by name die vertalende woordeboeke met Afrikaans en Engels as taalpaar, 'n dominante invloed gehad. Vergelyk in hierdie verband Gouws (1986), (1989), (1990) en Gouws en Ponelis (1992). Hierdie invloed het uitgegaan van 'n taalpaar waarvan albei die lede die enigste twee amptelike tale van die land was. Die amptelike status van hierdie tale het noodwendig daartoe gelei dat hulle oor die algemeen beter beheers is as ander nie-amptelike tale. Die hoë inligtingsdigtheid van hierdie woordeboeke het regstreeks verband gehou met die veronderstelde vertroudheid van die teikengebruikers met albei die behandelde tale. Sulke woordeboeke was dikwels uiters ontoeganklik vir die vermeende teikengebruiker. Dit was dié gevolg van 'n foutiewe siening van die gebruiker se behoeftes en naslaanbevoegdheids. Navorsing oor die geskiedenis van woordeboeke is nodig om die voortsetting van hierdie praktyk te voorkom en om te sorg dat sulke woordeboeke nie as rolmodelle vir ander vertalende woordeboeke dien nie; veral waar die teikengebruikers 'n mindere vertroudheid met albei die behandelde tale het.

Ook in die verklarende leksikografie moet rolmodelle met omsigtigheid

gekies word. Die oorheersende rol wat Engels tans in Suid-Afrika speel, mag nie daartoe lei dat Britse en Amerikaanse Engelse verklarende woordeboeke sonder meer as voorbeeld van gangbare leksikografie vir die Suid-Afrikaanse tale gesien word nie. Die eie aard van die teikengebruikers van Suid-Afrikaanse woordeboeke se leksikografiese behoeftes moet lei tot 'n vasstelling van die spesifieke tipologiese kategorie maar ook tot die vasstelling van die tersaaklike mikrostrukturele bewerking wat in plaaslike woordeboeke aangebied moet word.

Van die uiterste belang vir die Suid-Afrikaanse tale is dat alle gevestigde en voornemende leksikograwe nogmaals vanuit 'n historiese perspektief moet kyk na die ontwikkeling van 'n spesifieke taal se leksikografiepraktyk. Uit hierdie oorsig moet 'n herwaardering van bestaande projekte volg en 'n vasstelling van tipologiese, subtipologiese maar ook aanbiedings- en inligtingsbehoefte. Dit is 'n voorvereiste vir die sinvolle voortgang van enige taal se leksikografiese praktyk.

4.1.3 Woordeboekkritiek

Navorsing oor woordeboekkritiek word met reg binne die Wiegand-model as 'n terrein van die metaleksikografie gesien. Woordeboekkritiek is een van die min vorme van inspraak wat linguïes, metaleksikograaf en algemene woordeboekgebruiker in die leksikografiepraktyk het. Naas uitsprake oor die teoretiese en linguïstiese gehalte van 'n woordeboek is dit juis ook gebruikaspekte en die mate waarin die woordeboek teikengebruiker- en behoeftegerig is wat in woordeboekkritiek aan die orde kom. Nie net moet die leksikograaf of redaksie van 'n enkele woordeboek hulle by die herbewerking van die teks aan die kritiek op voorafgaande dele steur nie, maar leksikografiese beplanners moet die volgorde van 'n spesifieke taal se leksikografiese behoeftevervulling vasstel met inagneming van die uitsprake wat in die woordeboekkritiek na vore gekom het.

Hier kan twee voorbeelde genoem word waar woordeboekkritiek 'n regstreekse invloed op die leksikografiese praktyk gehad het. Die geskiedenis van die WAT bevat talle voorbeelde van ernstige kritiek op sowel die werkswyse as die inhoud van hierdie woordeboek. Kritici het hulle, soms ten regte maar ook soms ten onregte, veral sterk uitgespreek teen die aanbod in dele VI en VII. In deel VIII se toeligting word verwys na 'n deeglike strategiese beplanning wat vir die Buro van die WAT gedoen is en wat tot gevolg gehad het dat die redaksionele stelsel ingrypend hersien is. Dit word genoem dat die stelsel aangepas is by die hedendaagse stand van die taalkunde en die leksikografie en dat daar in die verfyning van die stelsel sover moontlik uitvoering gegee is aan die kommentaar en voorstelle deur resensente van deel VII. In die inleiding van deel IX word ook na die strategiese beplanning verwys en die feit dat verskeie eksterne kundiges betrek is by dié hersiening van die redaksionele stelsel.

Deel VIII en veral deel IX kan met alle reg beskou word as die beste leksikografiese werk wat tot nog toe deur die Buro voortgebring is. Aandag aan woordeboekkritiek in die beplanning van sy werk het tot 'n gehalteverhoging van die WAT gelei. Die redaksie het dit as een van 'n reeks prioriteite gesien en die toepassing daarvan lei tot 'n hoër graad van leksikografiese behoeftevervulling by hulle teikengebruikers.

In die internasionale leksikografiese ontwikkeling het die spesifieke behoeftes van die taalaanleerder gelei tot 'n tipologiese uitbreiding wat tot die publikasie van eiesoortige verklarende en vertalende woordeboeke gelei het. Ondanks die toevoeging van aanleerderwoordeboeke tot die internasionale mark het hierdie woordeboeksoort vir baie lank in die Suid-Afrikaanse woordeboekversameling ontbreek. Kritiek is uitgespreek op die afwesigheid van aanleerderwoordeboeke, vgl. byvoorbeeld Otto (1989). Leksikograwe het hulle daaraan gesteur en vandag help woordeboeke soos *Afrikaans / Zoeloe Woordeboek*, *Tweetalige Aanleederswoordeboek / Bilingual Learner's Dictionary* en *Basiswoordeboek van Afrikaans* met die vervulling van die taalaanleerder se leksikografiese behoeftes.

4.1.4 'n Algemene leksikografierteorie

Juis ter wille daarvan dat woordeboeke 'n taalkundige bewerking van die leksikon moet bied, is dit belangrik dat leksikograwe linguisties onderleg moet wees. In sy metaleksikografiese navorsing het Wiegand oor jare heen gepoog om 'n raamwerk te ontwikkel wat as riglyn vir leksikograwe kan dien in die aanbieding van linguistiese inligting. Onderliggend aan die noodsaak vir die formulering van 'n algemene leksikografierteorie was die probleem van praktiserende leksikograwe dat daar geen samehangende uiteensetting bestaan het om leiding te gee oor die keuse en aanbieding van die verskillende taalkundige inligtingstipes wat vir die kennisoordrag van verskillende woordeboeksoorte ter sake was nie. Gevolglik rig so 'n algemene teorie hom op grondliggende aspekte soos die doel van 'n woordeboek, organisasieteorie en die beplanning van woordeboekprojekte, leksikografiese taalnavoring wat fokus op die verkryging, toetsing en verwerking van data, asook leksikografiese taalbeskrywing wat woordeboektipologie en die teksteorie vir leksikografiese tekste insluit. Weer eens volg die vasstelling van 'n volgorde vir leksikografiese behoeftevervulling nie uit 'n stelsel van komplekse beginsels nie maar heel praktiese riglyne word binne 'n algemene leksikografierteorie aangebied.

Binne die Suid-Afrikaanse leksikografiese praktyk kan suksesvolle ontwikkeling grootliks ondersteun word deur vertroutheid met die wese van so 'n teorie. Doelgerigte beplanning staan nogmaals sentraal. Omdat 'n woordeboek van die leksikon verslag moet doen, moet 'n geldige en verteenwoordigende versameling leksikonitems opgeneem en bewerk word. Die korrekte manier van dataversameling asook taalkundig aanvaarbare bewerkingsmeto-

des kenmerk goeie leksikografiese produkte. Die korrekte hantering hiervan moet as voorvereistes vir sukses in enige leksikografiese bedryf gesien word.

5. Algemene navorsing, beplanning en leiding

Dit is belangrik dat leksikografiese beplanning vir elke taal en taalgemeenskap gedoen moet word. Hierdie beplanning moet nie van buite af op 'n spesifieke taal se leksikograwe afgedwing word nie, maar moet, met inagneming van die eiesoortige behoeftes van die spesifieke taal, deur linguïste en leksikograwe van en vir die spesifieke taalgemeenskap gedoen word. Die ideaal is dat elke taal oor 'n leksikografiese infrastruktuur sal beskik wat voorsiening maak vir genoeg praktiese leksikograwe maar ook vir diegene wat op metaleksikografiese vlak met navorsing en beplanning kan help. Die werklikheid van die Suid-Afrikaanse situasie dwing 'n mens egter al gou om na ander moontlikhede te kyk.

Die huidige stand van die Suid-Afrikaanse leksikografie maak dit onmoontlik om van praktiserende leksikograwe te verwag om volledig vertrou te bly met die vinnig ontwikkelende metaleksikografie. Alhoewel leksikograwe alles in hulle vermoë doen om ruimte te laat vir metaleksikografiese invloed op hulle werk, verhinder die omvang van hulle werkslas hulle meestal om op die hoogte van die jongste ontwikkelinge te bly. Dit is vir baie lank ook vererger deur die afwesigheid van 'n samehangende leksikografieteorie. As tussenganger tussen die teoretiese taalkunde en die woordeboekgebruiker was die leksikograaf in die onbenydenswaardige posisie dat hy / sy uit 'n magdom gegewens in die teoretiese taalkunde self moes besluit wat vir 'n spesifieke woordeboek se taalkundige aard en inligtingsoordrag ter sake is. Nie net moes die leksikograaf die keuse maak nie maar die toepassing daarvan het ook sy / haar verantwoordelikheid geword.

Die verskyning van 'n omvattende werk soos *Wörterbücher Dictionaries Dictionnaires. An International Encyclopedia of Lexicography* (Hausmann et al. 1989-1991), aangevul deur boeke wat in die reeks *Lexicographica Series Maior* verskyn asook artikels in toonaangewende tydskrifte soos *Dictionaries, International Journal of Lexicography, Lexicographica* en *Lexikos* het dit vir leksikograwe makliker gemaak om vinniger toegang tot onlangse metaleksikografiese bevindinge te kry. Enige praktiserende leksikograaf behoort dit as 'n belangrike prioriteit in diens van sy / haar eie woordeboekprojek te sien om kennis te neem van hierdie publikasies en self by te dra tot die uitbouing van sy / haar metaleksikografiese kundigheid. Maar steeds moet hulp van 'n algemener aard as 'n wesenlike aanvulling tot dié selfonderrig gesien word.

In hierdie verband is die leksikografiekursusse wat aan verskeie Suid-Afrikaanse universiteite aangebied word 'n tree in die regte rigting. Vir die meeste praktiserende leksikograwe is toegang tot hierdie kursusse nie prakties haalbaar nie. Opleidingskursusse soos dié wat deur die Buro van die WAT aan-

gebied word, is gevolglik van onskatbare waarde.

Opleiding kan as dié fondament vir die vervulling van Suid-Afrikaanse leksikografiese behoeftes gesien word. Hierdie opleiding moet egter deur navorsing ondersteun word en die bestek van die opleiding moet sodanig wees dat dit 'n wye verskeidenheid rolspele betrek. Die dae is verby dat net die leksikograaf opgelei moet word. Om 'n woordeboekkultuur in Suid-Afrika geskep te kry, is dit noodsaaklik dat die gemeenskap voorberei moet word om woordeboeke reg te kan gebruik. Dit bring mee dat onderwysers en onderwysebeplanners oortuig moet word van die waarde wat 'n komponent oor woordeboekgebruik in enige taalleerplan het. Skoliere en studente moet in woordeboekgebruik onderrig word. Hierdie opleiding moet egter ook uitgebrei word na diegene wat nie meer met skool, kollege of universiteit in aanraking is nie maar steeds op woordeboeke moet steun vir die bereiking van kommunikatiewe sukses in hulle daaglikse taalhandelinge. Om woordeboekgebruik aan sulke mense bekend te stel, moet nieregeringsorganisasies betrek word om as deel van gemeenskapsprojekte pasgeletterdes maar ook ander volwassenes in te lei in die woordeboekkultuur. Woordeboeke word wel deur leksikograwe saamgestel maar hulle word deur uitgewers gepubliseer. Ook hier is uitgebreide leksikografiese kennis noodsaaklik. Uitgewers moet met praktiserende en teoretiese leksikograwe saamwerk in die beplanning en herbeplanning van leksikografiese projekte.

Met inagneming van die groot behoefte aan opleiding in die breë Suid-Afrikaanse leksikografiese wêreld is PROLEKS, die *Program vir Leksikografie*, in die Departement Afrikaans en Nederlands van die Universiteit van Stellenbosch tot stand gebring. PROLEKS het leksikografiese navorsing en opleiding as sy belangrikste teikengebiede en wil juis probeer om op hierdie manier 'n rol te speel in die Suid-Afrikaanse leksikografiese behoeftevervulling. PROLEKS is nie 'n taalspesifieke projek nie en poog om al die leksikografiese rolspele by sy opleidingswerksaamhede te betrek.

6. Ten slotte

Die bewerkstelling van 'n voortdurende en produktiewe wisselwerking tussen leksikografiepraktyk en leksikografieteorie moet die basis vorm van leksikografiese behoeftevervulling in Suid-Afrika. Die wete aan die kant van die leksikograaf dat 'n woordeboek 'n praktiese gebruiksinstrument in die hande van 'n spesifieke teikengebruiker is met baie spesifieke behoeftes en vasstelbare naslaanvaardighede moet 'n beslissende invloed hê op sowel die inhoud as die aanbieding van woordeboeke. Metaleksikografiese navorsing wat onder meer lei tot die voortgesette opleiding van leksikograwe en woordeboekgebruikers lei nie net tot gebruikersvriendelike woordeboeke nie maar ook tot die skep van 'n woordeboekkultuur. Dit is die belangrikste sleutel tot die vervulling van leksikografiese behoeftes in Suid-Afrika.

Bibliografie

- Alberts, M. en P. Mtintsilana. 1988. *Leksikografiese ontleding van Afrikaanse en Afrika-tale leksikons*. Pretoria: Raad vir Geesteswetenskaplike Navorsing.
- Dolezal, F. 1986. How Abstract is the English Dictionary? Hartmann, R.R.K. (red.) 1986a:47-56.
- Du Plessis, M. 1993. *Tweetalige Aanleerderswoordeboek. Bilingual Learner's Dictionary*. Kaapstad: Tafelberg.
- Geeraerts, D. 1989. Principles of Monolingual Lexicography. Hausmann, F.J. et al. (Reds.). 1989-1991: 287-296.
- Gouws, R.H. 1986. Afrikaans Lexicography. Hartmann, R.R.K. (Red.). 1986a: 75-83.
- Gouws, R.H. 1989. *Leksikografie*. Kaapstad: Academica.
- Gouws, R.H. 1990. Information Categories in Dictionaries, with Special Reference to Southern Africa. Hartmann, R.R.K. (Red.). 1990: 52-65.
- Gouws, R.H. en F.A. Ponelis. 1992. The Development of Afrikaans and the Lexicographical Tradition. Zgusta, L. (Red.). 1992: 77-104.
- Gouws, R., I. Feinauer en F. Ponelis. 1994. *Basiswoordeboek van Afrikaans*. Pretoria: J.L. van Schaik.
- Harras, G. (Red.). 1988. Das Wörterbuch: Artikel und Verweisstrukturen. *Jahrbuch 1987 des Instituts für deutsche Sprache*. Düsseldorf: 394-408.
- Hartmann, R.R.K. (Red.). 1984. *LEXeter '83 Proceedings*. Tübingen: Max Niemeyer.
- Hartmann, R.R.K. 1986. Preface. Hartmann, R.R.K. (Red.). 1986a: vii-viii.
- Hartmann, R.R.K. (Red.). 1986a. *The History of Lexicography*. Amsterdam: John Benjamins.
- Hartmann, R.R.K. 1989. Sociology of the Dictionary User: Hypotheses and Empirical Studies. Hausmann, F.J. et al. (Reds.). 1989-1991: 103-111.
- Hartmann, R.R.K. (Red.). 1990. *Lexicography in Africa*. Exeter: University of Exeter Press.
- Hartmann, R.R.K. 1996. Lexicography as an Applied Linguistic Discipline. Hartmann, R.R.K. (Red.). 1996a: 230-244.
- Hartmann, R.R.K. (Red.). 1996a. *Solving Language Problems*. Exeter: University of Exeter Press.
- 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.). 1989-1991. *Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography*. Berlyn: W. de Gruyter.
- Ilson, R. (Red.). 1985. *Dictionaries, Lexicography and Language Learning*. Oxford: Pergamon Press.
- Kotzé, E. en P. Wela. 1991. *Afrikaans / Zoeloe-Woordeboek*. Kaapstad: Tafelberg.
- Kühn, P. 1989. Typologie der Wörterbücher nach Benutzungsmöglichkeiten. Hausmann, F.J. et al. (Reds.). 1989-1991: 111-127.
- Otto, A.N. 1989. *Kriteria vir 'n Afrikaanse aanleerderwoordeboek*. Ongepubliseerde D.Litt-proefskrif. Universiteit van Stellenbosch.
- Reichmann, O. 1988. Resümee der Tagung. Harras, G. (Red.). 1988: 394-408.
- Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Red.). 1984: 13-30.
- Wiegand, H.E. 1989. Der gegenwärtige Status der Lexikographie und ihr Verhältnis zu anderen Disziplinen. Hausmann, F.J. et al. (Reds.). 1989-1991: 246-280.
- Zgusta, L. (Red.). 1992. *History, Languages, and Lexicographers*. Tübingen: Max Niemeyer.

Lexical Culture*

Reinhard Rudolf Karl Hartmann, *University of Exeter, United Kingdom*

Abstract: The paper focuses on the mutual relations between linguistics and lexicography. Against the background of etymology and comparative-historical linguistic studies and using the words 'lexical' and 'culture' as examples, an attempt is made to demonstrate the relevance of morphology, grammar and semantics to the description of vocabulary and its lexicographic codification. Issues of intralingual variety and interlingual diversity can now also be addressed within the theoretical framework of contrastive textology, supported by computer-aided parallel text corpus analysis. However, lexicography is not entirely dependent on the results of linguistic research; it needs to move on and develop its own theoretical base. Recent progress in four fields of dictionary research has indeed contributed to an international upsurge of 'lexical culture'.

Keywords: LINGUISTICS (ETYMOLOGY, MORPHOLOGY, GRAMMAR, SEMANTICS), CONTRASTIVE TEXTOLOGY (PARALLEL TEXT CORPORA), LEXICOGRAPHY (REFERENCE SCIENCE, DICTIONARY RESEARCH), CULTURE

Opsomming: Leksikale kultuur. In hierdie artikel word gekonsentreer op die gemeenskaplikhede tussen die linguistiek en die leksikografie. Teen die agtergrond van die etimologie en vergelykend-historiese linguistiese studies en met 'lexical' en 'culture' as voorbeelde, word 'n poging aangewend om die relevansie van morfologie, grammatika en semantiek aan te dui in die beskrywing van die woordeskate en die leksikografiese kodifisering daarvan. Kwessies van buitetaalige verskeidenheid en intertaalige uiteenlopendheid word ook aangespreek binne die teoretiese raamwerk van kontrastiewe tekstologie, ondersteun deur rekeneaar gesteunde parallelle tekstokorpusanalise. Leksikografie is egter nie heeltemal afhanklik van die resultate van linguistiese navorsing nie; dit moet verder vorder en 'n eie teoretiese basis ontwikkel. Onlangse ontwikkeling in vier velde van woordeboeknavorsing het inderdaad bygedra tot 'n internasionale oplewing van 'leksikale kultuur'.

Sluitelwoorde: LINGUISTIEK (ETIMOLOGIE, MORFOLOGIE, GRAMMATIKA, SEMANTIEK), KONTRASTIEWE TEKSTOLOGIE (PARALLELE TEKSKORPUSSE), LEKSIKOGRAFIE (VERWYSINGSWETENSAP, WOORDEBOEKNAVORSING), KULTUUR

Lexical culture may be a newish phrase, but it is not a new phenomenon. Indeed, human curiosity in words and their ever-changing forms and meanings lies at the beginning of any scholarly effort to understand them. In this sense, the study of words surely predates linguistics as an academic discipline. Often it was triggered by the need to understand words in texts, in religious texts, in legal texts, and in literary texts, for instance. These are also, in many

* Plenary lecture held at the 8th Annual Conference of ALASA, Stellenbosch, R.S.A., July 1995.

cultures around the world, the early forerunners of linguistic interpretation. To give just one example, from India, in the tradition of Sanskrit scholarship, where the exegesis of religious texts, namely the Vedic hymns, eventually led to the impressive linguistic work of Panini and other grammarians at least four centuries before Christ, producing a description of the structure of words which anticipated the rules of generative grammar by over 2,000 years!

In the Western tradition of linguistics, lexical interest was for a long time focussed on etymology, or the explanation of the origin of words and their change throughout the historical development of a language. This in turn led to the creation of historical-comparative linguistics, or 'philology', especially since 1786 when William Jones stipulated a genetic link between various so-called 'Indo-European' languages from Sanskrit and Greek through Latin to most of the modern Romance, Germanic, Slavonic and Celtic languages.

Knowledge of this kind shows us that our phrase 'lexical culture' has a Greek and Roman pedigree, *lexical* being derived from the Greek words *lexis* and *lexicon* which mean 'word' and 'word-book', respectively, and *culture* (literally) rooted in Latin *culture* which in turn comes from the verb *colere* 'to till' or 'cultivate'. Both *lexical* and *culture* are thus very old words with impeccable Indo-European family connections. If we had more time, we could go on tracing their historical evolution and their cognates in other languages; we would discover, and historical-comparative linguists and historical dictionaries such as the OXFORD ENGLISH DICTIONARY would assist us, that there are fascinating linguistic interconnections, on the one hand, between *lexis* in Greek with *legein* 'to read' and *logos* 'speech' and their cognates in other languages, such as Latin *legere* 'to read' and possibly *lex* 'law', and, on the other hand, between *cultus* in Latin with *colonia* 'settlement' and *collum* 'neck' and its cognates in other languages, such as Greek *kuklos* 'cycle', Middle Dutch *hals* 'neck' and English *wheel*.

Today, etymology is no longer academically very respectable, and historical-comparative linguistics is not in great demand in the universities, but most of the world's great all-purpose dictionaries still record lexical information of this kind, and their buyers and users expect to find it in them.

Meanwhile, the term 'derivation' has acquired a modern sense in addition to its historical one: since the middle of our century linguists seem to be less interested in the (diachronic) processes of lexical evolution, but rather more in the (synchronic) processes of word-formation: How are new words produced from the lexical resources of any one language, and how does 'borrowing', or copying, words from other languages work?

This is the province of morphology, the study of the structure of words in terms of the constituent parts that help to shape them, like roots and affixes, free morphemes and bound morphemes, and derivational and inflectional elements.

When we claim that we are now more interested in morphology than in etymology, we are perhaps also making value judgements about such things as

scholarly solidity and historical uncertainty, even (I am exaggerating a little) declaring our preference for fact over fiction. When we say that the morphological resources of English allow the formation of adjectives like *lexical* and *cultural* from nouns like *lexicon* and *culture*, we are observing a regular pattern by inspecting data available to us all, which is an established principle of scientific methodology. However, when we say that English *lexical*, a word first documented in 1836, comes from *lexicon*, borrowed from Greek via Middle French in 1603, and is thus cognate with certain words in other languages, we also grudgingly admit (as does the BARNHART DICTIONARY OF ETYMOLOGY, 1988, from which I have taken this information) that datings depend on the accidental nature of the records on which they are based, and the evidence becomes more and more shaky the further back we go in time, and the more we have to rely on theoretical reconstructions.

Morphology is only one of several branches of linguistics; together with syntax, it forms the field of grammar which has dominated the scene for the last 40 years or so. Indeed, for part of this period it might be argued that there has not been a 'lexical culture', but rather a 'grammar culture' which has tended to relegate vocabulary to lists of items, regarding the dictionary merely as a collection of irregularities that cannot be explained by syntactic rules. One famous British linguist, John Lyons (1971: 514) put it like this:

We can regard the lexicon as a kind of appendix to the grammar — an appendix in which we find, appropriately indexed, all the information that we need to know about particular lexemes or their associated forms and cannot derive from anything else that the grammatical or phonological analysis of the language tells us about them.

By turning their focus of attention away from the lower-level constituents of words towards the higher levels of phrases, clauses and sentences, linguists were in danger for a while of losing sight of the basic lexical units altogether. By regarding words like *culture* and *lexicon* merely as instances of the syntactic category 'noun' and words like *cultural* and *lexical* merely as members of the word-class 'adjective', are we not missing a lot of essential features of these items? (This is one area where the so-called love-hate relationship between linguistics and lexicography could be potentially explosive.)

The aspect of language where 'grammatical culture' has overshadowed 'lexical culture' most is, of course, meaning or semantics. We have already seen how our two favourite words have changed historically and how difficult their interpretation can be. There are certainly multiple senses, as, with *culture*, we move from the cultivation of plants and animals (as in *agriculture* and — of special significance here in Stellenbosch — *viticulture*) to the cultivation of intellectual activities like art, music and education (think of phrases like *high culture* or *cultured person*), and on to further meanings as used in special fields like anthropology, biology and the media; cf. the entry *culture* in the OXFORD

COMPANION TO THE ENGLISH LANGUAGE (McArthur 1992: 274).

To elucidate these various senses or meanings, we need a well-developed semantics. Unfortunately, as another British linguist reminds us (Cruse 1990: 139), semantics is still rather young and relatively unsophisticated:

It is probably true to say that the most influential work to date in semantics has been somewhat theoretical in orientation, and has been directed preponderantly towards elucidating and accounting for the logical properties of sentences within the framework of some system or other of formal logic. There has been relatively little in the way of descriptive work ...

Descriptive work is arguably just as important as theory, you will agree.

It might be argued that the descriptive semantics that linguists like Cruse are demanding has already been achieved by lexicographers who have always striven to supply explanations of the lexical units of the language in the form of definitions in dictionaries. There are even some specialised types of dictionaries available which provide such information, e.g. synonym dictionaries and thesauruses (like Laurence Urdang's excellent and original OXFORD THESAURUS, 1991, which arranges words of similar meaning in alphabetic, not thematic, order and gives example sentences to distinguish shades of meanings between them).

While such reference works undoubtedly exist, at least for the major languages of the world, they cannot meet all the various needs users might have, and they do not usually treat more than the core vocabulary, and that only superficially. So we must continue to encourage a two-way debate between linguistic semantics, on the one hand, and practical lexicography, on the other, about what needs to be done and what can be done to describe in detail the semantic behaviour of lexical items in the context of language use.

A number of theoretical approaches and descriptive techniques are available. From these, I can only briefly mention three.

The first is usually called 'lexical field' theory, or 'semantic field' theory, a way of analysing lexical items into conceptual sets or domains. This is associated with German linguists like Jost Trier and Walter Porzig in the 1930's; you may also have come across names like Baldinger, Geckeler, Lipka and Lutzeier. I myself have experimented with the cooking verbs, pipe fittings, and words like *plane*, *level* and *stratum*, but our words *lexical* and *culture* are still waiting to be described in detail (a good summary of this approach is Adrienne Lehrer's book *Semantic Fields and Lexical Structure*, 1974).

The second approach is labelled 'componential analysis' or 'distinctive feature analysis', the former being associated with American anthropology of the 1940's and 1950's, the latter with transformational-generative grammarians, notably Katz and Fodor who produced the famous example of *bachelor* broken down into its sense elements in the form of a family tree. There is also a Euro-

pean version of this theory which goes under the name of 'structural semantics'; the most well-known instance is the analysis of pieces of furniture for sitting on, presented in the form of a matrix of columns and lines, with + and - signs entered in the spaces, attributed to the French linguist Bernard Pottier.

The third approach I want to single out here is the much more recent offshoot of cognitive linguistics and psychology which is pursued under the banner of 'prototype semantics' and popularised by George Lakoff, Anna Wierzbicka and Jean Aitchison and applied to lexicography by Dirk Geeraerts. Geeraerts (1994) gives the example of *trousers / pants / blue jeans* and discusses some of the factors that determine their respective salience and variation in particular contexts.

Two aspects are often overlooked when these semantic issues are debated, and I have spent many years of my career trying to draw people's attention to them: one is intralingual variation, the other is interlingual contrast. It surely matters, I would have thought, whether a semantic description is done not only for the purpose of establishing the similarities and differences between the analysed lexical items in question (as the three approaches that I have mentioned try to do), but also to check if the style, register or text type in which they typically occur have an impact on the relations between them. Thus, a word like *culture* does not have a single or basic meaning, but various uses in communicative situations. So in the discourse community of social history, *culture* is often synonymous with *civilisation*, while in political texts it can mean 'race' (think of the phrase *multicultural society*), while in medicine the term *culture* can be paraphrased by something like *growth of micro-organisms*.

The other factor in semantic description apart from intralingual variation is interlingual contrast. As we all know, there is no such thing as straightforward one-to-one lexical equivalence, especially between pairs of languages which are not genetically related, and this causes great difficulties in translation, a process with which I am sure most of you are familiar. If you are, you will also be aware of a particularly problematic aspect of this interlingual matching, or lack of fit, namely what is commonly referred to as 'culture-specific vocabulary', or those lexical items that are used to label things, concepts and institutions unique to a particular society. This is an area where linguists sometimes give up, and leave the field completely to ethnographers, translators, and bilingual lexicographers (cf. Hartmann 1992).

While admittedly these problems of intralingual variation and interlingual contrast seem at times unsolvable, we must not lose heart, and indeed help seems to be at hand, both theoretically and practically. In a book which I published 15 years ago, entitled *Contrastive Textology* (Hartmann 1980), I made out a case for an integrated approach to this problem. I saw it as the combination of intralingual variety studies or discourse analysis and interlingual contrastive analysis (hence the term 'contrastive textology' which unfortunately has not caught on, although the desirability of such an approach has since been acknowledged by other linguists, too). Contrastive statements about a pair of

languages only make sense if they are based on the analysis of so-called 'parallel texts', defined as examples of discourse from corresponding genres. This has remained a theoretical demand or dream as long as we have not had the means to carry it out on a large scale.

Today, computer technology gives us the opportunity to attempt a new generation of contrastive text linguistics for the benefit of semantics, translation and bilingual lexicography. I cannot go into the intricacies of 'corpus linguistics', as the literature is too complex to review and the operational procedures are too diffuse to present here. Let me just hint at the essentials. The recent boom in monolingual text corpora, especially of English, has tended to minimise interlingual or crosslinguistic work, although even monolingual corpora can have a contrastive dimension when they give us data about different regional varieties (such as British vs. South African English), about intralingual registers (such as spoken vs. written), or about different periods of the same language.

It is extremely difficult to get a good overall picture of the field of bilingual text corpora, as their scope and design features vary according to whether they are intended for linguistic description, foreign-language teaching, bilingual lexicography, human translation, machine translation, artificial intelligence research, or speech technology, and reports are scattered over very diverse periodicals and conference proceedings. I find it helpful to distinguish two types of parallel texts: 'bitexts' or translationally matched texts (such as European Union reports and laws and their translated versions in other languages), and 'paired texts' which are independently written texts from two languages, but matched in terms of similar communicative settings (such as the manifestoes of the political parties in Britain and Germany). Interesting work has been done, both on aligned bitexts and on matched paired texts, to extract lexical equivalents, which is a promising technique of potential benefit to translators and bilingual lexicographers.

May I end this section by presenting a brief illustration of what I have done on an English-only text corpus with our two words *lexical* and *culture*. One of our students on the Exeter M.A. Course in Lexicography, Lucia Cortese, recently enjoyed a practical placement at COBUILD in Birmingham. I asked her to check and print out all occurrences of the two words, *lexical* and *culture* from the 'Bank of English' corpus of over 200 million words. The material came, by fax, laid out in the form of concordance pages, with the words in question in a central column together with preceding and following context line by line. You will be interested to know that there were 81 citations of *lexical* and exactly 100 citations of *culture*. (See appendix.)

The fact that the two words did not occur together, not even once, simply means that the corpus has no record of the phrase *lexical culture*, and if it does not occur in a large corpus, the likelihood is that it may be a nonce-formation. (If my paper had been processed in Birmingham, Lucia would of course have found quite a number of instances.) The concordance printouts in the

Appendix tell us a number of facts about these two words and their collocations that we could not have predicted simply from native speaker intuition alone, e.g. that *lexical* collocates most frequently with nouns that it modifies, such as *database* and *item*, while *culture* rarely forms noun compounds as first item (in such phrases as *culture gap* or *culture studies*; *culture shock*, surprisingly, was absent), but often forms the last element in collocations like *our culture*, *French culture*, *working-class culture*, *nature and culture*, *civic culture*, *black culture*, *post-modern culture*, and *lesbian culture*.

Before I move on to more strictly lexicographic topics, let me briefly check over the ground we have covered so far. We started by looking for evidence of 'lexical culture' in linguistic studies. We found some in etymology and historical-comparative linguistics, and in the dictionaries inspired by these diachronic perspectives. As we went forward into the structural linguistics of our own century, we saw the emergence of the pioneers of semantics (such as lexical field theory), and also some progress in the field of morphology. But we also noted that during the period (after the 1950's) of the predominance of grammatical theory, linguists pushed lexical considerations to one side. Now, the pendulum has swung again in the other direction, and we can observe further advances in semantics, including contrastive studies, in discourse analysis, in computational linguistics, and in psycholinguistics.

Today, there is a greater willingness on the part of linguists to acknowledge their indebtedness to dictionaries as repositories of lexical knowledge and codification, and on the part of lexicographers to utilise ideas and techniques supplied by linguists. However, while I am stressing this increasing two-way traffic, I do not believe that lexicography is only dependent on progress in linguistics; lexicographers need to develop their own institutions and theories which are appropriate to their specific requirements, and also borrow from other fields such as information technology.

The terminological distinction that is sometimes made between 'lexicology' and 'lexicography' can be misleading. Lexicology, where it is recognised as such, is a branch of linguistics dealing with aspects of the vocabulary (historical, morphological and semantic). Lexicography is not merely applied lexicology, but a related set of lexically based activities which has its own theoretical premises and professional practices. And a lot of these are not 'linguistic' even, so much so that among some of my colleagues we have recently been talking in terms of a wider context of 'reference science', which would include the theory and practice of making reference works of all kinds, including encyclopedias, maps, directories etc., by reference professionals, for the benefit of a variety of users with a variety of reference needs and reference skills (we note, in passing, that 'reference' in this wider sense is only partly connected with the sense in which it is used in semantics).

May I, in the final section, venture a few thoughts on the state of this field of lexicography which I see as being positioned within a wider reference perspective, but still as part of 'lexical culture'. Dictionaries and all other types of

reference works are indeed cultural products serving cultural purposes, which is why the famous French lexicographer Alain Rey — who is also interested in encyclopedias and in intercultural translation — calls the general-purpose explanatory dictionary 'le dictionnaire culturel' (Rey 1987), and why the most recent volume of the *Lexicographica* Series Maior, a festschrift in honour of the Czech-American doyen of lexicography, Ladislav Zgusta, has the title *Cultures, Ideologies and the Dictionary* (Kachru and Kahane 1995).

I have been very fortunate in my academic career to be associated in a modest way with the rise and establishment of lexicography as an interdisciplinary and international field. Exactly 30 years ago, in the title of one of my first review articles, I asked the deliberately provocative question "Is lexicography making progress?" (Hartmann 1965). Today such a title would be quite unthinkable in the light of all the associations, conferences, journals and other publications on the subject (cf. the *International Encyclopedia of Lexicography* edited by Hausmann *et al.* 1989-91), in addition to the flood of dictionaries and other reference books on the market world-wide.

But this growth has been uneven, and some of the progress on the theoretical side of lexicography has been an uphill struggle, I must admit. During my 20 or so years at Exeter, a small provincial university in the county town of Devon in Southwest England, I managed to organise five meetings that have made a contribution to progress by bringing like-minded people together, in 1978 at the Lexicography Seminar of the British Association for Applied Linguistics, in 1980 at the International Summer School on Applied Linguistics and Dictionary Making, in 1983 at the LEXeter Conference on all Aspects of Lexicography (which also became the first EURALEX Congress), in 1986 at the Seminar on the History of Lexicography, and in 1989 at the Workshop on Lexicography in Africa, all with published proceedings. In 1984 we established the Dictionary Research Centre, and since 1987 we have run 10 international one-week courses in Lexicography called InterLex.

Encouraged by the success of all these events, we explored the idea of mounting a European diploma course at Master's level, which we designed with the help of an ERASMUS grant and a consortium led by people at Exeter, Lille and Amsterdam. This M.A. in Lexicography is now in its fourth session. The subject of training is of course fraught with limitations and laced with compromises, but the course attracts half a dozen students each year (from almost as many different countries) who all manage to obtain a practical placement with a dictionary publisher.

My involvement with all these ventures, with the *Lexicographica* series, and visits at centres in other countries has given me an awareness of the possibilities of the field, but also of the enormous differences between the various national, linguistic and cultural traditions. The results of a survey I attempted of the lexicographic scene in Europe was published in *Lexikos* (Hartmann 1993). Before that, I had published the papers of a modest workshop on lexicography in Africa (Hartmann 1990). After my 1995 visit to South Africa I am in a much

better position to compare notes and see where the real differences lie in terms of lexicography between Europe and Africa.

Finally, a few words about my view of four aspects or branches of 'dictionary research', as I prefer to call the enterprise of metalexicography.

The lexicography of a country or language depends not only on the external history of that country or language, but also on the internal cultural traditions that help to shape successive generations of dictionary-making. It is therefore imperative to explore all facets of dictionary history (cf. Hartmann 1986).

Another aspect of dictionary research is dictionary typology, or the classification of the many genres of reference works that we find in various periods, countries, or cultures. Of the many types that we have studied at Exeter, the 'bilingualised' learner's dictionary has offered some intriguing new features (cf. Hartmann 1994).

Innovation may occasionally come from dictionary criticism, a relatively old field which has been bedevilled by prejudice and arbitrary personal opinions. More recently, attempts at rational evaluation of dictionaries against sets of more objective standards have shown promise (cf. Osselton 1995).

All these branches of dictionary research can profit from an empirical look at dictionary users, their reference needs and reference skills which originally cause and ultimately test the compilation of such works. Several postgraduate projects at Exeter have made significant contributions to this field (cf. Diab 1990, Stark 1990).

In conclusion, I would like to say that I am optimistic, both in terms of the 'lexical culture' that is building up in the relations between linguistics and lexicography, and the way lexicography is being established as a respectable discipline. The future looks bright, but having recognised this lexical culture, we must nurture it in every way we can. This includes attention to training, to more ambitious research, to better documentation facilities, and to more intensive contacts between people, disciplines and countries around the world.

References

- Cruse, D.A.** 1990. *Language, Meaning and Sense: Semantics*. Collinge, N.E. (Ed.). *An Encyclopedia of Language*: 139-172. London and New York: Routledge.
- Diab, T.** 1990. *Pedagogical Lexicography. A Case Study of Arab Nurses*. Tübingen: M. Niemeyer.
- Geeraerts, D.** 1994. Varieties of Lexical Variation. Martin, W. et al. (Eds.). *EURALEX 1994 Proceedings*: 78-83. Amsterdam: Vrije Universiteit.
- Hartmann, R.R.K.** 1965. Is Lexicography Making Progress? *The Incorporated Linguist* 4: 42-45.
- Hartmann, R.R.K.** 1980. *Contrastive Textology. Comparative Discourse Analysis in Applied Linguistics*. Heidelberg: J. Groos.
- Hartmann, R.R.K.** (Ed.). 1986. *The History of Lexicography*. Amsterdam: J. Benjamins.
- Hartmann, R.R.K.** (Ed.). 1990. *Lexicography in Africa*. Exeter: University of Exeter Press.

- Hartmann, R.R.K. 1992. 300 Years of English-German Language Contact and Contrast: The Translation of Culture-specific Information in the General Bilingual Dictionary. Blank, C. (Ed.) *Language and Civilization. A Concerted Profusion of Essays and Studies in Honour of Otto Hietsch*: Vol. II, 310-327. Frankfurt: P. Lang.
- Hartmann, R.R.K. 1993. General Lexicography in Europe. *Lexikos* 3: 67-82.
- Hartmann, R.R.K. 1994. Bilingualised Versions of Learner's Dictionaries. *Fremdsprachen Lehren und Lernen* 23: 206-220.
- Hausmann, F.J. et al. (Eds.). 1989-91. *Wörterbücher / Dictionaries / Dictionnaires. An International Encyclopedia of Lexicography*. 3 volumes. Berlin: W. de Gruyter.
- Kachru, B.B. and Kahane, H. (Eds.). 1995. *Cultures, Ideologies, and the Dictionary. Studies in Honor of Ladislav Zgusta*. Tübingen: M. Niemeyer.
- Lehrer, A.J. 1974. *Semantic Fields and Lexical Structure*. Amsterdam: North-Holland.
- Lyons, J. 1977. *Semantics*. 2 volumes. Cambridge: U.P.
- McArthur, T. (Ed.). 1992. *The Oxford Companion to the English Language*. Oxford: U.P.
- Osselton, N.E. 1995. *Chosen Words. Past and Present Problems for Dictionary Makers*. Exeter: University of Exeter Press.
- Rey, A. 1987. Le dictionnaire culturel. *Lexicographica International Annual* 3: 3-50.
- Stark, M. 1990. *Dictionary Workbooks. A Critical Evaluation of Dictionary Workbooks for the Foreign Language Learner*. Exeter: University of Exeter Press.

Appendix: COBUILD concordances of 'lexical' and 'culture'

[The author is grateful to COBUILD, Birmingham, for permission to reproduce these concordances, and to Lucia Cortese for obtaining them.]

<	<tr> Another alternative is to work toward some sort of	lexical	ordering of planes in accordance with general moral	>
<	those of assigning different weights and of establishing	lexical	priorities are undeniably hard to establish and	>
<	the process of picture classification bypasses any	lexical	or logogen system, whereas when names for pictures	>
<	whereas when names for pictures have to be found,	lexical	search and retrieval will become necessary." If this	>
<	by technical terms, and it has had its intrinsic	lexical	and syntactic productivity exploited by Hegel and	>
<	follows a different route. It does not specially involve	lexical	changes, nor syntactic changes (as with Hegel), though	>
<	s work with Generalized Phrase Structure Grammar and	Lexical	-Functional Grammar.Frederick Newmeyer, Linguistic	>
<	Criseyde's heart and the weakness of women: an essay in	lexical	interpretation", Studia Neophilologica, vol. 54 (>
<	they know # I'm the last person to try and trap rap in a	lexical	ghetto of # Yo!" and # Hey, motherf # er!" but,	>
<	Cobuild. Collins and Birmingham University International	Lexical	Database run by MX. A computer fed each week with	>
<	I think <ZF1> the <ZFO> the target skills are control of	lexical	items er and control of structural items # <ZF1> The	>
<	er of the language in context. Target skills control of	lexical	items control of structural items task type text	>
<	perhaps you use in your teaching the concept of linked	lexical	items which in the kind of description I use is called	>
<	s very useful for them to have basic concentration at a	lexical	level when they've actually taken part <ZGY> in a	>
<	neatly into the model that I'm proposing for our new	lexical	database. [laughs # <FOX> Mhm.<FOX> We can't talk	>
<	<ZGY> So stage one we chose erm a few of the commonest	lexical	items in the languages. This was because many of our	>
<	how <ZF1> what <ZFO> what we found <ZGY> very common	lexical	items <ZGY> Then from that experience we deliberately	>
<	s true but why don't I just out all my <ZGY> <MOX> Which	lexical	realization <ZGY> <FOX> Cos you don't know the cost	>
<	ZGY> publisher <ZGY> etcetera and you've got a 1+ a new	lexical	database which contains ei erm information pertaining	>
<	database which contains ei erm information pertaining to	lexical	items. And this is where some of the points that have	>
<	for us phonetics will be one relation within that new	lexical	database which could also contain [pause # information	>
<	or er corpus would have er erm a field anyway in the	lexical	database. How much more information was added to it	>
<	thing the figure in the erm <FOX> Mm.<MOX> er	lexical	database would be updated. Erm [pause] and that would	>
<	we for instance added in some of the fields of the old	lexical	database which you were interested in for instance a	>
<	a <ZFO> a <ZFF> s+ <ZFO> starting point for that new	lexical	database the sooner it can be circulated to other	>
<	again one thing that the [pause] that s+ possible new	lexical	database would have is actually all the attested forms.	>
<	pause] FX you were wanting some information from the old	lexical	database weren't you.<FOX> Have we done five?<FOX>	>
<	it more than once?<MOX> <ZGY> <MOX> Can you use it for a	lexical	database <ZGY> use it as a <ZGY> <MOX> Yes.<MOX> <ZGY>	>
<	times <ZGY> <MOX> <ZGY> <MOX> The admin database the	lexical	database <MOX> <ZGY> <MOX> the grammar database.<MOX>	>
<	except possibly having a syntax field in the new	lexical	database which connects with the tagger there is no	>
<	data on a grammatical basis.<FOX> Yes.<MOX> We only have	lexical	<FOX> Yeah.<MOX> <ZGY> <FOX> But this grammar database	>
<	Mm.<FOX> going to be a massiver thing than the original	lexical	database <MOX> Yeah that's why I said link not <ZF1> I	>
<	connection between lexis [pause] and what they're doing,	Lexical	items must be mentioned at some point in what they're	>
<	might be a good idea to do it both for sort of erm fully	lexical	passives <ZGY> and for those and then together.<MOX>	>
<	<ZGY> in the gr in grammatical ones # <ZF1> In <ZFO> in	lexical	ones you can start off with the basic frequency of the	>
<	<ZF1> I've <ZFO> I've begun looking at these er semi-	lexical	er cohesive items like stuff and thing. Perhaps we	>
<	could think of them as being on the borderline between	lexical	cohesion and substitution <MOX> Yeah.<FOX> in which	>
<	ought to do if it was <MOX> You'd think.<MOX> a typical	lexical	item in a nom+<FOX> Mm.<MOX> at the end of a nominal	>
<	class they're it's <ZGY> that gets increasingly more	lexical.	<MOX> Yeah. Yeah.<FOX> Yeah.<FOX> Yeah.<MOX> So I mean	>

<	that in 1914 died violently, passing into some other	culture	whose form is still unclear to us. Think of its	>
<	that has been learned from the animal and tissue	culture	studies, in order to determine the maximum tolerated	>
<	his visit is far less influential for them than its	Culture	One planners think it is.<t> American Catholics in	>
<	orthodoxy but as true sacramentally-imbued pastors that	Culture	one leaders make themselves welcome in Culture Two and	>
<	It was getting hotter, and desert was spreading, and the	culture	they had lived by for three thousand years was failing	>
<	guests' problems on your own. Other expressions of the	culture,	like Feldman's complaint policy and his declarations	>
<	be <FCH> given <FCH> meaning. Much of what we call	culture	and civilization consists in efforts people have made,	>
<	her own embryos, one after the other, to places of	culture	and edification. Jesse, Woodson, James, and Frank were	>
<	and freedom.Gayle and David are an example of how this	culture	-wide pattern works. For the past five years, Gayle has	>
<	to the surface. There's a rhythm of change in a	culture	where you push, then you pull back a little bit and	>
<	by its adherents in the historical setting of Western	culture.	His concern with social and political questions led to	>
<	stirring the pot will cook up an ever more vibrant civic	culture,	improving the quality of life for us all. Those of us	>
<	Fri. 7:30 p.m.Having a Wonderful Time (1938) A	culture	-seeking typist vacations at a mountain resort and	>
<	Speaking to foreign journalists, the deputy minister for	culture	of the Afghan government said that the foreign	>
<	Jovic called for an .. for .. bilateral ties in the ..	culture,	educational, scientific, technical and border	>
<	and belongs to a new class of virus.Now they want to	culture	it in the laboratory.<h> THE FARMING WORLD 1622british	>
<	and they are as much a part of Greek history and	culture	as the Hellenistic majority. There are small and	>
<	into Europe they brought with them a more sophisticated	culture	than the Neanderthals had # And this may certainly	>
<	lover valuable increases her value. In a materialist	culture,	this kind of thinking is common. How does one know	>
<	touch with the pop music, the mini-skirts, the teenage	culture	of the 1970s. Her peers were for the most part	>
<	the absence of correspondence between polity and	culture,	inherent in 'Canadian audiences' consumption of United	>
<	Random House).Almond, G. and Verba, S. (1963), The Civic	Culture	(Princeton, NJ: Princeton University Press).Althusser,	>
<	which Weber carved out a position for himself within	culture,	with which he effected his own insertion into	>
<	culture Weis (1985, p. 7) adopts a hegemonic view of	culture	which sees culture as semi-autonomous (Gramsci, 1971).	>
<	Clearly, the majority of people within almost every	culture	no longer believe that the only true experience of	>
<	into what we know today as Sanskrit. The Indo-Aryan	culture	was decidedly patrist, evidence of ancestor-worship of	>
<	100 people whose language one does not speak and whose	culture	one does not understand should not be underestimated.	>
<	an economic "base".As Worsley remarks, 'the concept of	culture	has been virtually ignored by those social scientists	>
<	abilities: some effects of age, schooling, and	International Journal of Psychology, vol. 14 (1979),	writ large would become the essence of the modern	>
<	educationalsystem; in the urban context, this peasant	culture	and society are tending towards individualism and	>
<	consequently in tension with Bellah's view that American	culture	exists. Already in the early 1970s, ideas were	>
<	acknowledges that an alternative to the prevailing	culture	and Commerce <LTH> Paris is indisputably one of the	>
<	<LTH> PHOTO WITH CAPTION # <LHH> The Capital of French	Culture	in Oberammergau and Salzburg,but came away imbued with	>
<	as Germanophiles, following happy experience of German	culture	You'd be amazed at how many people call us in	>
<	direct about what they want or need. It's part of the	culture.	but Peter Muller, the Amandari's brilliant architect,	>
<	world hotels make an irritating nod towards the local	culture,	or that of china's occupation will be dependent in	>
<	this as a poignant symbol for the fate of Tibet's living	culture,	that has diversity of age, colour, class, cultural	>
<	were writers of fiction or poetry, representative of a	culture		>

<	of any formal education and his public championing of a	culture	of loud and loutish C2-ishness. This finds its most	>
<	have always used it as the language of literacy and	culture,	and in Pakistan, where the mother tongues of the vast	>
<	invite the Japanese in. It'd wipe out any remaining	culture,	but they'd make the air breathable." <LTH> Are you	>
<	weren't enough) of how much Elvis looted from black	culture.	The photo selection is good (if not better) than most	>
<	Hackney Homeless tent, things get slicker. YOGHURT	CULTURE	are about smooth grooves, the mood of Motown soul and	>
<	to make things better. <LTH> <LHH> LEO <LTH> The call of	culture	is very strong for Leos this week. In fact you could	>
<	Programmes were removed from the Ministry of Sport,	Culture	Youth Affairs to the Ministry of Education. and	>
<	not sufficiently attuned to the new Council's # business	culture	to work for Westminster. No thought is given to their	>
<	but our main concern is to set about creating a lesbian	culture.	From the start we have organised discussions,	>
<	muscles (or lack of them). And for women who mix in a	culture	where it's 'in' to reveal breasts, the string vest is	>
<	and audience a lesson in the evils of a Euro-centric	culture,	Everyone cheers, and you can't help but wonder if she,	>
<	short, there is emerging a global mass button-thumping	culture,	with inane, if not regressive values, and we continue	>
<	Roots, an anthology of gay history, sex, politics and	culture	presents a wide range of essays and a troubling amount	>
<	where contact with the modern African-Caribbean	culture	a mixture of fascination and revulsion. <LTH> Those	>
<	10.30-m # <KPD> 4.80 <KPD> 3.50 concs). <LTH> QUEER	CULTURE	# The Future of Lesbians and Gay Men in Theatre and	>
<	the power to create laws that promote its language and	culture	# If approved by lawmakers in Canada's 10 provinces,	>
<	Muslim country or a country of completely different	culture	# Oh, I have to tell you that the level of knowledge	>
<	has taped a message for the American people # Iraq's	culture	and information minister says Iraq has officially	>
<	<< Audio footage of an Indian chant)<< Berkes: Our	culture	is important, one boy tells us later, we're proud of	>
<	series on the family # In the last few decades, American	culture	has changed quite a bit # The American family has	>
<	9205050211 UKRAINIANS STRUGGLE TO REGAIN IDENTITY UN,	CULTURE	Noah Adams, host:<t> The vote for independence in the	>
<	class, and Gwen Macsai's hilarious send up of our body	culture	told me why # Thanks, Gwen # By the way, I am an avid	>
<	what else are we missing out on in terms of Japanese	culture?	<t> Reid: Well, there are really some good movies-	>
<	You know, there's no doubt about it # I call this a	culture	war # We're kind of, you know--there's an old	>
<	Schneiderman (Medical Ethicist): Maybe we are in a	culture	of cure rather than in a culture of care # This is a	>
<	but- <t> IRENA BUDO, Hungarian Teacher in Ukraine: In	culture	and in democracy- <t> Ms. SZOLVACI: But spiritual	>
<	is where it's at the city's where we live that's and our	culture	and our lives basically are linked to cities now in	>
<	were so integrally connected with western history and	culture	that in fact it was impossible er for those resources	>
<	have written about the concept of organizational	culture.	I'm thinking of books like In Search of Excellence	>
<	major focus of my P H D is to look at this concept of	culture	and see how it can be applied in a more rigorous	>
<	you think that's due to nature or due to culture.<FOX>	Culture.	<M01> Right. Right. Thank you. pause <FOX> I think we	>
<	of this year, examining almost every aspect of Jewish	culture,	but ignoring the Holocaust. This omission, highly	>
<	life to the city is now regarded as history, as museum	culture,	then all Glasgow is diminished. <t> The Second City	>
<	before 1995, said meaningfully: 'It's a question of	'culture	in Europe, not European culture # Aaah .. <t> It	>
<	that created the world's greatest free enterprise	culture	should appear here, beneath the myth he was himself so	>
<	state and military zone, and their ancient Buddhist	culture	destroyed or driven into exile. Stalinism and nazism	>
<	done. Bloom in particular saw that transmission of the	culture	is the rock upon which liberal democracy might	>
<	20 outlets are planned by 1995. <t> But bridging the	culture	gap is not always without its problems. <t> In Hong	>

<	firms or launching their own businesses where male	culture	does not dominate. <t> In my experience, however,	>
<	And Major, the man who was railing against Britain's yob	culture	a month ago, was there to get what turned out to be a	>
<	for stifled entrepreneurial vigor. The political	culture	that created the conditions in which crack selling is	>
<	developments in technology and to our accelerated	culture.	But within what remains of that shabby, embarrassing	>
<	Among the infinite echoes of our sorry post-modern	culture,	anything and nothing goes. Everything around us is	>
<	the sixties, he averred, his view of the emerging new	culture,	would have been brighter.<t> Peter Hall, on his Royal	>
<	hands in apocalyptic horror at the dangers of American	culture,	while at the same time enjoying a passionate love	>
<	owed much to the same kind of argument. Working-class	culture	was inherently racist, officers would say. Once the	>
<	Charter of 1955, which did indeed call for 'the doors of	culture	to be opened to all # thus placing culture, however	>
<	That book was subtitled Tales for an Accelerated	Culture.	In the novelty of that description, and a narrative	>
<	live here among ourselves to defend our history and our	culture."	A project that they obviously feel Le Pen can help	>
<	the garden as a site for the interaction of nature and	culture,	and incorporates fossils, skulls, encrusted tools and	>
<	rather the part that can be absorbed, let us call our	'culture	acquisition device". That means I can line it up with	>
<	own grammar, and computers, too. We are going from a	culture	dominated by literal logic to a culture in which there	>
<	was to exterminate his people, their religion and their	culture;	and he called for an international commission to	>
<	in the power to make proposals on health, education and	culture	(though only their cross-border aspects) and on	>
<	You can take individuals and scatter them into another	culture.	If you take a Japanese or a Chinese and you dump him	>
<	cheeks. For Filipinos, superstitious and sentimental by	culture,	the era of 'Steady Eddie' has started well.<p> That	>
<	many Jews, the middle-class ones often embraced French	culture.	As early as the 1920s several Jewish women there	>
<	to be exposed to television programs which reflect our	culture,	our multiculturalism, our ethics and our values	>
<	keys in locks to language, music, geography, maths and	culture.	She said the school, established in 1978, aimed to	>
<	for a variety of disciplines. In Aboriginal art and	culture	lessons for nine to 12-year-olds, they'll learn about	>
<	while showing off to the world, but he can't destroy our	culture,"	he said. He can try and make us go to discos and	>
<	her whole life as an Aborigine out of touch with her own	culture	in a society which largely ridiculed and rejected her	>

Lexicographical Resources in A Multilingual Environment: An Orientation

Willy Martin, *Free University of Amsterdam, The Netherlands*

Abstract: This article considers dictionaries as lexical information / knowledge sources to be derived from a deeper, underlying, lexical database. These dictionary-tokens or -instantiations are *inter alia* specified by the users' needs. As a case in point of such a derivation meeting the needs of a multilingual society, a bidirectional bilingual learner dictionary is presented. Specific tools, such as editors with reversal function, and models, such as the hub-and-spoke model, are discussed as means to function within the lexicographical infrastructure of a multilingual society.

Keywords: LEXICAL DATABASES, BIDIRECTIONAL DICTIONARIES, LEARNER DICTIONARIES, DICTIONARY EDITORS, REVERSIBILITY, HUB-AND-SPOKE MODEL

Opsomming: Leksikografiese hulpmiddels in 'n meertalige omgewing: 'n oriëntering. In hierdie artikel word woordeboeke beskou as leksikale inligtings- / kennisbronne wat gebaseer moet word op 'n grondiger, onderliggende, leksikale databasis. Die karakter en inhoud van hierdie woordeboeke word o.a. bepaal deur die gebruikers se behoeftes. 'n Tweerigting tweetalige aanleerderswoordeboek wat voorsien in die behoeftes van 'n meertalige omgewing, dien as voorbeeld van so 'n generering. Spesifieke hulpmiddels, soos teksredigeerders met omkeerbare funksies, en modelle, soos die naaf-en-speekmodel, word bespreek as metodes waarbinne die leksikografiese infrastruktuur van 'n meertalige omgewing kan funksioneer.

Slutelwoorde: LEKSIKALE DATABASISSE, TWEERIGTINGWOORDEBOEKE, AANLEERDERSWOORDEBOEKE, WOORDEBOEKREDIGEERDERS, OMKEERBAARHEID, NAAF-EN-SPEEKMODEL

Survey

Introduction

1. Dictionaries
2. Different Kinds of Dictionaries
3. The Changing Lexicographical Landscape: Dictionaries versus Lexical Databases
4. 'Dictionaries' in a Multilingual Environment
5. Aspects (of the Construction) of a Multilingual Lexicographical Infrastructure

5.1 General Objectives

5.2 Tools

5.3 Models

5.4 Actors

Bibliography

Introduction

Being asked by the organizers of this seminar* to deal with the question 'which dictionaries are needed to fill the needs of a multilingual country such as South Africa', I have preferred to answer this question in a fairly *general* way. First of all because of the fact that I can *not give specific* answers to such a question myself. It is for the speakers of this country who come after me to deal with needs and preferences of the respective linguistic communities within South Africa.

Secondly, I have considered it much more informative for you (and feasible for me) to provide you in a first instance with a general and in a sense '*empty*' framework for the construction and organization of a multilingual infrastructure. I do believe that such a framework is needed as a background against which to situate, compare and specify the diverse, concrete needs, preferences and products. In a sense what is presented here is in line with the frame approach I have been advocating in other publications of mine (see e.g. Martin 1994), providing people with slots which they have to specify and fill out themselves.

1. Dictionaries

In accordance with the general approach just advocated, I will start from a very broad and even traditional definition of what is called a 'dictionary'. Defining it as:

- * a book
- * containing information on (aspects and / or parts of) the vocabulary
- * of one or more languages
- * by making use of two dimensions:
 - a macrostructure
 - and a microstructure

From the above it will have become clear that what is meant here is a printed dictionary and that the distinction between macro- and microstructure (the

* The seminar was held on 12 April 1996 in Stellenbosch under the title *Lexicography as a Financial Asset in a Multilingual South Africa*.

vertical and the horizontal dimension) — although in itself an artificial convention — will serve both to bring together and to differentiate between such objects as e.g. explanatory monolingual dictionaries, pronouncing dictionaries, thesauri, frequency lists, monodirectional bilingual dictionaries, terminological dictionaries and encyclopedias, to name but a few.

2. Different Kinds of Dictionaries

Although the above-mentioned objects all can be said to belong to the 'type' dictionary, this does not mean that we can not make any distinctions between them as 'tokens'. Following Geeraerts 1984 in this respect, we will argue that parameters which characterize dictionaries are linked to

- * the selection and / or the presentation
of
- * their macro- and / or their microstructure

In other words, all differences between 'dictionary-tokens' will be differences between either the selection and / or presentation of either their macro- and / or their microstructure. Distinct surface markers or indicators to differentiate between instantiations of the 'type' dictionary are then (also see Al-Kasimi 1977, quoting Malkiel 1959):

- * scope (both macro- and microstructurally speaking: what kind of language population is the dictionary aiming at?, which data categories does it want to account for?)
- * linguistic perspective (e.g. does the dictionary describe language synchronically or diachronically?)
- * entry arrangement (e.g. alphabetical, reverse alphabetical, chronological, onomasiological)
- * number of languages (one or more)
- * user group(s) (this 'marker' is, of course, of an other order, namely based on function, rather than on features, this way it precedes or supersedes the already mentioned ones, user group(s) implying a certain scope, perspective etc.; it is kept here both because of its high degree of relevance and its relative distinctness).

By means of parameters such as the above one can now characterize different kinds of dictionaries. It would be of great help both for users and for makers of dictionaries to have / make this information available. Not so much in order to come to a dictionary categorization / classification per se — the above parameters moreover much more lead to a parametrization than to a fixed or rigid classification — but because such a characterization / specification is of great help both in the planning, the design, the use and the evaluation of dictionaries.

In other words, a characterization of a dictionary such as the English-Dutch Van Dale as

- * a dictionary
- * meant for native speakers of Dutch
- * for them to understand and translate
- * contemporary English
- * of both spoken and written 'texts'
- * of a general nature

(see Martin-Tops 1989², Introduction, p. 13) is not just a superfluous metalexicographical minor detail, but a useful and necessary datum in the description of the lexicographical landscape of a language.

3. The Changing Lexicographical Landscape: Dictionaries versus Lexical Databases

One of the striking differences between lexicography nowadays and lexicography a couple of decades ago, is, among others, that, with the advent of computers in lexicography, both an evolution in breadth and one in depth has taken place. As to the evolution in breadth: a shift can be observed from just one object or product of interest, the printed or paper dictionary, to a whole range of them. In Martin 1995a in addition to the traditional paper dictionary the following objects of interest are listed:

- * computer-aided / based dictionaries (i.e. dictionaries developed with the help of computers)
- * machine-readable dictionaries (i.e. dictionaries of which a machine-readable version exists)
- * electronic dictionaries / term banks (dictionaries of which the data are electronically stored and which dispose of retrieval software)
- * machine dictionaries (dictionaries for NLP-systems, showing, as a rule, a higher degree of formalization of the data than is the case with paper dictionaries)
- * lexical databases (dictionaries in database form, with a high degree of formalization, and retrieval software)
- * AI-lexicons (comparable to lexical databases, involving as an extra next to lexical knowledge also knowledge about specific State-of-Affairs)

Actually this list is open-ended and other 'types' such as hypertext dictionaries, multimedia dictionaries etc. could be added; however what is important, is that, next to this broadening of the field, also a 'deepening', an evolution in depth, has taken place. As a rule, one can say that nowadays underlying a printed dictionary there is a lexical database from which it is derived, the latter

also being 'deeper', in that it is both more explicit and yet more general, i.e. not directed to just one user group. Eight years ago in 1988 at the 3rd Euralex Conference in Budapest Al and Martin accordingly could claim therefore that the best way to come to user friendly dictionaries was for them to be derived from non-user oriented lexical databases. To quote the above-mentioned article:

'We would like to suppose that every new dictionary that is published nowadays has been derived from an underlying database (which can, of course, be more or less sophisticated). Furthermore we suppose that different types of users are in need of different types of dictionaries. (...) This does not imply, however, that one has to build up a completely separate database for each type of dictionary. *For several reasons it is preferable to set up one "subjacent", fundamental database which is not user-oriented, and to derive from it as many user-oriented front-end databases as there are types of dictionaries.* Editing, updating and further completion of the bare lexical data are activities which — in a certain sense — should be unrelated to the final products a publisher may have in mind. They concern the fundamental, non-user-oriented database. Front-end databases on the contrary, are typically product-oriented. They contain specific selections which depend entirely on the needs of the users for whom the dictionaries involved are intended" (Martin-Al 1990:393).

Figure 1 (taken from Heid 1991) clearly illustrates this change in focus. In the centre of the figure the lexical database is to be found whose primary objective is to represent lexical knowledge. This knowledge is acquired and imported through interfaces from sources such as text-corpora, other dictionaries, databases, human informants) and the like (see left-hand side of figure). At the right-hand side concrete products (dictionaries) are to be found which are exported via interfaces, derived from the central lexical database and oriented towards specific users, thus leading to specific applications.

What is needed first of all therefore as a lexicographical resource in my opinion, is not so much one or more concrete products to relieve the specific needs of certain language users with, but much more the underlying database with which to relieve not only these, but also other possible and important needs.

4. 'Dictionaries' in a Multilingual Environment

However true the above observation may be, it does not exempt one from having to conceive the possible derivations and products, the front-end databases that, in our case, might be of particular interest in a multilingual society.

It is rather obvious that next to lexicographical products which are to be found in a monolingual society (such as monolingual dictionaries e.g.), a mul-

tilingual society needs other products too, such as bilingual or translation dictionaries. Less obvious it may seem that a multilingual environment actually may require a special kind of translation dictionary, e.g. a multifunctional one serving the needs of both the source and the target language users. Indeed, given the fact that a country such as the new South Africa has eleven official languages and so fifty five different language pairs, producing mono-functional / monodirectional dictionaries (and so at least four dictionaries per language pair) (see Kromann et al. 1991) would involve two hundred and twenty different dictionaries. Clearly this would become too costly an affair. Moreover if a multilingual society is really to function properly, it should dispose of learner dictionaries, preferably of bilingual bidirectional learner dictionaries so that speakers of different language groups may learn each other's language. This does of course not alter the point-of-view that there should be just one database (per language) to derive all these different products from. On the contrary the more variety the different front-ends show, the more the need and usefulness to start from just one common single non-oriented lexical database will be felt.

By way of example and because of the fact that the notion 'bidirectional learner dictionary' is fairly new, I will dwell somewhat longer on this concept (see Martin 1987 and Hannay-Wekker 1995). Learner dictionaries are not to be seen here as primary teaching tools but as specific kinds of dictionaries with a supporting role in the language learning process. As such they may contribute to improving the learner's receptive and productive knowledge of the target language. However where Martin 1987 still starts from a monodirectional learner dictionary, Hannay-Wekker 1995 present a bidirectional one, implying that the dictionary should function for two language groups at the same time, serving as both an L_2-L_1 and as an L_1-L_2 dictionary.

An example entry of such a (experimental) bidirectional English-Dutch dictionary follows (see Hannay-Wekker 1995:28). It shows the following structure:

1. **Headline** (containing the headword and basic information about it)
2. **Profile** (containing a set of meaning descriptions and basic coded information on these translations, so to be able to make a choice between given possibilities)
3. **Frame** (listing the most frequent discursual, syntactic, collocational and prepositional environments of the head word organized according to distinctions in the profile field)
4. **Translations** (including frequent contextual occurrences of the headword plus their translations)
5. **Notes** (providing information on synonyms and usage)
6. **Expressions** (comprising idioms, pragmatic formulas and proverbs)

The slots: 'headline' and 'profile' are obligatory, the other ones are optional.

Example of bidirectional English-Dutch entry

HEAD	ANXIOUS <adj.-bn.> [fon] <anxiously; anxiety, anxiousness>
PROF	1 [worried] bezorgd + over => [stronger; mostly predicative] ongerust + over
	2 [causing worry] angstig
	3 [=> be anxious] erg graag willen + dat
FRAME	1 s.o. is anxious about sth s.o. has an anxious look / expression an anxious mother
	2 an anxious time an anxious question an anxious situation
	3 s.o. is anxious to do sth s.o. is anxious for s.o. else to be / do sth s.o. is anxious that s.o. else should be / do sth s.o. is anxious for sth
NOTES	1 syn.: <i>worried, concerned; anxious</i> benadrukt de vrees van de betrokkene
TRANS	2 anxious anticipation \ bange verwachting
	3 she is anxious for her mother to be present \ zij wil erg graag dat haar moeder er bij is; he was anxious to get away \ hij stond te popelen om weg te gaan

As stated before a bidirectional dictionary is multifunctional. In the above case of a bidirectional learner dictionary English-Dutch it is meant to be

- (a) a productive dictionary of English for the speaker of Dutch
- (b) a receptive dictionary of Dutch for the speaker of English
- (c) a productive dictionary of Dutch for the speaker of English

Clearly then as some of the information will be primarily meant for one type of language user and even be superfluous for the other (see e.g. Frame, meant for TL-speaker only), it is necessary for such a dictionary to have a clear structure allowing for different paths to be followed. Figure 2 gives an idea of how this can be achieved.

5. Aspects (of the Construction) of a Multilingual Lexicographical Infrastructure

After having stressed the importance of theoretical insights both for the representation (see section 3) and for the application of lexical knowledge (see section 4), in this section I would like to take into account some aspects of the implementation side too, paying some attention to the pragmatics of setting up

a multilingual lexicographical infrastructure and so dealing with such issues as general objectives, tools, models and actors.

5.1 General Objectives

I take for granted that the set-up of a multilingual infrastructure in a multilingual environment as such is not to be questioned: such a structure is simply to be regarded as a basic social benefit for the environment in question. However I hold the view that the multilingual infrastructure one should aim at, should lead to the production of lexicographical products (mono / bi / multilingual)

- # of a high-level quality
- # within a coherent, anticipatory framework (thus superseding short term planning)
- # on an economically justified basis.

This should imply multifunctional, re-usable and linkable resources. Such resources presuppose appropriate tools, one of which, namely OMBI, will be presented in the next section as a case-in-point.

5.2 Tools

OMBI, an acronym based on the Dutch word group 'Omkeerbare Bilinguale Lexicale Databanken' (= Reversible Bilingual Lexical Databases), is an editor which has been developed during the academic year 1994-1995 by the Dutch software house SERC (= Software Engineering Research Centre, Utrecht, The Netherlands) under the auspices of the CLVV — a committee of which I happen to be the chairman — (Commissie voor Lexicografische Vertaalvoorzieningen = Committee for Lexicographical Translation Resources). This Committee is an intergovernmental body of lexical experts set up in 1993 by the Ministry of Education and Science of both Flanders and the Netherlands in order to improve and stimulate the production of bilingual dictionaries and lexical databases with Dutch as a source or target language. The Committee has launched up till now several lexicographical projects which are, commercially speaking, non-viable, yet of great social relevance. In this respect projects such as Turkish-Dutch v.v., Arabic-Dutch v.v., Polish-Dutch v.v. etc. need to be mentioned.

However not only is it the Committee's task to have concrete products realized, but also to see to it that, if needed, adequate lexicographical tools and infrastructure are provided for.

The construction of OMBI is to be situated within this second domain, aiming at providing lexicographical teams with a generic and powerful editing tool.

Actually OMBI's main characteristics come down to the following:

- it functions as an *editor*
- which is *generic*
- having *importing and exporting* facilities
- and the power to *reverse* lexical databases, trying to do so in *as accurate as possible* a way

As the last aspect is the most innovative I will briefly deal only with this characteristic.

The fact that OMBI can reverse translational relations and directional databases in general makes it rather unique. While the editing function is busy creating a bilingual database $X \ll Y$, and as such taking in translations from X to Y , OMBI simultaneously stores the reversed counterparts, thereby building a reverse database $Y \ll X$. The end result is a non-directional bilingual database, from which databases and / or dictionaries in both directions can be automatically derived at a subsequent stage.

In order for the process and outcome of reversal to be non-trivial, the tool should not merely state that if word form x is a translation of word form y , then word form y is a translation of word form x . This is in many cases not only too limited a conclusion, but also a wrong one: only rarely is translation a straightforward symmetrical relation between word forms.

The first, highly important observation about translation relations is that it is not *words* that are translated into other words, but rather *words in a specific meaning*. The English word *horse* is a translation of the Dutch word *paard*, but only in the meaning of the latter as [certain animal], not in its meaning [certain chess piece]. This insight has had a fundamental influence on the architecture of the databases that OMBI builds. The database distinguishes between Form Units or FUs (word forms) and Lexical Units or LUs (meanings): every Form Unit (e.g. *horse*) can have one or more meanings (e.g. 1—[certain animal], 2—[certain chess piece], etc.); only a LU (which always belongs to an accompanying FU) can be translated by a LU into another language (see fig. 3).

The second important observation is that translation, and reversal of translation in particular, only holds if certain conditions are met. In OMBI the translation relation is analysed into four relevant parameters that influence reversibility, and which therefore have to be specified and taken into account in 'calculating' whether the reversal of the relation is valid or not. The four parameters are the following:

- conceptual equivalence
- pragmatic contrast
- variant status
- lexicalization status

It is not the place here to discuss the OMBI-calculus (see e.g. Martin-Tamm 1996 in this respect), however seeing some of its results may give one an idea of how it works and of the role it could play in the establishment of both a coherent and sound and yet economical lexicographical infrastructure. Starting from the following INPUT (Dutch-English e.g.):

honger <zn.> 1 <n.-telb.zn.> <neutral, contemporary, neutral, neutral, general, general> [desire to eat] *hunger; appetite (1.a) (idiom) ik heb honger als een paard I'm starving; I'm so hungry I could eat a horse; I'm hungry as a hunter; (1.b) ik heb toch een honger! Am I hungry!;* 2 <n.-telb.zn.> <neutral, contemporary, neutral, neutral, general, general> [lust] *lust; appetite; hunger (2.a) (idiom) honger naar geld en goed lust for material things; greed;* 3 <n.-telb.zn.> <neutral, contemporary, neutral, neutral, general, general> [famine] *famine; hunger.*

paard <zn.> 1 <telb.zn.> <neutral, contemporary, neutral, neutral, general, general> [animal] *horse (1.a) (idiom) het paard achter de wagen spannen put the cart before the horse;* 2 <telb.zn.> <neutral, contemporary, neutral, neutral, general, sports> [gymnastics] *horse; vaulting horse;* 3 <telb.zn.> <neutral, contemporary, neutral, neutral, general, sports> [chess] *knight.*

raspaard <zn.> 1 <telb.zn.> <neutral, contemporary, neutral, neutral, general, general> *purebred horse; pedigree horse; full-blooded horse; thoroughbred.*

automatically yields the following OUTPUT (English-Dutch) for horse:

full-blooded <adj.> 1 <adj attr> <neutral, contemporary, neutral, neutral, general, general> [(1.a) (lexical collocation) *full-blooded horse raspaard.*

horse <n.> 1 <countable> <neutral, contemporary, neutral, neutral, general, general> [animal] *paard (1.a) (idiom) put the cart before the horse het paard achter de wagen spannen; (1.b) (lexical collocation) purebred horse raspaard; (1.c) (lexical collocation) pedigree horse <rel.> raspaard; (1.d) (lexical collocation) full-blooded horse raspaard;* 2 <countable> <neutral, contemporary, neutral, neutral, general, sports> [gymnastics] *paard;* 3 <countable> [idiom.] (3.a) (idiom) *I'm so hungry I could eat a horse ik heb honger als een paard.*

5.3 Models

An action plan for a multilingual infrastructure not only involves a delineation of objectives and of tools to realize them, also models, in the sense of overall frameworks and / or approaches, are needed.

In this respect I would like to briefly present a model which may be of particular interest, given a multilingual environment, namely the hub-and-spoke model (see Martin 1995 and Mashamaite 1995). The hub-and-spoke model actually is inspired by air traffic organization in which certain airports (hubs) act as centres to and from which flights from other airports (spokes) operate. The key contention of this model for lexicography is that several bilingual dictionaries can be made by linking the lexical items of the spoke languages (source languages selected for the envisaged bilingual dictionary) to those of a common hub language (serving as the target language in the language pair). Suppose e.g. that one has a region in which English, Lesotho sa Leboa and Tshivenda are accepted as official languages, then what actually should be provided for are three lexicons which are linked to each other forming a hub-and-spoke configuration (see fig. 4).

The linking itself consists in supplying for and making explicit the kind of relationship there exists between the lexemes of the spoke languages and their equivalent hub lexemes. This equivalent relationship is rendered along the lines outlined in the preceding section, making use of the four parameters mentioned there, namely: conceptual equivalence, pragmatic equivalence, lexicalization status and variant status. So e.g. 'setimela' in Lesotho sa Leboa could be linked to 'train' in English through

- * conceptual equivalence = complete
- * pragmatic differences = nil
- * lexicalization status = lexicalized
- * variant status = main

so that it could be reversed without any restrictions: train = setimela, and that it could be linked to Tshivenda 'tshidimela' which is linked in the same way to English 'train' as is 'setimela'.

The consequence of applying this model in a multilingual environment such as South Africa with eleven official languages, would drastically reduce the work needed for the production of bilingual dictionaries. Instead of some hundred (actually 110) bilingual dictionaries to be made without any co-ordination at all, just ten lexical databases would need to be constructed and linked up all in the same way with a central (hub) lexical database so that

- (a) Spoke and hub can be reversed
- (b) spoke_i and spoke_j can be linked (and reversed)

Of course both calculi will involve blockings such as they will need modifications and gap-filling as well. Yet it is my contention that the model leads to high quality, consistency and economy.

5.4 Actors

It will have become clear from the preceding that the setting up of a lexicographical infrastructure is a complex process involving many actors such as lexicographers, metalexicographers, publishers, subsidizers, (official) language planners, not to forget the users. Each of them in turn show different subtypes and categories. The exact role of each of them is difficult, not to say impossible to define without knowledge of the exact contextual factors. May it therefore suffice here for me to state that next to the main issues already dealt with one action not yet mentioned here is of primary importance for the success of the whole undertaking, namely that of the co-ordination of all activities and of all actors involved. The fact that this very Seminar is set up within this spirit of co-ordination and co-operation, bringing together representatives and experts from several fields and domains, is however proof of it that exactly this aspect has not been lost sight of.

Bibliography

- Al-Kasimi, Ali M. 1977. *Linguistics and Bilingual Dictionaries*. Leiden: Brill.
- Geeraerts, D. 1984. Dictionary Classification and the Foundations of Lexicography. *ITL: Review of Applied Linguistics* 63: 37-63.
- Hannay, M. and H. Wekker (Eds.). 1995. *Towards a Database for General Translation Dictionaries and Bilingual Learner Dictionaries, with Special Reference to Dutch and Portugese*. The Hague: CLVV-report.
- Heid, U. 1991. *A Short Report on the Eurotra-7 Study*. Stuttgart: Unpublished manuscript.
- Kromann, H.P. et al. 1991. *Principles of Bilingual Lexicography*. Hausmann, F.J. et al. (Eds.). *Wörterbücher, Dictionaries, Dictionnaires* III. Berlin, New York: De Gruyter. 2711-2728.
- Martin, W. 1987. Niet ieder (leer)woordenboek is een leerwoordenboek. *Lexicologie. Toegepaste Taalkunde in Artikelen* 27: 60-68.
- Martin, W. 1994. 'Schemata für Wissensrepräsentationen und Wörterbuchdefinitionen'. Hyvärinen, I. und R. Klemmt (Eds.). *Von frames und slots bis Krambambuli: Beiträge zur zweisprachigen Lexicographie*. Jyväskylä. 29-46.
- Martin, W. 1995a. Maschinelle Lexikographie: ein Blick in die Zukunft. Hitzemberger, L. (Ed.). *Angewandte Computerlinguistik*. Hildesheim: Olms. 1-21.
- Martin, W. 1995b. Government Policy and Bilingual Lexical Databases: the Action Plan for Dutch. *Second Language Engineering Convention*. London. 133-144.
- Martin, W. and G. Tops (Eds.). 1989². *Van Dale Groot Woordenboek Engels-Nederlands*. Utrecht / Antwerpen: Van Dale Lexicografie.
- Martin, W. and B. Al. 1990. User-orientation in Dictionaries: 9 Propositions. Magay T. et al. (Eds.). *Budalex '88 Proceedings*. Budapest. 393-399.
- Martin, W. and A. Tamm. (forthcoming). OMBI: An Editor for Constructing Reversible Lexical Databases (to be published in Euralex-96-Proceedings).
- Mashamalte, K.J. 1995. *The Hub-and-spoke Model: A Recipe for Making Bilingual Dictionaries between African Languages in South-Africa*. Doctoraalscriptie Lexicologie. Vrije Universiteit, Amsterdam.

Figure 1: Processes in Lexicography

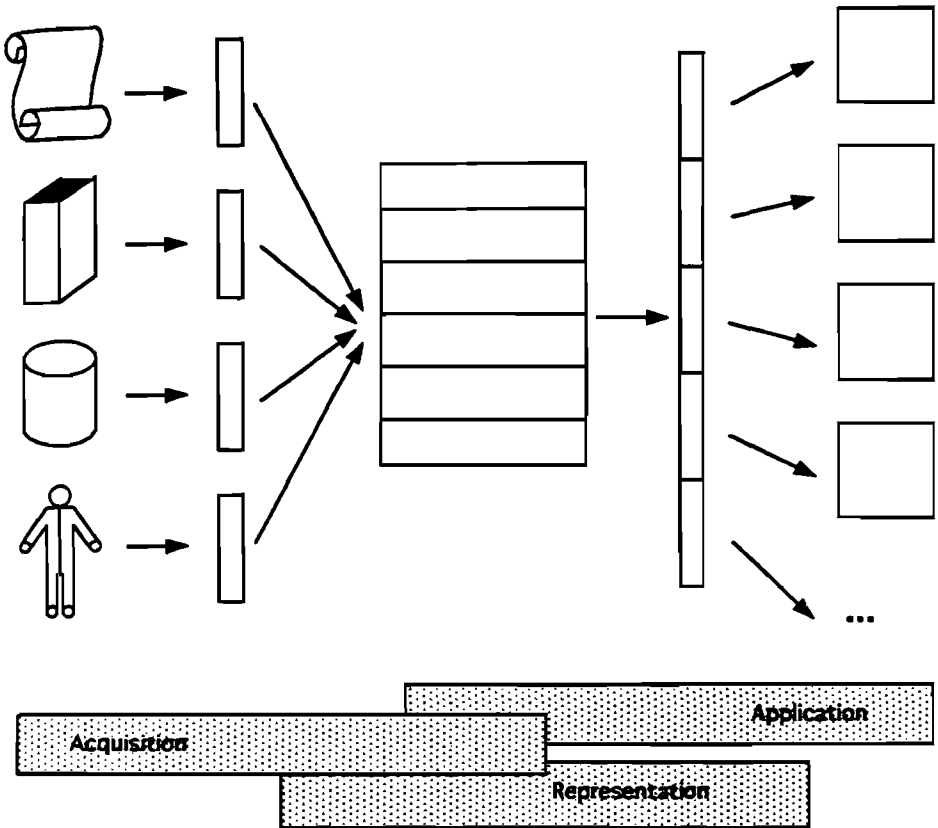
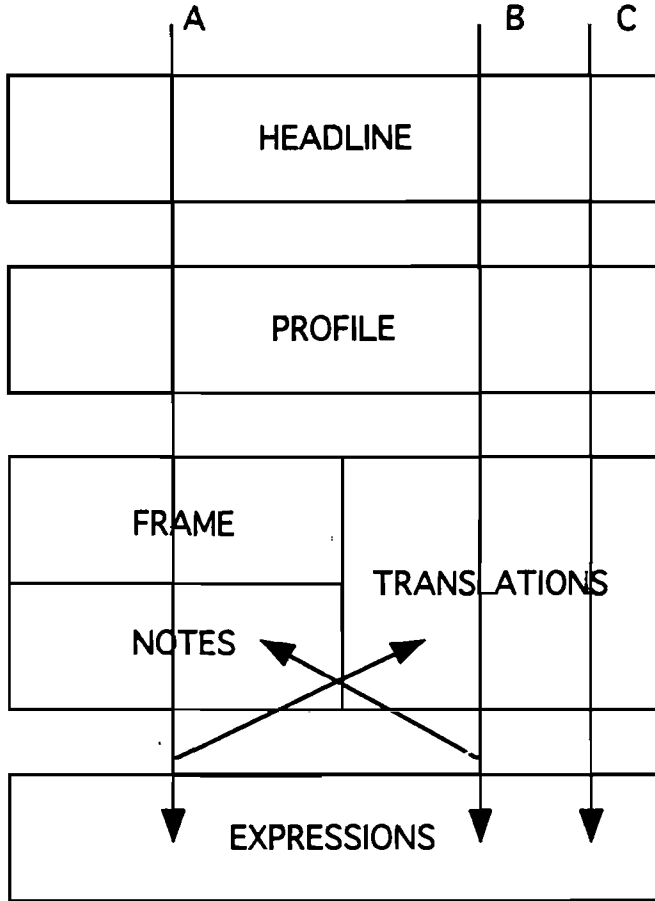
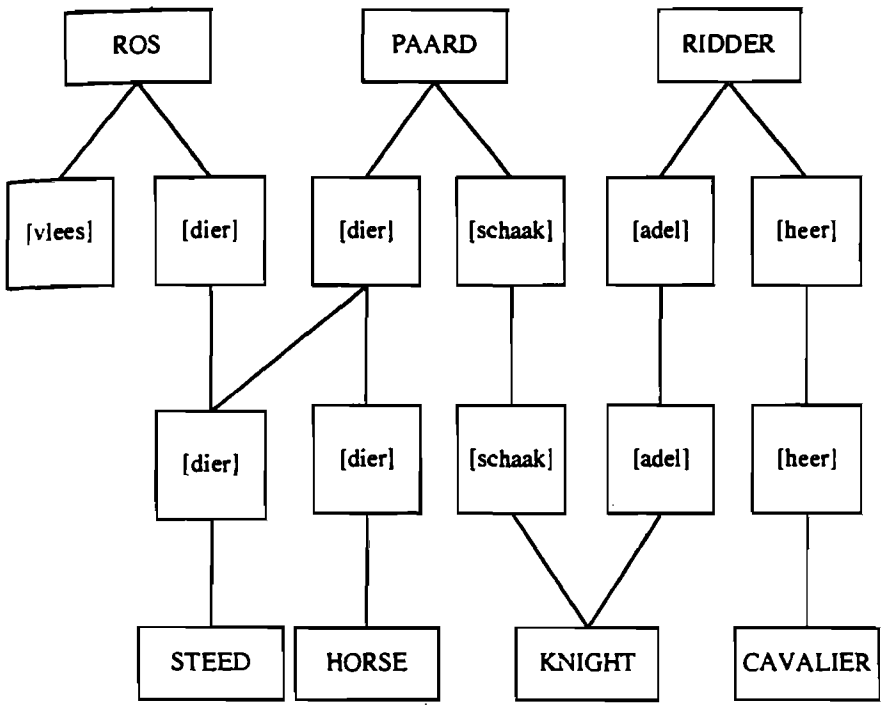


Figure 2: Pathways through an entry in a bidirectional English-Dutch dictionary

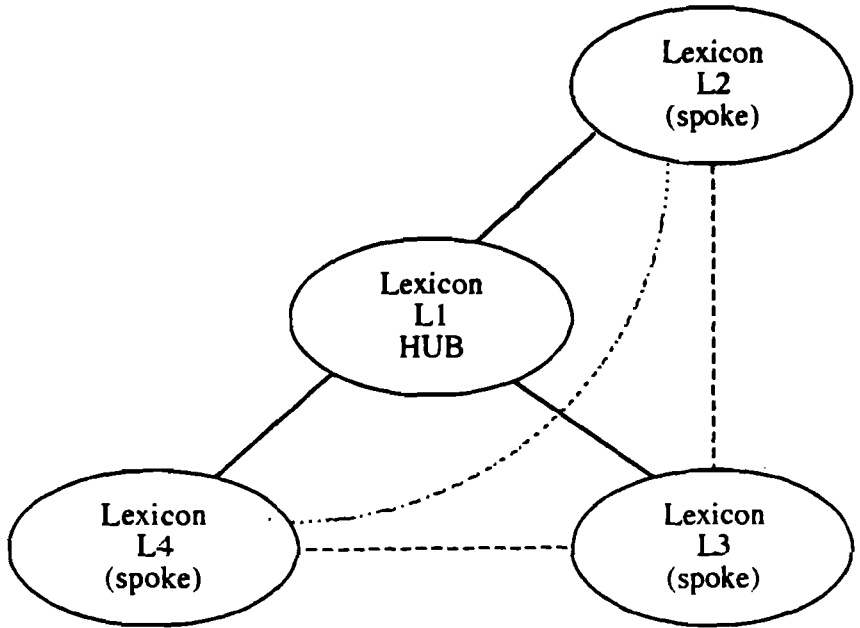


- A = the productive user of English,
- B = the receptive user of English,
- C = the productive user of Dutch

Figure 3: Interlingual Links between Lexemes



Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2011)

Figure 4: Hub-and-Spoke-model

— = Direct Link (from hub to spoke and vv)
- - - = Indirect Link (from spoke via hub)

Authenticating the Vocabulary: A Study in Seventeenth-Century Lexicographical Practice

N.E. Osselton, *University of Newcastle, England*

Abstract: Thomas Blount is unique among English compilers of the seventeenth century in systematically naming books and authors as evidence for the use of words entered in his dictionary. Such documentation of the vocabulary would now be associated with scholarly historical dictionaries such as the Johnson and the OED, rather than with a small dictionary for general use. In his *Glossographia* (1656) Blount differs from these by referring mainly to contemporary writings (and without giving quotations) and in being *selective* in his attestations, for about one in twenty of all entries. An analysis is presented of the types of words thus treated, and of the varied purposes served by his references to external sources: he sees the need to authenticate the less stable parts of the English vocabulary of his day — neologisms, exotic terms, semi-technical and learned words. Blount probably derived this lexicographical technique from the Latin dictionaries of his time. A century earlier than Dr Johnson, he confronted some of the same problems in justifying entries made in his dictionary, and produced different but equally valid answers to them.

Keywords: AUTHENTICATION, DICTIONARY, ENGLISH, HARD WORDS, LEXICOGRAPHY, MONOLINGUAL, RENAISSANCE, VOCABULARY

Opsomming: **Stawing van die woordeskat: 'n Studie van sewentiende-eeuse leksikografiese praktyk.** Thomas Blount is uniek onder Engelse samestellers van die sewentiende eeu deurdat hy boeke en outeurs sistematies noem as getuienis vir die gebruik van woorde wat in sy woordeboek opgeneem is. Sodanige dokumentasie van die woordeskat sou tans geassosieer word met wetenskaplike historiese woordeboeke soos Johnson en die OED, eerder as met 'n klein woordeboek vir algemene gebruik. In sy *Glossographia* (1656) verskil Blount hiervan deur veral te verwys na tydgenootlike werke (en sonder om aanhalings te gee) en deur *selektief* te wees in sy stawings, vir omtrent een uit twintig van alle inskrywings. Daar word 'n analise aangebied van die tipes woorde wat so hanteer is, en van die verskillende doelstellings wat deur sy verwysings na eksterne bronne bereik word: hy sien die noodsaaklikheid daarvan in om die minder stabiele deel van die Engelse woordeskat van sy tyd te staaf — neologiesmes, eksotiese terme, semi-tegniese en geleerde woorde. Blount het hierdie leksikografiese tegniek waarskynlik afgelei van die Latynse woordeboeke van sy tyd. 'n Eeu voor dr. Johnson het hy sommige van dieselfde probleme ervaar om inskrywings wat in sy woordeboek gemaak is, te regverdig, en hy het ander maar net sulke geldige antwoorde daarop gelever.

Slutelwoorde: EENTALIG, ENGELS, LEKSIKOGRAFIE, MOEILIKE WOORDE, RENAISSANCE, STAWING, WOORDEBOEK, WOORDESKAT

Introduction

The authentication of words entered in a dictionary may be *indirect*, through description in the preface or on the title-page of the compiling procedures which have been followed; or *direct* (especially in the larger or more scholarly dictionaries) by reference to authors or the inclusion of quotations in the lemma itself. In the more modern mode, authentication may be by reference to a computer corpus from which the whole has been derived. This urge to validate dictionary entries is however a longstanding one, and the occasion to do so has varied from age to age in accordance with the perception of the lexicographer's task, as well as the current view of the state of the language.

Within the English tradition it is Dr Johnson who is generally credited with setting the pattern for lexicographical authentication: he provided quotations for most — though not all — of the words entered in his Dictionary of 1755. In the more modest works of his predecessors no quotations had been given and there was in general little in the way of external documentation for the vocabulary. Among them the recusant scholar and antiquary Thomas Blount, in his *Glossographia* (1656), is unique in his extensive use of *direct authentication* in entries such as the following:

Blateron or Blateroon (*blatero*) a babler, an idle-headed fellow. Mr How.
Spatiate (*spatior*) to walk abroad, to travel, to wander, to go jetting up and down. Bac.

The word *blateroon* appears to have been a nonce-formation on the part of 'Mr How.', i.e. James Howell (1594-1666) and the only citation for the word in the OED is from his *Epistolae Ho-Eliaanae* of 1645. *Spatiate*, for 'stroll, wander' occurs first in Lord Bacon's *Sylva* (1626), being used there in a figurative sense ('The fixing of the Minde upon one Object of cogitation, whereby it doth not spatiate and transcurre'); the word survives into the nineteenth century, though it has never been as well-established in English as the corresponding verb *spazieren* (*gehen*) in German. *Blateroon* and *spatiate* were thus both of them rare, and both of them were recent acquisitions: two good reasons for the lexicographer to add the authors' names to show the user that, unusual or outrageous as these particular items might seem, they nevertheless represented genuine English usage.

The citing of authorities for words he entered is one of several new features in Blount's dictionary — he is for instance also the first English compiler to include etymologies. Starnes and Noyes (1946, chapter V), while recognizing such innovations, present the *Glossographia* as being in the main an archetypally derivative 'hard word' collection, but the evidence produced later by Schäfer on how Blount took over material from other people's glossaries (1978) has shown the need for further study of the variety of his compiling methods.

The following description of the range and style of Blount's external refer-

ences is necessarily selective and impressionistic, but it may serve to illustrate how innovative Blount was, and also to establish his place more clearly within the English and continental European lexicographical tradition.

Authentication in lexicography

Nowadays we tend to take our dictionaries on trust, and assume that if a word is entered, then it really exists, and is usable. The compilers may explain their principles of selection in the front matter (which most users will never read), and they may (or may not) include guidance on usage and word status in the individual lemma. But it is the simple fact of inclusion which carries most weight with the user.

What then is the case for buttressing the status of a particular word by saying that it has been used by a particular author? — just in the way, that is, in which we should feel that it is proper to give credit to the author of a quotation which we use. This may be done for archaic words (used by Shakespeare, Milton, etc.) as in Chambers (1995). It may also be the practice in dictionaries of slang (Ayto and Simpson 1992) and in even quite modest dictionaries of new words (Tulloch 1992). But then all such words are felt to be in some way peripheral, and this no doubt explains the editors' determination to shelter behind the hard fact of attestation. It is fair to say that today little need is felt for the authentication of words in general (as opposed to historical) dictionaries.

Historical dictionaries form a category on their own, with needs of their own. Here, the names of the authors, ranged in chronological order, have always been acknowledged as the necessary objective evidence for statements about the lexical history of the language which the compilers wish to make, and the quotations normally subjoined are first-hand historical documents on the basis of which the reader may form an independent judgement. This is as true of Johnson's Dictionary (1755) as it is of the OED. Both of these works are in their several ways concerned with *current* use as well, and a fine array of reputable and recent quotations may serve also to confer an air of authenticity upon a word for which they are given. On the other hand it is noticeable that in the COD, which may still in some sense be regarded as a boiled-down version of the OED, such quotations are absent: entries in smaller general dictionaries have always been considered to be self-authenticating. Blount differs from other lexicographers of his age (and of ours) in not making this assumption. He therefore documents those words where status might be in doubt, or where he felt the reader should be directed to further sources of information.

The Scale of Authentication in Blount

The *Glossographia*, with approximately 12,000 headwords, is a modest octavo volume with well spaced-out columns, and the typical lemma will contain only about twenty to twenty-five words. Some headwords are indeed given much more extensive treatment (especially legal terms such as *waif*, which takes up a page and a half), but in this compact and practical volume the lexicographer could not permit himself the leisurely, discursive style of grander folio reference books such as Minsheu's *Ductor in linguas* (1617), and certainly not the parade of authors and quotations which was to characterize historically-orientated dictionaries from Johnson onwards. It is then the more surprising that documentation of sources (and therefore authentication of usage) is squeezed in at all. An analysis of three blocks of consecutive entries (from near the beginning, in the middle, and near the end of the alphabet) showed that for roughly one word in five some kind of documentation is given to external sources. The statistics are as follows:

<i>Sequence of entries</i>	<i>number of entries</i>	<i>entries with sources</i>	<i>%</i>
consummate to decarck	300	66	22
mortmain to nymphous	300	50	17
vespilone to wurt	300	74	25
Totals	900	190	21

Seventeenth-century dictionaries are notorious for their inconsistency, and the promise at the beginning of the alphabet has often faded before we get to the end. But Thomas Blount shows a steadiness of purpose in documenting his less well-known words, from *abbord* 'to go near the shore ... Florio', to *zoographer* 'a Painter or one that draws the pictures of beasts. Dr Br[owne]', some 2500 words in all.

The Range of External Sources

Given the great range of writings drawn upon, it is not hard to take Blount literally when he tells us that the dictionary had 'taken me up the vacancy of above Twenty years' (Sig. A3 *recto*). An impression of these scholarly siftings may be set out under several heads.

Literary Well-known literary figures whose names are given as an authority for including particular words include Lord Bacon (1561-1626), Sir Thomas Browne (1605-82), Geoffrey Chaucer (c. 1343-1400), John Cleveland (1613-58), Abraham Cowley (1618-67), Robert Herrick (1591-1674), Ben Jonson (1572-1637), Sir Thomas More (1477-1535), Sir Walter Raleigh (1554-1618) and Sir Henry Wotton (1568-1639). There are also references to the *Eikon Basilike* (1649) — then commonly supposed to be the work of King Charles I — and to various lesser-known writings such as Owen Felltham's *Resolves Divine, Moral, Political* (1620), Joshua Sylvester's translation of *Du Bartas his Devine Weekes and Workes* (1605), and the political allegory *Dodona's Grove* (1640) by the royalist pamphleteer James Howell (1594-1666).

For a work which is limited in size and selective in its use of citations, this makes quite an impressive roll-call, though the absence of authors such as Spenser, Sidney or Shakespeare may seem odd to us today. But with the exception of Chaucer, who is a special case in the history of English lexicography (Kerling 1979) and Sir Thomas More, the emphasis is upon recent or contemporary writers — often quite new publications — and in this respect Blount is closer to our ideas of what a modern dictionary should be than Dr Johnson was when he resolved 'to admit no testimony of living authours'.

The items which Blount backs up by reference to such literary writings may be characterized as *the less common words in general use* — non-technical, but of low frequency; such as *amnesty*, *effronterie* (both from *Eikon Basilike*), *foliage* (from Herrick's poems), and the Gallicisms *cumble* 'the top-fulness, height or overplus of any thing' and *clientele* from Howell. The authors most commonly referred to are Lord Bacon and Sir Thomas Browne. Typical of Bacon are *lubrefaction* 'a making slippery' (the last recorded instance of the word), *scissile* 'easie to cut' and *sibilate* 'to whistle or hiss'; and of Browne *locotion* (the first recorded instance of the word), *populosity* (a current variant of *populousness*) and *protuberant*. These are fair examples of the 'hard words' whose right to a foothold in the English language was so hotly debated in the sixteenth century. In his prefatory matter (Sig. A4 verso) Blount explains that the language had of late become greatly enriched by the introduction of such words; previously they would have sounded harsh, but use of them by 'our best modern Authors' had now made them familiar 'even to vulgar capacities'. Thus, he has added their names so that nobody should think he was either inventing the words, or recommending their use 'let every ones *Genius* and the quality of the Subject they treat be their own *Dictator*'. It is a fair statement of principle on the authentication of words in any dictionary.

The work by Sir Thomas Browne which he drew upon most frequently was the *Pseudodoxia Epidemica* of 1646 (Osselton 1995, chapter 6). The astonishing range of recondite topics covered in that work made it a treasure-house of terms such as *corrodible*, *cosmical*, or *crustaceous* which were undoubtedly technical or semi-technical, but yet bore the sanction of use by a major literary figure. What more could the lexicographer desire?

Travel Books Exotic words have always posed problems for lexicographers: it is not always easy to decide when words relating to foreign customs or institutions, flora, fauna, etc., may qualify for inclusion. Blount makes good and intelligent use of some of the better-known travel books of his age to buttress his decisions. The handsome folio volumes of George Sandys' *Relation of a Journey begun Anno Dom. 1610 ... Containing a description of the Turkish Empire, of Ægypt, of the Holy Land, of the remote Parts of Italy, and Ilands adjoyning* (1615) gave him the backing he needed for words such as *coral*, *permagies* ('a sort of little Boats among the Turks') and *tarantula*. From Sir Thomas Herbert's *Relation of some yeares travaile begunne anno 1626 into Afrique and the greater Asia* (1634) he takes *buzzar* ('a Market-place among the Persians', i.e. *bazaar*), *catamite*, *mosche or mosque*, *pistachoes* and many others. For *caravan*, *mufti* and *timariots* (Turkish landholders with an obligation of military service) he refers to his namesake Sir Henry Blount's *Voyage into the Levant* (1636).

Perhaps the most striking instance of how effective the use of such travel books could be is the entry for *cha*, the early spelling for *tea*, which was then a newly imported commodity. The earliest recorded instance of the word in the OED (with the spelling *chaw*) is dated 1616. Blount's entry reads

Cha, is the leaf of a tree in *China*, about the bigness of a myrtle, which being dried in iron Sives over the fire, and then cast into warm water, serves for their ordinary drink. *Hist of China*. f. 19.

The Spanish writer Alvaro Semmedo's *Imperio de la China* was published in Madrid in 1642, but the English translation of it (*The History of that great and renowned Monarchy of China*) did not come out until 1655, the year before the word appeared in Blount, whose definition (including the detail about the myrtle leaf) is taken from that work. Here then was a brand-new term not merely exotic, but lexicographically relevant because tea was at the time becoming established as a beverage in England (Pepys first drank tea in 1660). By turning to the very latest travel book Blount could not only authenticate his entry, but was also able to provide the dictionary users with the kind of cultural background information on this novelty which they would certainly want to know.

Religious writings Given the scale and intensity of religious controversy and persecution at the time, the mid-seventeenth century dictionary compiler had a clear duty to his users to provide some reasonable cover of the specialized vocabulary to be found in the enormous output of religious books and pamphlets of the day. Here is a selection of titles of religious writings referred to in individual lemmata in Blount's dictionary:

Broughton, Richard. *The Ecclesiasticall Historie of Great Britaine, deduced by Ages*. 1633. [arch-flamins, cynosure]

- Cressy, Hugh (Serenus Cressy OSB). *Exomologesis: or, a Faithfull Narration of the occasion and Motives of a Conversion unto Catholike Unity*. 2nd ed. 1653. [abstention, mystical]
- Digby, Sir Kenelm. *Two Treatises. In the one of which the Nature of bodies; in the other, the Nature of mans soule, is looked into*. 1645. [bigot]
- Featley, Daniel. *The Dippers Dipt*. 1645. [anabaptists, coaxation, hemerobaptist]
- Gregory, John. *Notes and Observations upon some Passages of Scripture*. 1646. [hellenistical, neophite]
- Hammond, Henry. *A Paraphrase and Annotations upon the Books of the New Testament*. 1653. [phasm, phylactery]
- James I, (King). *Daemonologie, in forme of a Dialogue*. 1597. [necromancer]
- Montagu, Walter. *Miscellanea Spiritualia or Devout Essaies*. 1648. [charlatan, piquant, resipiscence]
- Rushworth, William. *The Dialogue of W. Rushworth, or Common Sense in the Choise of Religion*. 1640. [autographer]
- Spottiswoode, John (Archbishop). *The History of the Church in Scotland*. 1639. [lollards]
- Taylor, Jeremy. *A Discourse of the Liberty of Prophecyng*. 1647. [patrocinate, relegation]

The combination here of ecclesiastical history, works of devotion, biblical commentary, sectarianism and controversy would have been creditable even as a special reading programme for the current religious vocabulary. Blount often took more than just the words themselves from his sources. At *mystical* ('secret, hidden, sacred') there is a sub-entry as follows:

Mystical Theology, is nothing else in general but certain Rules, by the practise whereof, a vertuous Christian may attain to a nearer, a more familiar, and beyond all expression more comfortable conversation with God by arriving unto, not onely a beleef, but also an experimental knowledge, and perception of his divine presence, after an expressible manner in the soul, &c. *Cressy fol. 437*.

In form, this is a definition, but the text is lifted quite literally from Hugh Cressy's famous account of his conversion (except that Cressy says 'after an inexpressible manner ...'). We shall never know for sure what Blount's methods of compiling were, but many of the names listed above (Cressy, Digby, Rushworth, Spottiswoode) crop up in his private correspondence (Bongaerts 1978). It seems reasonable to assume that in making many of the more interesting entries in his dictionary he was simply drawing upon books which came his way or were in his library.

The instances in the above examples of more general, non-sectarian words such as *coaxation* ('the croaking of Frogs or Toads') and *piquant* ('Fr. pricking, sharp, piercing, stinging, nipping') shows that even if Blount was conducting a trawl of religious writings for theological and ecclesiastical terms, he must also have been alert (as any lexicographer should be) to the possibilities of picking up other items too.

Where it is appropriate, he backs up entries by reference to the Bible, and

this can be useful, as in *paraclete* 'an Advocate or Patron; a comforter. Our Saviour Jesus Christ is so called, 1. Joh. 2.1.', as also for the Pauline reference (Phil. 3.2) for the use of *conciSSION* in the sense of 'schism'. It must be added that Blount is generally even-handed in his definitions, avoiding the *odium theologicum* which commonly attends the treatment of religious words in other dictionaries of the time.

The Law Later in life, Thomas Blount was to produce his own *Law Dictionary* (1670), and it is therefore not surprising to find legal terminology well documented in this earlier and more general work. The entries for many legal terms incorporate references to laws and statutes, often (though not always) taken over from Minsheu (1617):

Hambles, is the plural of the French (*hable*) signifying as much as a Port or Haven of the Sea. This word is used, *Anno 27. H 6. cap. 3*
Libel (*libellus*) literally signifies a little book, but by use it is the original Declaration of any Action in the Civil Law. *An. 2 H. 5. ca. 3.* and *An. 2. Ed. 6. ca. 13.* It signifies also ...

The reference can (as with *hambles*) serve to authenticate a less well-known word, or else (as with *libel*) direct the reader to a source of further information. There are numerous references to Cowell (1607, 1637), the standard legal dictionary of the day, but Blount also draws on more specialized writings for his law terms. The legal sense of *ambidexter* for a double-dealing juror 'that takes of both parties, for the giving his Verdict' is backed up by a reference to Richard Crompton's edition of Fitzherbert's *Office and Auctoritie de Justices de Peace* (1583). Information from John Manwood's *Brefe Collection of the Lawes of the Forest* (1592) is quoted in the explanation of *agist* and *paunage* ('money taken by the Agistors for feeding Hogs with the Mast of the Kings Forest'). For *headborow*, now known to us mainly from Shakespeare's *Taming of the Shrew*, he refers to a treatise on *The Duties of Constables* (1583) by William Lambard. The same author provides Blount with a rather dubious justification for entering the word *eirenarchy* — 'Mr. Lamberd wrote a book called *Eirenarchie*, or the Office of a Justice of Peace'; it is true that Lambard wrote such a book, but since the word in his title is *eirenarchia*, Blount's case for putting it in as an English word is not a strong one. No lexicographer is infallible — certainly not Blount — but at least we can say that the users of his dictionary would come across a wealth of legal information there, often with pointers to where they could find more.

Technical terms Legal terminology is only one of the special vocabularies to which the compiler of a general dictionary is inclined to give house room (however reluctantly) and coverage of 'Terms of Art' (as technical terms were then known) was both a headache and a topic of title-page boasting for many

of the compilers of seventeenth-century dictionaries. The technique in the *Glossographia* is to refer to 'authorities' selectively, where highly specialized terms seemed to call for documentation. Here are some representative ones:

- ARCHITECTURE Wotton, Sir Henry. *The Elements of Architecture*. 1624. [cockle-stayres 'a pair of winding stairs'), *mosaique* 'a most curious work wrought of stones of divers colours']
- ASTRONOMY Fulke, William. *A Goodly Gallerie with a most pleasant Prospect ... to behold the natural Causes of all Kynde of Meteors*. 1563 (repr. 1655). [comet, exhalation]
- CHRONOLOGY Hopton, Arthur. *A Concordancy of Yeares ... Containing the Use of the English and Roman Kalendar*. 1612. [nones]
- FARRIERY Markham, Gervase. *Markham's Maister-peece; or What doth a Horse-man lacke*. 1610. [pissuprest 'or the Stone, is a disease in a Horse ... the suppression of the Urine']
- GEOGRAPHY Wright, Edward. *The Description and Use of the Sphaere*. 1613. [latitude, longitude]
- HERALDRY Legh, Gerard. *The Accedens of Armory*. 1562. [scarpe, 'a Scarf, worn by Commanders in the Field; and so named in heraldry']
- MATHEMATICS Digges, Leonard. *A geometrical Practise, named Pantometria*. [pantometrie]
- MEDICINE Read, Alexander. *The Manuall of the Anatomy or Dissection of the Body of Man*. 1638. [ligament]
- MILITARY AFFAIRS Barriffe, William. *Militarie Discipline; or, the Young Artillery-man*. 1635. [plottons 'a term of war'; i.e. platoons]

Sometimes the form of his entry suggests that he may (very rightly) himself have entertained doubts about the legitimacy of a term:

Pantometrie (Gr.) a measuring of all kinde of quantities; It is the title and subject of a Mathematical Book, set forth by one Mr. Digs.

But Blount was living in a time when the vocabulary of scientific writings in the vernacular was expanding enormously, and he deserves credit for the lexicographical field-work which enabled him to document so many of the newer technical terms of his time.

History The writings of historians provided Blount with attestations for a wide range of specialist words. Such are *creticism*, *hagiographer*, *quinquireme* and *petalism* ('a kinde of banishment for five years among the Siracusians') from Sir Walter Raleigh's *History of the World* (1611). Of these *creticism* 'the Art of coyning or inventing lyes' was evidently coined by Raleigh himself. It occurs in book 5 of the *History* (Blount gives the page reference) and appears never to have been used by anyone else. Another and more recent world history — Peter Heylyn's *Cosmographie ... containing the Horographie and Historie of the whole World* (1652) — furnishes him with documentary back-up for items such as (under the letter P) *pamphili* ('a kinde of great Boats in Italy'), *parallels*, *peninsula*, *pergamenous* ('of or belonging to ... Parchment') and *Protestants*. Classical antiquity is represented by Thomas Godwin's *English Exposition of the Roman Antiquities* (1614:

amphoral, ballista, chorus, etc.); the Dark Ages by Verstegan's *Restitution of decayed Intelligence* (1605: *wassail, werewolf, Woden*). There are references to John Stow's *Annales of England* (1592) for the explanation of *curfew* and *Danegelt*, and the neat definition of the adjective *wardwit* ('quit of giving money for keeping Watch and Ward') has an added page reference to his *Survey of London* (1598). Blount opts for a recent attestation (Sir Richard Baker's *Chronicles of the Kings of England*, 1643) for the word *heptarchy*, a word much in use in descriptions by renaissance historians of the supposed seven distinct kingdoms of Anglo-Saxon England. Even more up-to-date is the reference under the word *autocrasie* ('self-subsistence, a being able to subsist of ones self) to Hamon L'Estrange's *Reign of King Charles*, published only two years before the dictionary came out.

Words with particular historical associations (*danegeld, pikemen*) still pose problems for lexicographers today, and it made good sense for Blount to attach external references to some of them, while making use of his reading among the historians for the documentation of more general words as well.

Reference works When Blount protests (over modestly) in his address to the reader that he has 'done little with his own pencil', he signals a debt to earlier lexicographers. He has, he says, 'extracted the quintessence of Scapula, Minsheu, Cotgrave, Rider, Florio, Thomasius, Dasypodius, and Hexams *Dutch*, Dr Davies *Welsh Dictionary*, Cowel's *Interpreter*, &c. and other able Authors, for so much as tended to my purpose' (Sig. A5 *recto*). This honest acknowledgement is highly unusual for its time, and is also of interest because it provides first-hand evidence of the patchwork of different language combinations which the mid-seventeenth century lexicographer felt he could call upon in making up his word-list: Greek and Latin (Scapula), Latin and German (Dasypodius), Latin and English (Thomas, Rider), French and English (Cotgrave), Italian and English (Florio), Dutch and English (Hexham), Welsh and English (Davies) and polyglot (Minsheu). Dictionary compilers have always plundered each other's works; the surprising thing here is that Blount should have been so open about his indebtedness.

It is possible to identify 'other able Authors', notably his predecessor John Bullokar, whose *English Expositor* (1616) is cited especially frequently in the early letters of the alphabet. Legal dictionaries referred to include those of John Rastell and of Sir John Skene, as well as that of Cowell. There are references also to '*Seaman's Dict.*', i.e. *The Sea-man's Dictionary* (1644) by Sir Henry Manwayring, and to Thomas Wilson's *Christian Dictionarie* published in 1612.

The advantage of taking over words from other people's dictionaries is that suitable ready-made definitions may be to hand. Blount adopts many of the less common words from Bullokar (such as *amphysician, Paracelsian* and *sassafras*) complete with definition, and at times (e.g. *corrosive*) adds Bullokar's name to support an additional meaning for a word. Other dictionary definitions are adapted or expanded: the explanation given for the verb *rumidge* ('to remove any Goods or Luggage out of a place. Seamen use it for ...') is adapted

from Manwayring's fuller nautical account of the matter, and in Blount's definition of *heretick* as 'a person wilfully and stiffly maintaining false opinions against the Scriptures and Doctrine of the Church after due admonition', it is only the five words *and Doctrine of the Church* which have been added to what he found in Wilson's dictionary — a mild expansion, and no doubt in accord with views which would be held by a recusant compiler.

Blount also drew upon specialist word-lists. Schäfer (1978) has pointed out that 27 of his entries are lifted from the *Short Table of some hard words and phrases* appended to Edmund Bolton's *Elements of Armories* (1610). This glossary contains 54 terms, many of them having to do with heraldry (*diagonal, imbricate, oval, etc.*). The implication is that when Blount adds the note '*El. of Ar.*' to his entry for the word *analogy* this might give the false impression of a lexicographer finding his examples in context, when in fact he was cribbing them from the back of the book. However, other words from Bolton which occur in Blount (e.g. *parallelistion*) do not appear in the glossary, and we may therefore take him quite literally when he says of it 'I finde the worde used in the *Elements of Armories*'.

To draw on other men's word-lists, as Blount does, is in any case not to his discredit. Such secondary documentation (the use of reference works, as opposed to original quotations) was later to be widely used by Johnson (De Vries 1994), and it also accounts for a fair proportion of the citations given in the OED (Osselton 1995 chapter 14).

The Style and Function of References given within the Lemma

Simply to append the name of Lord Bacon (or Howell, or Sandys, or anyone else) to an entry in your dictionary can be a compact way of authenticating a particular word: it at least establishes that the word has a real basis in English usage. But in his *Glossographia* Blount varies the application of this rather crude device; he addresses the reader, points him towards sources of further information, indicates frequency, adds first-person comments and so on. The following list illustrates some of the lexicographical formulae evolved by him.

1. Appending the name of an author Addition of an author's name in italics at the end of an entry is easily the commonest mode of authentication. Blount was somewhat haphazard and inconsistent in the way he did this: *S. Wa. Ra.* means Sir Walter Raleigh, but then so does *Raw*. The fact that so many names and titles are truncated (*Bac., Dr Br., Cow., Dub. [=Du Bartas], Fel., Shep. Fa. Counc., etc.*) suggests a readership of those who would be familiar with the contemporary literary scene.

2. Appending the title of a work In other cases he will go further, and indicate the title of a particular work from which a word has been taken (either

with or without the name of the author). Thus for *aluminous* we have 'Vulg. Er.', for *polymorphean* he adds 'Tr. of Mass', for *mutes* he adds 'Dr Taylors Gram.', for *toman* 'Herb tra.' [i.e., Herbert's Travels] and (rather more vaguely) for *serenade* we are given 'Mr. Cowley in his Poems'. From a purely lexicographical point of view it is clearly of importance to know whether the word has been located in a specialist theological text, a general work of history, a poem, or some other work, though Blount may of course have intended the fuller reference simply as a courtesy to his reader.

3. Giving multiple references Some entries are graced by the name of two authors. For *piquant* we are given 'Master Montagu in his *Miscel.* and my Lord Bacon in his *Hen 7*, and the entry for the verb *amortize* reads '(from *mors*) to deaden, kill or slay. Lord Bacon and Chaucer' — here Blount may have felt the urge to supply the name of at least one modern writer.

4. Providing detailed finding-places At times, though not frequently, page-references are given. Thus, *amphoral* has the addition 'Godwin. 143', *cornage* 'Cam. Britan. fo. 609' and *cosmical* 'Vul. Er. 224'. This was for Blount not a matter of mere arid scholarly documentation; he was in many cases pointing to sources of useful background information. The entry for *Palmeto Wine*, for instance, reads as follows:

Palmeto Wine, is a sweet and pleasant juyce (like Muscadine or Allicant) coming from the *Palmeto* Tree in the Isle *Mauritius*. *Herb tr.* 210.

Any reader who troubled to turn to page 210 in Herbert's *Travels* would be rewarded by finding there an engraving showing the peculiar shape of the Palmeto palm, together with the information that its wine 'purges the belly and helps obstructions', and that by making incisions in three trees the travellers 'in lesse than one houre ... bowzed their bellies full'.

5. Use of formulae of the type 'it is used in ...' Sometimes Blount addresses information on usage more casually to his readers: of the word *adiaphorous* he comments 'Dr Taylor useth it in his *Liberty of prophecying*'. Of *habergion* ('a little coat of Mail') he says 'It is used in Scripture, *Rev. 9.9.*'; elsewhere we have 'I read it in the *Protestants Apology*' (*recreant*); 'I finde the word used in the *Elements of Armories*' (*parallelisation*). The general effect of such comments is not so very different from what we find under 1 or 2 above. But especially where the first person pronoun intrudes, it reveals to us the lexicographer at work with his books, at times suggesting reservations on his part, a mild surprise at having found the word there: as he says in his preface, it was not his purpose 'to become an Advocate for the use of such Words'.

6. **Citing an author as authority for additional information** The references given above all apply to the whole entry. But Blount sometimes shelters behind a named author for part of his definition:

Vitrification (*vitrificatio*) glassiness, or the making of Glass. *Vitrification* (says Dr. Brown) is the last work of fire, and when that arrives, humidity is exhaled ...

So also the poet Cowley is quoted as an authority for information about the temple built to the god *Janus* 'which in time of peace was shut, and in war open'; under *neophite* we read 'The Neophite or new Convert (says Mr Gregory, p. 111) received a Taper lighted, and delivered by the *Mystagogus*'; and under the word *hurricane* (spelt *haracana*) Howell is cited as a source of information for the belief that the devil often appears to pagans in a hurricane, and Herbert for the opinion that hurricanes occur in some countries once in every nine years.

7. **Directing the reader to a source of further information** Blount sometimes makes use of the second person pronoun. He thus addresses the reader directly in the entry on *probate*: 'See more of this, if you please, in *Cowel*'. Having explained that a *hayward* is a sworn official of a court, he adds 'and the form of his oath you may see in *Kitchin, fol. 46*', and in more general terms he explains that *pyramidography* is 'a writing, treating, or describing of *Pyramids*; whereof one Mr. *Greaves* has wrote a book.'

8. **Citing an author as authority for a particular meaning** *Septuary*, Blount says 'may be applied to any thing composed of the number seven' and adds a note of a Brownean sense: 'Dr. *Br.* uses it for a week, consisting of seven days'; *rebeck* is a fiddle, but '*Chaucer* uses it for an old trot'; *scytale* is said to have two distinct meanings, one being a field mouse, and the other a small staff 'and in this last sence my Lo. *Bacon* uses it, in his *Advancement of Learning, fol. 268*.'

9. **Citing authors' opinions on near-synonyms** The verb *to exhilarate* is given as 'to make one merry, to refresh', but Blount then goes on '*Exhilaration* (says my Lord *Bacon*) has some affinity with joy, though it be a much lighter motion'. He also gives Sir Walter Raleigh's views on the distinction between *predestination*, *providence* and *prescience*, and quotes no less an authority than King James in his explanation of *necromancer* when he says 'The difference between *Necromancers* and *Witches* *K. James* (in his *demonologie*) hath taught in a word; the one (in a sort) command, the other obey the Devil'.

10. **Citing authors for variant derivational forms** In the seventeenth century, variant derivational forms were rather commoner in English than they are now. You could say either *pyramidal* or *pyramidical*, *regress* as well as *regression* could

be used for the noun, and there were many others. In the case of *exploremēt* and *exploration* (bracketed together with the definition 'a search, a tryal or searching out') Blount thought it necessary to say 'Doctor *Brown* useth both in his *V. Er.*' The evidence of the OED shows that he indeed does so: the text of the *Vulgar Errors* provides the first attestation of the form *exploremēt*, and also contains *exploration* in the now obsolete sense in which Blount defines it: a striking instance of acute textual observation on the part of the lexicographer. Elsewhere Blount notes the word *felicitate* (from Felltham), adding 'and I have read *Felicify*' — a pity he does not say where he had read it, since the earliest attestation in the OED is from Hooker in 1683.

11. Citing an author for a particular spelling Under *wea-bit* or *way-bit* Blount notes 'I find it w[r]itten *Wea-bit* in Mr *Cleavelands Poems*, but I should rather think it *Way-bit*; *quasi*, a *bit* or part of the *way*.' Clearly he was opting here for the spelling that would be most helpful to a reader, though he still puts the Cleveland form in the prime alphabetical position. The 'intervention' of the compiler (involving the use of the first-person pronoun) to express such personal preferences was not uncommon in the early dictionaries (Stein 1986).

12. Citing an author as evidence for word frequency Under *peere*, that is *pier* (which Blount says rather tentatively 'seems properly to be a Fortress made against the force of the Sea, for the better security of ships') he notes 'You shall read the word often in *Sands Travels*'. A few pages earlier, he notes that the Spanish word *parabien* (for 'a welcoming') is 'often used in *Ariana*'. Such comments again reveal that kind of textual and lexical alertness which compilers (and contributors to dictionaries) have always needed.

Lexicographical Models

The practice of thus authenticating words in a monolingual dictionary was quite unknown to Blount's predecessors Cawdrey, Bullokar and Cockeram. As for so many other features in the monolingual vernacular dictionaries of the renaissance, the model for this particular innovation is probably to be sought in the well-established Latin and bilingual dictionaries of the period.

Latin and Greek dictionaries such as those of Scapula (1652) — who is referred to by Blount in his preface — very commonly noted the names of classical writers as authority for the inclusion of particular words or word senses. In the Latin polyglot dictionary of Calepinus (1644) lists are provided in the front matter of the authors thus cited in the body of the work; these are divided (according to the conventional classification) as belonging to the *Golden Age* (for instance, Plautus, Caesar, Livy), the *Silver Age* (Seneca and others), the *Iron Age*, etc. The user would thus have ready guidance (if he wanted it) on the degree of stylistic respectability conferred by the names of those authors referred to in

the body of the dictionary.

A calibrated authority-scale of this kind would hardly do today, though dictionary compilers may still attach greater weight to attestations from recognized writers than from (say) obscure journalists. But in the seventeenth century, when the everyday Latin of scholarly communication had become so contaminated by the practical needs of the modern world, at least the broad distinction between classical Latin and non-classical Latin ('Latin Barbarous') was of importance, and this is indeed reflected in the titles and make-up of dictionaries such as that of Littleton (1693). The two bilingual Latin and English dictionaries which Blount acknowledges as sources both carry on this tradition of naming classical authors. Thomas (1587) has entries such as

Bucolicus, a, um. *Ovid.* Pertaining to neate or oxen.
Citissime, Adverb. Plin. By & by, incontinent.

Roughly 60 percent of all entries include such references to 'authoritative' writers. The dictionary of Rider (1589) — far more extensively drawn on by Blount — is bi-directional. The English-Latin part (which forms the bulk of the work, with full definitions) contains no references to writers at all. But these are included in the smaller Latin-English section, wedged in between the headword and the column-references to the main entries in the English-Latin part:

Detrunco, As, Accusat. Cicero. 136,49. 384,50. 385,45 ...
Hypopygium, i. locus sub turri. Calep.
Populus, li., f. Virg. 1076,12.

Here, beside Cicero and Virgil, we have a reference to the dictionary of Calepinus: the very same combination of primary and secondary sources that is to be found in Blount.

When we turn to the vernacular bilingual dictionaries which are known to have been used by Blount, there is little to be found which might have served as a model: nothing in the Dutch and English Hexham (1647); in Florio's Italian and English dictionary (1611) a fine list of about 200 authors and books 'that have been read of purpose for the collecting of this Dictionarie', but no actual references in the lemmata; and in the French and English dictionary of Cotgrave (1611) not much more than the occasional note 'Rab.' against some of the more remarkable Rabelaisian creations (*démantibulé* 'unjawed, or, whose jaws are dasht in pieces', *magistrostralement* 'dunsically').

It seems then likely that Blount adopted the practice of authenticating words in his dictionary from existing patterns in Latin lexicography, though (as we have seen) he used the device less mechanically than did his originals.

Blount and the English lexicographical tradition

When seen in the context of later lexicographical history, Thomas Blount's *Glossographia* may be looked upon as a hybrid. It cannot by any stretch of the imagination be called a historical dictionary: it is a smallish work for everyday use, aimed at the educated, perhaps, but hardly at scholars. Yet the names of authors and source references are included in the lemmata after the fashion which we have now come to associate with the requirements of historical lexicography. When Blount names authors, these are very commonly his near contemporaries — writers of travel books, histories, works on theology, architecture, heraldry and so on which will have formed the common reading matter of the educated gentleman of his day. As we have seen, he was in fact singularly successful in picking up early citations for lesser-known words.

Unlike the compilers of our great historical dictionaries, to whom *all* the words of the language equally deserve scholarly documentation, Blount authenticates words only *selectively* — in about one in twenty of all the entries. It would certainly have seemed nonsense to him to quote 'authorities' for a common and uninteresting word such as *wicket* ('a small gate'), which Johnson was later to provide with a battery of quotations from Spenser, Milton, Dryden and Swift.

As the survey given above will have shown, the words for which Blount *did* provide attestations were especially those which were new, exotic, rare, specialized in their sense or somewhat technical. These are, of course, the very types of words for which general dictionaries of the language always have been, and always will be most often consulted. But at mid-seventeenth century, when the vocabulary of English was in a peculiar state of flux, and when dictionaries contained large numbers of ghost words — mainly thinly-disguised Latinisms created by compilers to swell their vocabulary — there was real instability in the vocabulary, genuine doubts about whether certain words could be said to 'exist' in the language or not (Jones, 1953). The habit of appending glossaries to scholarly writings in the vernacular had been widespread in the sixteenth century, and these will in some sense have served to legitimize words which might otherwise have been called in doubt. But by Blount's time this practice had gone into a decline (Schäfer 1989) thus creating a gap which the new monolingual dictionaries could fill. The authentication of usage was then not a mere scholarly fad or a lexicographical luxury. The seventeenth-century reader really needed to know: and if Bacon, or Cowley, or Sir Henry Wotton had used the words, then it was quite safe for him to do so too.

The *Glossographia* went through five editions down to 1681, and the pattern of authors named in it underwent a process of gradual modification which cannot be looked at now. If we discount the largely derivative, and anonymous, *Glossographia Anglicana Nova* (1707) Blount has found no significant imitators. A great innovator, he is thus to be found citing 'authorities' for particular words a century before Johnson. Yet there is no reason to suppose that Johnson

(who knew the Latin tradition well enough) copied the idea from him, extending its cover and adding in quotations as well. The hybrid dictionary of Blount stands on its own, but his distinctive and original technique of *selective authentication* made perfectly sound lexicographical practice in his own day, and he used it with both intelligence and discrimination.

References

(a) Dictionaries

- Ayto, J. and J. Simpson. 1992. *The Oxford Dictionary of Modern Slang*. Oxford: Oxford University Press.
- Blount, Thomas. 1656. *Glossographia*. London: Thomas Newcomb.
- Blount, Thomas. 1670. *A Law Dictionary*. London: Thomas Newcomb.
- Bullockar, John. 1616. *An English Expositour*. London: John Legatt.
- Calepinus, Ambrosius. 1644. *Dictionarium septem linguarum*. [Editio novissima]. Venice: Ex Typographia Baretiana.
- Chambers English Dictionary. 1995. Edinburgh: Chambers.
- Cotgrave, Randle. 1611. *A Dictionarie of the French and English Tongues*. London: Adam Islip.
- Cowell, John. 1607. *The Interpreter: or Booke Containing the Signification of Words*. Cambridge: John Legate.
- Dasypodius, Petrus. 1537. *Dictionarium latinogermanicum et germanicolatinum*. Strasburg: W. Riheilius.
- Florio, John. 1611. *Queen Anna's New World of Words, or Dictionarie of the Italian and English Tongues*. London: Melchior Bradwood.
- Hexham, Henry. 1647. *A Copious English and Netherduytch Dictionarie*. Rotterdam: Aermout Leers.
- Johnson, Samuel. 1755. *A Dictionary of the English Language*. London: W. Strahan.
- Littleton, Adam. 1693. *Linguae Romanae Dictionarium Luculentum Novum ... Representing I. The English Words and Phrases before the latin ... II. The Latin Classic ... III. The Latin-Proprie Names ... IV. The Latin Barbarous ... V. The Law-Latin*. Cambridge: W. Rawlins.
- Minsheu, John. 1617. *Ductor in linguas. The Guide into Tongues*. London: John Browne.
- Rider, John. 1589. *Bibliotheca Scholastica. A Double Dictionarie*. Oxford: Joseph Barnes.
- Scapula, Johannes. 1652. *Lexicon Graeco-Latinum ... Editio nova*. Leiden: J. Blaeu.
- Thomas, Thomas. 1587. *Dictionarium linguae Latinae et Anglicanae*. London: Richard Boyle.
- Tulloch, S. 1992. *The Oxford Dictionary of New Words*. Oxford: Oxford University Press.

(b) Other sources

- Bongaerts, T. 1978. *The Correspondence of Thomas Blount (1618-1679). A Recusant Antiquary*. Amsterdam: APA-Holland Universiteits Pers.
- Congleton, J.E. and E.C. Congleton. 1984. *Johnson's Dictionary. Bibliographical Survey. 1746-1984*. Terre Haute: Dictionary Society of North America.

- De Vries, C.M. 1994. *In the Tracks of a Lexicographer. Secondary Documentation in Samuel Johnson's Dictionary of the English Language (1755)*. Utrecht: LED.
- Jones, R.F. 1953. *The Triumph of the English Language*. London: Oxford University Press.
- Kerling, J. 1979. *Chaucer in Early English Dictionaries*. Leiden: University Press.
- Osselton, N.E. 1995. *Chosen Words. Past and Present Problems for Dictionary Makers*. Exeter: University Press.
- Schäfer, J. 1978. 'The working methods of Thomas Blount'. *English Studies* 59, 405-8.
- Schäfer, J. 1989. *Early Modern English Lexicography*. Oxford: Clarendon Press.
- Starnes, D.T. and G. Noyes. 1946. *The English Dictionary from Cawdrey to Johnson 1604-1755*. Chapel Hill: University of North Carolina Press.
- Stein, G. 1986. 'Definitions and first person involvement in Thomas Elyot's Dictionary'. In D. Kastovsky and A. Szwedek (Eds.). *Linguistics across Historical and Geographical Boundaries*. Berlin: Mouton, 1465-74.

Die negende deel van die *Woordeboek van die Afrikaanse Taal*

Ilse Feinauer, *Universiteit van Stellenbosch, Stellenbosch, Suid-Afrika*

Abstract: *Woordeboek van die Afrikaanse Taal Volume IX.* Volume nine of the *Woordeboek van die Afrikaanse Taal (WAT)* is the first product of a completely new system used by the Bureau of the WAT: it is the first volume of the new chief editor, the editorial procedure was overhauled entirely, and WAT IX was the first volume to be produced fully by desktop publishing. This review article wants to investigate whether the new system really brought about lexicographical improvement as such. The general form and presentation of this dictionary are discussed, but the focal point of the article is the alphabetical part of the dictionary. The ways in which the latest lexicographic theory is embodied in WAT IX, as well as the extent to which WAT IX accommodates criticism on previous volumes are also discussed. Concerning form and presentation, WAT IX is improved to such an extent that it features as one of the most user-friendly Afrikaans descriptive dictionaries. Macrostructural gains in WAT IX are the extensive inclusion of sublexical and multilexical lemmas, and the ordering of abbreviation lemmas. Microstructurally the lemmas are dealt with more thoroughly and a larger spectrum of linguistic information is included. Unfortunately the improvements are not applied consistently in WAT IX.

Keywords: ABBREVIATIONS, AFFIXES, CITATIONS, DATABASE, EDITORIAL USAGE EXAMPLES, EXAMPLES, HOMONOMY, IDIOMS, LEXICAL LEMMA, LEXICAL MEANING, LEXICOGRAPHY, MACROSTRUCTURE, MICROSTRUCTURE, MULTILEXICAL LEMMA, OVERALL-DESCRIPTIVE DICTIONARY, POLYSEMY, PRONUNCIATION GUIDANCE, RACISM, REVIEW, SEXISM, SUBLEXICAL LEMMA, TYPOGRAPHY

Opsomming: Deel nege van die *Woordeboek van die Afrikaanse Taal (WAT)* is die eerste produk van 'n heel nuwe bestel van die Buro van die WAT: dit is naamlik die eerste deel van die nuwe hoofredakteur, die redaksionele stelsel is ingrypend hersien, en WAT IX is volledig as kantoorpublikasie tot stand gebring. Hierdie resensieartikel wil vasstel of die nuwe werksbenadering werklik vernuwing in die woordeboek as sodanig teweeggebring het. Daar word ondersoek ingestel na die algemene vorm en voorkoms van die woordeboek, maar die artikel konsentreer hoofsaaklik op die alfabetiese deel van die woordeboek. Aandag word gegee aan die toepassing van die jongste leksikografieteorie, maar ook aan die mate waarin WAT IX aanpassings gemaak het op grond van kritiek wat op vorige dele gelewer is. Wat vorm en voorkoms betref, is WAT IX so aangepas dat dit een van die gebruikervriendelikste Afrikaans verklarende woordeboeke is. Makrostruktureel toon WAT IX leksikografiese wins in sy opname van subleksikale en multileksikale lemmas, en die plasing van afkortinglemmas. Mikrostruktureel word die lemmas beter bewerk en die spektrum van taalkundige inligtingtipes is meer uitgebrei. Die verbeterings word egter nie oral konsekwent deurgevoer nie.

Sleutelwoorde: AFFIKSE, AFKORTINGS, DATABASIS, HOMONIMIE, IDIOME, LEKSIKALE BETEKENIS, LEKSIKALE LEMMA, LEKSIKOGRAFIE, MAKROSTRUKTUUR, MIKROSTRUKTUUR, MULTILEKSIKALE LEMMA, OMVATTENDE WOORDEBOEK, POËME, POLISEMIE, RASSISME, RESENSIE, SEKSISME, SITATE, SUBLEKSIKALE LEMMA, TIPOGRAFIE, UITSPRAAKLEIDING, VOORBEELDMATERIAAL

1. Inleiding

Die negende deel van die *Woordeboek van die Afrikaanse Taal* open vir die Buro van die WAT 'n wêreld van eerstes: Dit is die eerste deel onder die hoofredakteur dr. D.J. van Schalkwyk, nadat die redaksionele stelsel ingrypend hersien is, dit is die eerste deel in die geskiedenis van die Woordeboek wat volledig as kantoorpublikasie op rekenaar tot stand gebring is, en dit is die eerste deel wat in net minder as drie jaar voltooi is — dit terwyl deel nege die letter L volledig dek.

Hierdie resensieartikel wil vasstel in watter mate die nuwe werksbenadering 'n vernuwing in WAT IX teweeggebring het, nie net wat die inhoud nie, maar ook wat die algemene vorm en voorkoms betref. Groter klem sal op die alfabetiese deel van die woordeboek (477 bladsye en meer as 6 500 verklaarde en ongeveer 2 500 onverklaarde lemmas) gelê word. Hier sal veral aandag gegee word aan die toepassing van leksikografieteorie in die hantering van enkele woordeboekartikels, en aan die mate waarin WAT IX hom steur aan kritiek wat op die voorafgaande dele gelewer is.

2. Vormlike vlak

In die Inleiding (p. i) meld die Hoofredakteur dat die redaksionele stelsel sterk vereenvoudig is met o.a. as doel om 'n "toegankliker produk vir die gebruiker te lewer". Saam met die volledige rekenarisering wat opwindender tipografiese moontlikhede bied, het die Redaksie bepaald daarin geslaag om 'n heelwat toegankliker produk te lewer. Op die vormlike vlak toon WAT IX verbeteringe wat dit een van die gebruikervriendelikste Afrikaanse woordeboeke ooit maak.

2.1 Toeligting by die gebruik van die woordeboek

Die belangrikste vernuwing by die **Toeligting** is die skematiese voorstelling van verskeie artikelinskrywings (pp. x-xi): Die skema maak deel uit van die afdeling **Terminologieverklaring** waar 'n verduideliking van die vernaamste leksikografiese vakterme gegee word. Reeds in deel VIII bring die verklaring van leksikografiese terme hier groter duidelikheid vir die gebruiker, maar met behulp van die skema in deel IX maak die terme nou leksikografies sin deurdat

dit as woordeboekteks aangebied word. Die doel en status van elke onderafdelinkie word aangetoon, en die gebruiker word terselfdertyd verwys na die gepaste paragraaf in die *Toeligting* waar die bepaalde tipe inskrywing bespreek word.

Die verklaring van die etikette (p. xiv) wat in deel IX gebruik is, word ook verfynd. Reeds deur die gebruik van die korrekte leksikografiese term *etiket* word die gebruiker beter leksikografiese teoretiese ingelig as wat in vroeëre dele die geval was. Terwyl deel VIII slegs *gebruiklikheidsaanduiding* in die *toeligting* bespreek, tref die gebruiker in deel IX ook etikette aan wat die lemma binne 'n spesifieke register of 'n spesifieke spesialiteitsveld plaas. Die agt *gebruiklikheidsetikette* van deel VIII word in deel IX afgeskaal tot vier: sodoende kan die gebruiker makliker daartussen onderskei. Vergelyk hier Feinauer (1993: 289-290) se kritiek op die aanduiding van *gebruiklikheidsetikette* in die *toeligting* van deel VIII. Wat wel steeds in deel IX ontbreek, is die gebruik van etikette wat moderner lemmas en neologismes aandui.

'n Belangrike amendement by die *Toeligting* is dat die lys *Redaksionele Afkortings* (p. xxi) in deel IX die erkende afkortings van die *Afrikaanse Woordelys en Spelreëls* gebruik. *w.*, *b.n.*, *s.*, *m/d* en *v/d* word nou tereg met *ww.*, *b.nw.*, *s.nw.*, *m.d.* en *v.d.* vervang. Afgesien daarvan dat die Redaksie hierdeur toon dat hulle gehoor gee aan vroeëre kritiek (o.a. Gouws 1985 en Feinauer 1993), is hulle ook hierdeur die gebruiker meer ter wille; hy of sy hoef nie langer net vir die *WAT* 'n ander stel afkortings na te slaan of te memoriseer nie.

2.2 Tipografie

Die heelwat beter ordening van die onderskeie inskrywings in elke artikel help die gebruiker om die inligting daarin moeiteloos te ontsluit. Onder andere word die makliker ontsluiting bewerkstellig deur bloot elke inligtingstipe op 'n nuwe reël te begin. Die ruimte wat die leksikograaf hierdeur inboet, wen hy aan toeganklikheid vir die gebruiker. Ander tegnieke om die leesbaarheid van die artikels te verhoog, is die plasing van die homoniemonderskeidingsnommers voor die lemma, en daar volg geen leestekens na die lemma of enige ander syfer wat die polisemiese onderskeidings rangskik nie. Dit sorg vir 'n heelwat rustiger bladspieël in deel IX as in vorige dele. Struktuuraanwysers soos ≈ en die vierkantige blokkie ■ is ook gebruik in plaas van ekstra stelle hakies of ander tipes syfers wat heelwat meer verswarend sou wees. Volgens Otto (1994) maak hierdie tipe merkers iets makliker vindbaar met die oog.

Elke uitdrukking wat in 'n woordeboekartikel verskyn, word in *WAT* IX ook aan die begin van 'n reël gelys, maar om die vindbaarheid van uitdrukkinge verder te verhoog, word hulle deur soekwoorde ingelui. Die soekwoord sowel as die uitdrukking word vet gedruk. Afgesien van die groter vindbaarheid wat die vetdruk van uitdrukkinge meebring, het dit ook 'n baie belangrike leksikografiese funksie: vir die eerste maal word uitdrukkinge nou volledig

bewerk. In WAT IX het alle uitdrukkings, soos reeds in deel VIII ook, volle verklarings en voorbeeldsinne waarin die betrokke uitdrukking optree. Vergelyk par. 3.1.3 vir 'n bespreking van WAT IX se hantering van multileksikale lemmas.

3. Inhoudelike vlak

Die toets vir WAT IX se geslaagdheid as omvattende woordeboek sal hoofsaaklik op die inhoud berus, alhoewel die vorm natuurlik ook belangrik is. Volgens Gouws (1989: 68) is dit "die ideaal dat so 'n woordeboek die hele woordeskat van 'n taal moet weergee en dat die artikels die volle spektrum van taalkundige inligtingstipes moet bevat". Daar sal nou aandag gegee word aan WAT IX se woordversameling; die makrostrukturele inhoud, en aan enkele taalkundige inligtingstipes, die mikrostrukturele inhoud.

3.1 Makrostrukturele inhoud

Hier gaan dit om hoe omvattend die woordversameling is wat in WAT deel IX betrek is, maar ook om watter ander tipe leksikale items volwaardige lemma-status kry, bv. subleksikale lemmas, multileksikale lemmas, en afkortings as lemmas.

3.1.1 Woordversameling

Omdat die WAT 'n omvattende sinchroniese woordeboek is, moet sy woordversameling die mees verteenwoordigende wees. Volgens Gouws (1994) bied WAT IX die volledigste optekening tot nog toe van leksikale items wat met die letter L begin, en ook vir Otto (1994) stel WAT IX nie teleur wanneer dit by volledigheid kom nie. Dit is te verwagte, want die databasis van die Buro van die WAT word as die mees verteenwoordigende versameling Afrikaanse woorde en uitdrukkings beskou. Vergelyk hier ook die Hoofredakteur op p. iii: "Die Buro beskik reeds oor die mees omvattende en verteenwoordigende streektaalversameling."

Enkele voorbeelde word aangehaal om te illustreer dat 'n groot spektrum van Afrikaanse woorde gedek word:

Streektaalvorme: lepaai, liefman, liefvrou, lafskaai

Groepsvorme: lelie (weermagstaal), laitjûr (visserstaal), langaan (vissers-
taal), lambaar (Kaapse Maleiers)

Geselstaal / omgangstaal: leste, liewerste, liewerster, langes, laitie, lady,
ladismitter, Ligbloue

Verouderde vorme: laerlêer, ledikant, lemier

Vaktaal: lanugo, laparoskopie, laktoon, labiodentaal

Leenvorme uit ander tale as Engels: leotard, l'envoi, Lectori Salutem

Nuutskoppinge / moderne vorme: luikrugmotor, luisterlied, laserdrukker, liefdesbaba

Dit spreek vanself dat naas bogenoemde tipes algemene neutrale Afrikaanse woorde ook opgeneem is.

Volgens Geeraerts (1991: 291) is die leksikografie "not a disembodied activity taking place in a vacuum: it is an act of communication in a real-world context, and both aspects are essential to the lexicographical decisions to be made". Hoe lyk hierdie "communication in a real-world context" vir Afrikaans? Wat betref die omgangstaal, oftewel informele geselstaal, is Afrikaans deurspek van Engelse woorde. Gouws (1995: 303-304) merk heeltemal tereg op dat indien 'n Engelse leenwoord reeds geïnternaliseer is as deel van die Afrikaanse woordeboek moet dit veral in informatiewe omvattende woordeboeke soos die WAT opgeneem word. Die WAT IX slaag in 'n groter mate as vroeëre dele daarin om erkenning te gee aan die Suid-Afrikaanse werklikheid deur bv. die opname van leksikale items soos *lady*, *ladyskap*, *ladybird*, *lanie*, *lorrie*, *lewwentrie*, *laitie*. Vergelyk ook die omvattende polisemiese en homonimiese opname van die onderskeie geselstaalvorme van *lyk*:

²lyk 2 (*geselstaal*) Begrafnis: ...

³lyk ... ww., het gelyk. (*geselstaal*) Hou van, of daarvan hou: ...

⁴lyk ... ww., het gelyk. (*streektaal*) Sê of te kenne gee: ...

In geen ander Afrikaanse verklarende woordeboek word aandag aan enige van hierdie betekenis van *lyk* gegee nie, terwyl hulle, met die uitsondering van ⁴lyk frekwent gebruik word.

Die WAT IX voldoen egter nog nie ten volle aan een van die belangrikste vereistes volgens Gouws en Ponelis (1992: 87) vir 'n omvattende eentalige woordeboek nie, naamlik om die taalgebruik van die verlede en die hede te weerspieël nie. Wat die taalgebruik van die verlede betref, is die opname in die WAT IX bo verdenking: meer as genoeg ouer leksikale items word aangetref. Die probleem kom by die taalgebruik van die hede voor: daar word nie naastenby genoeg Engelse leenwoorde opgeneem nie. Die geweldige invloed van Engels op die Afrikaanse spreektaal word nie voldoende gemanifesteer nie. Vergelyk byvoorbeeld die volgende leksikale items wat in deel IX opgeneem kon word, weliswaar met die etiket *geselstaal / omgangstaal*: *lover*, *lift*, *looks*, *lagoon*, *lawn*, *lawnmower*, *lumberjacket*, *landlord / -lady*, *laundromat*, *leopard crawl*, *lifesaver*, *loaf*, *loafer*, *lobbie*, *life jacket*, *lob* (tennis), *locker*, *long drop*, *luckie*, *luck*, *luckie dip*, ens. Meerwoordige leksikale items wat opgeneem sou kon word, is bv. *love* en *charity* en *likes* en *dislikes*. Indien die realiteit van die grootskaalse ontlening aan Engels in die Afrikaanse omgangs-

taal weergegee moet word, kan die WAT nie lukraak enkele Engelse woorde opneem en talle ander buite rekening laat nie.

Die oorsprong van hierdie dilemma kan tweeledig wees. Eerstens kan dit wees dat die WAT soos heelwat eentalige Afrikaanse standaardwoordeboeke, die taal van die ideale Afrikaansspreker reflekteer en nie dié van die tipiese spreker nie (Gouws 1995: 305). Indien die WAT hom hieraan skuldig maak, geld dit net ten opsigte van die stigmatisering en gevolglike weglating van Engelse leenwoorde wat weer op oordrewe purisme en onvanpaste voorskryfbaarheid dui. Deurdát WAT IX en in 'n mindere mate ook deel VIII reeds begin het met die opname van hierdie leksikale items moet daar wat laasgenoemde dele betref, 'n ander sleutel vir die probleem gesoek word, naamlik die databasis.

Een van die grootste probleme waarmee leksikograwe gekonfronteer word wanneer hulle 'n databasis saamstel, is volgens Gouws (1991: 279) dat dit primêr uit geskrewe taal bestaan en nie uit gesproke taal nie. Om enigsins as ekstensiewe databasis gereken te word, móét daar voldoende optekening uit sowel die skryf- as spreektaal wees. In 1994 toon Gouws steeds skeptisisme oor die werklike verteenwoordigendheid van die WAT-databasis. Dié skeptisisme word grootliks gedeel, hoofsaaklik omdat talle hoogs frekwent gebruikte leksikale items uit die omgangstaal nie opgeneem is nie, maar minder bekende of minder frekwent gebruikte items wel. Om enkele voorbeelde te noem: *lady-skap* word tereg opgeneem, maar die alledaagser *ladylike* en *ladiesman* nie. Die heel vreemde *levoor* as geselstaalvorm vir *ivoor* word opgeneem, maar *lipstiek* as heelwat gebruikliker benaming vir *lippestif*, *lippestiffie*, *lippeverf*, *lipsel*, *lipstif*, *lipstiffie*, of *lipverf* nie. Drie van die lemmas, naamlik *lippestif*, *lippestiffie* en *lipsel* het boonop etikette wat hulle as minder gebruiklik merk. *Lipstiek* is selfs al in skryftaal opgeteken. Vergelyk *De Kat* van November 1994: 54 vir voorbeelde soos *rooi lipstiek*, *verslaaf aan lipstiek*, *om lipstiek aan te sit*, *lipstiekgrap*, ens. Soms lewer die databasis wel goeie omgangstaalvorme op, maar dan geskied die optekening nie volledig genoeg nie. By *lanie* word slegs die naamwoordvorm opgeneem terwyl *lanie* meer frekwent in die spreektaal as adjektief gebruik word. Alleen *lewwentrie* en *lewwie* word aangetref, maar nie die algemeen gebruikte *lewwetrie* nie. Dit word nie eers as wisselvorm by *lewwentrie* aangedui nie. Ook by gewone woorde word die tekortkominge van die databasis duidelik wanneer *lietsjie*, *lietsjieboom*, *lietsjiemot* en *lietsjievrug* as verklaarde en *lietsjiebedryf* en *lietsjieboord* as onverklaarde lemmas verskyn, maar *lietsjiesap* is nêrens te vinde nie. Oor die rol van die databasis by die bewerking van gewone woorde word in par. 3.2.2 meer gesê.

Volgens die Hoofredakteur (p. i) is "Leksikale items slegs as lemmas opgeneem indien die bewyse in die databasis vir hulle bestaan as gevestigde Afrikaans bo verdenking was. ... Leemtes van hierdie aard wat in die Woordeboek mag voorkom, is dus toe te skryf aan tekortkominge in die databasis." Tot tyd en wyl die databasis van die WAT in alle opsigte die sosiolinguïstiese werklikheid van die Afrikaanse taalgemeenskap reflekteer, sal die samestellers van

komende dele hulle nie ten volle op die databasis kan verlaat, of hulle daardeur kan laat intimideer nie, en sal op eie inisiatief ook lemmas moet opneem. Die eerste prioriteit moet sekerlik wees om 'n meer omvattende woordeboek daar te stel, en nie om aan die databasis getrou te wees nie.

Die WAT het in die verlede kwaai kritiek verduur oor sowel die opname van vakterme as die definiëring van vakterme. Vergelyk bv. Gouws (1985: 15) en Feinauer (1993: 297) waar hulle dit het oor die uitvoerige hantering van vakterminologie en oor die verklaring van vakterme wat nie altyd verstaanbaar en bruikbaar is vir die gebruiker nie. Volgens Feinauer (1993: 297) is die hantering van vakterme in deel VIII beter as in die voorafgaande dele. In deel IX word vakterme opgeneem soos van 'n omvattende woordeboek verwag word, maar volgens die Hoofredakteur (p. i) word "hoogs gespesialiseerde vakterme wat slegs deur die vakspesialis benut word, nie opgeneem nie. Vaktaal is slegs opgeneem indien daar bewyse in die databasis gevind is dat sulke terme ook in die algemene spreek- en skryftaal gebruiklik is". Indien hierdie opname van vakterme konsekwent toegepas is, sou dit beteken dat vakterme aansienlik minder sou wees as in voorafgaande dele. Op die oog af lyk dit asof dit inderdaad die geval is. 'n Voorbeeld waar die vakterm opgeneem is met sitate uit die algemene skryftaal is bv. *laparoskoop*, *laparoskopie* en *laparoskopies* met sitate uit onderskeidelik *Die Burger*, *Huisgenoot*, en *Sarie*. Direk daaronder verskyn egter *laparotomie* met sitate alleen uit die tydskrifte *Geneeskunde* en *Agro-animalia*. Hier is dus 'n vakterm met geen optekening in die algemene skryftaal nie, deel van die makrostruktuur. Dieselfde gebeur met die vakterm *labialiseer* met vier sitate uit taalkundige artikels en werke. Geen voorbeeld van 'n vakterm kon opgespoor word met enige poëem of sitaat uit die spreektaal nie. Indien konsekwent slegs gebruiklike vakterme uit die algemene skryftaal opgeneem word, sal die vakterminologie in komende dele van die WAT verder verminder.

Ten gunste van die hantering van vakterminologie in deel IX kan gesê word dat die definiëring sodanig aangepas is dat dit bruikbaar en verstaanbaar vir die gewone gebruiker is. Hier gee WAT IX dus duidelik gehoor aan vroeëre kritiek. Vergelyk bv. die volgende definisies wat vir die leek deursigtig behoort te wees:

ligand ... Enigeen v.d. atome of atoomgroepe wat a.d. sentrale atoom, gew. 'n metaalatoom, verbind is om koördinasieverbindings te vorm: ...

labiodentaal ... I ... Spraakklank wat m.d. botande teen die onderlip gevorm word: ...

3.1.2 Subleksikale lemmas

Deurdat daar in WAT IX ruim opnames is van subleksikale en multileksikale items met volle lemmastatus word gestalte gegee aan die opnamebeleid wat deur die Hoofredakteur op p. i gestel word, naamlik: "dat die lemmakeuse 'n

weerspieëling moet wees van die leksikale items waaruit die woordeskat van Afrikaans bestaan". Die WAT is wat deel IX betref, dus beslis nie meer net 'n woordgerigte woordeboek nie.

Eerstens wat subleksikale items betref: WAT IX neem die hele reeks op wat in Gouws (1989: 85-96) geïdentifiseer word, naamlik woorddele wat sowel woord- as stamoptrede het, en waarvan die stamoptrede 'n semantiese pluswaarde ten opsigte van die woordoptrede het, woorddele wat slegs 'n stamoptrede het, asook affigale gevalle (p. ii). Affikse wat in WAT IX opgeneem is, is: **-lei, -lik, -liks, -ling, -lings, en -loos**. Mikrostruktureel word affigale lemmas tot 'n groot mate identies aan gewone leksikale lemmas bewerk: uitspraakleiding word byvoorbeeld deurgaans gegee. Een van die opvallendste verskille egter is dat die grammatikale inligting by affigale lemmas nie soos by leksikale items direk na die uitspraak volg nie, maar deurgaans as deel van die definiens aangebied word. Vergelyk bv.

-liks ...

1 Agterv. van afgeleide bw., m.d. bet. "elke", waarby die stam gew. 'n temporele s.nw. is, bv. *daagliks, halfuurliks, jaarliks, ...*

2 Agterv. van afgeleide bw. m.d. bet. "mate, graad of duur", waarby die stam 'n b.nw. is, bv. *grootliks, kortliks en nouliks*.

en

-loos ... Agterv. waarmee byv. afl. gevorm word, m.d. bet. "sonder" of "met 'n minimum van", bv. *aarselloos, aksentloos, betekenisloos, blaarloos, ...*

Hier word beslis die verkeerde leksikografiese praktyk met die bewerking van die affigale lemmas gevolg: die definiensdeel van die definisie moet gereserveer word vir die oordra van slegs semantiese inligting van die betrokke lemma. Daar bestaan geen regverdiging vir die vermenging van grammatikale besonderhede en semantiese inligting nie. Die grammatikale besonderhede hoort doodgewoon tussen ronde hakies na die uitspraakleiding soos die geval is by leksikale lemmas. Affigale lemmas hoef nie soos funksiewoorde (setsels, lidwoorde, voegwoorde) met hulle yl semantiese waarde gehanteer te word waar die definiens onder andere ook aandui hoe hulle gebruik word nie (Feinauer 1994: 66). Affigale lemmas het meer as genoeg semantiese waarde om 'n volwaardige mikrostrukturele bewerking te verdien. 'n Verdere punt van kritiek by die bewerking van affigale lemmas is dat hulle gebruik nêrens kontekstueel geïllustreer word nie. 'n Lys eenwoordsitate word wel verskaf om die affigeringsmoontlikhede te illustreer, maar geen kollokasies of sinne word as óf sitate óf poëme verskaf nie. Die semantiese of grammatikale waarde van die betrokke affiks word dus nie prakties belig nie. Alle affigale lemmas word minstens konsekwent hier verkeerd bewerk.

Hoe lyk die bewerking van subleksikale lemmas wat nader aan leksikale lemmas is, in dié sin dat hulle soms woordstatus kan hê? 'n Mens sou verwag dat hoe duideliker die woordstatus van die subleksikale lemma word, hoe nader sal die bewerking wees aan dié van leksikale lemmas. Dit gebeur egter nie: deurgaans word grammatikale inligting leksikografies onjuis as deel van die definiens aangebied. Vergelyk hier die hantering van die byna affigale **-logie** en die subleksikale lemma met sterk woordstatus **ligtekop-**:

-logie ...

1 Slotkomponent van selfst. komposita m.d. bet. "leerstelling, teorie of wetenskap", bv. *antropologie, biologie, ekologie, etimologie, ...*

2 Slotkomponent van selfst. komposita m.d. bet. "woord, rede, spraak of gesprek", bv. *doksologie, tetralogie, toutologie, trilogie.*

ligtekop- ... Aanvangskomponent van selfst. komposita, wat aandui dat die persoon wat m.d. tweede komponent benoem word, ligte hare het, bv. *ligtekopdogter s.nw., ligtekopnooi s.nw., ligtekopseun s.nw.*

Die streeplemma **-lap** is die enigste subleksikale lemma wat opgespoor is waar die woordsoortaanwysing soos by leksikale lemmas gewoon direk na die uitspraakleiding volg; daarnaas word daar ongelukkig steeds grammatikale inligting as deel van die definiens aangetref:

-lap [lap] s.nw. Slotkomponent van enkele pejoratiewe persoonsname wat 'n persoon aandui wat is, of doen soos deur 'n b.nw. of ww. as eerste lid benoem word, bv. *dronklap, paplap, smeerlap, suiplap.*

Wat by hierdie tipe subleksikale lemmas wel aangetref word wat nie by die affigale lemmas voorkom nie, is etikette. By sommige streeplemmas word hulle heel korrek tussen hakies voor die definiens aangetref. Vergelyk:

-liede ... (*verhewe; verouderend, behalwe in 'n aantal gevestigde ss.; minder gebruiklik*) Slotkomponent van selfst. komposita waarin **-liede** in bet. parallel is aan *lui ... edelliede, jongeliede, ...*

laeweerstand- ... (*ongewoon*) Sien LAEWEERSTANDS-, bv. *laeweerstandtoets s.nw.*

Direk onder bogenoemde lemma word **laeweerstand-** aangetref met gebruiklikheidstoeligmag ook deel van die definiens. Dit mag natuurlik net so min soos die grammatikale leiding wat deurgaans by streeplemmas definiensmateriaal uitmaak. Vergelyk:

laeweestands- ... Selde ook *laagweerstand-* en *laeweestand-*. Aanvangskomponent van selfst. komposita, wat aandui dat dit wat m.d. tweede komponent benoem word, 'n lae weerstand besit of bied, ... bv. **laeweestandsdraad** s.nw., ...

Ook by spellingleiding wat by sommige streeplemmas gegee word, is daar inkonsekwente bewerking: by **ligtekop-** word die spellingleiding heeltemal korrek tussen hakies voor die definiens geplaas, maar by al die polisemiese onderskeidings van **laat-** word die spellingleiding voor die grammatikale inligting in die definiens geïnkorporeer.

ligtekop- ... (*selde ook los geskryf v.d. daaropvolgende s.nw.*) Aanvangskomponent van selfst. komposita, wat aandui dat die persoon wat m.d. tweede komponent benoem word, ligte hare het, bv. **ligtekopdogter** s.nw., **ligtekopnoui** s.nw., **ligtekopseun** s.nw.

laat- ...

I 1a Selde ook los geskryf v.d. daaropvolgende s.nw. Aanvangskomponent van selfst. komposita wat 'n laat of latere stadium van 'n geologiese tydperk ... aandui bv. **Laat-Barok** s.nw., **Laat-Bronstydperk** s.nw. ...

b Dikw. ook los geskryf v.d. daaropvolgende s.nw. Aanvangskomponent van selfst. komposita wat 'n laat of latere stadium i.d. geskiedenis of ontwikkeling van 'n taal aandui; ... bv. **Laat-Grieks** s.nw., **Laat-Hebreeus** s.nw. ...

2 Soms ook los geskryf v.d. daaropvolgende s.nw. Aanvangskomponent van selfst. komposita wat 'n laat stadium, voorkoms, handeling of ontwikkeling aandui; ... bv. *laataand*, *laatgeld*, *laatherfs* ...; **laatbesmetting** s.nw., **laatbestryding** s.nw., **laatbetaling** s.nw. ...

II 1 Selde ook los geskryf v.d. daaropvolgende b.nw. Aanvangskomponent van byv. komposita of afleidings wat iets aandui in 'n laat of latere stadium ... bv. **Laat-Bisantyns** b.nw., **Laat-Frans** b.nw., **Laat-Goties** b.nw. ...

Afgesien van die foutiewe plasing van spellingleiding in die definiens van **laat-** word daar aan die einde van die artikel as opmerking heelwat bykomende spellingleiding gegee; dié keer leksikografies heeltemal korrek gehanteer.

By hierdie subleksikale lemmas word net soos by die affigale lemmas geen kontekstuele leiding in die vorm van poëme of volsinsitate gegee nie. Die gebruiksliding is net soos by die affigale lemmas in die vorm van lyste eenwoordsitate waarin die subleksikale lemmas se optrede as aanvangs- of slotkomponente van komposita geïllustreer word. Hierdie eenwoordsitate word elders wel volledig behandel. By sommige subleksikale lemmas soos **ligtekop-**, **laeprys-** en **laeweestand-** egter, neem die voorbeeldmateriaal die vorm aan

van onverklaarde of sublemmas: hulle is vet gedruk en grammatikale leiding sowel as hoofklemaanduiding word verskaf.

Afgesien van die kritiek wat uitgespreek is teen die bewerking van subleksikale lemmas moet die WAT IX gelukkigewens word met die makrostrukturele hantering van subleksikale lemmas. Alleen deur die konsekwente opname van affigale lemmas onderskei die WAT IX hom van vorige dele van die WAT en van ander eentalige Afrikaanse woordeboeke. Subleksikale lemmas word ook omvattend en gedetailleerd opgeneem. Polisemiese verskille (-lings) en selfs homonieme (¹-ling, ²-ling) word aangetoon. Vergelyk die subleksikale lemmas -lui wat net soos hulle woordekwiwalente korrek as drie homonieme aangebied word:

- ¹-lui ... Slotkomponent van byv. komposita waarin -lui in betekenis parallel is aan *lui* (¹LUI), bv. *aartslui*, *donkielui*, *vreklui*.
- ²-lui ... Slotkomponent van ww. komposita waarin -lui in betekenis parallel is aan *lui* (²LUI I), bv. *aflui*, *inlui*, *uitlui*.
- ³-lui ... Slotkomponent van selfst. komposita waarin -lui in betekenis parallel is aan *lui* (³LUI), en wat dikw. sinonieme komposita met -mense en (of) -manne, soms -liede en selde ook -mans, vorm deur met dieselfde aanvangskomponent as hulle te verbind, bv. *edellui*, *jongelui*, *hoëlui*, *kooplui*, ...

In die *Verklarende Handwoordeboek van die Afrikaanse Taal* (1994: 632) (voortaan HAT) word byvoorbeeld slegs die -lui aangetref wat die meervoudsvorm aandui by komposita soos *seelui* en *jongelui*. In HAT word die subleksikale lemma -lui ook verkeerdelik as homoniem saam met die leksikale lemmas *lui* georden:

- lui*¹ mv. ... Mense ...
- lui*² Agtervoegsel by 'n aantal s.nw. om die mv. aan te dui, bv. *jongelui*, *seelui*, *werklui*.
- lui*³ b.nw. en bw. Sonder lus om te werk of vir enige inspanning: ...

In WAT IX word bogenoemde leksikografiese fout dat leksikale en subleksikale lemmas in 'n homonimiese verhouding tot mekaar staan, nêrens aangetref nie. Die subleksikale lemmas word soos die leksikografiese teorie vereis, selfstandig gehanteer: die koppelteken aan die subleksikale lemma veroorsaak 'n vormlike verskil met die leksikale lemma. Daarom kan hulle volgens Gouws (1989: 138) nie as homonieme geklassifiseer word nie.

By enkele subleksikale lemmas, bv. *dié* wat van 'n eienaam afgelei is, word selfs etimologiese inligting aangetref. Vergelyk hier *lilliput*-:

- lilliput*- ... (na Lilliput, denkbeeldige land in J. Swift se *Gulliver's Travels*, waarin alles, ook die inwoners, baie klein is) ...

3.1.3 Multileksikale lemmas

Aangesien die hantering van uitdrukings as multileksikale lemmas in die verlede soveel leksikografiese aandag ontlok het, word hier hoofsaaklik op uitdrukings gekonsentreer. In WAT IX het alle uitdrukings soos reeds in deel VIII ook, volle verklarings en voorbeeldsinne waarin die betrokke uitdrukking optree: uitdrukings word dus heelwat omvattender bewerk as wat die geval is met sublemmas in die meeste Afrikaanse verklarende woordeboeke. Vergelyk Gouws (1990: 270) en Hausmann en Wiegand (1991: 349) vir die term **sublemma**.

'n Leksikografiese wins wat vernuwend in WAT IX aangewend word, is die gebruik van dubbelpuntlemmas. Dit is veral waardevol vir die aanduiding van uitdrukings wat nie in 'n bepaalde lemma se artikel geplaas kan word nie omdat die uitdrukking geen leksikale item bevat wat selfstandig as lemma kan optree nie. Die uitdrukking as geheel dien dus as 'n multileksikale lemma waarvan die eerste kernkomponent as alfabetiese soekwoord in die vorm van 'n dubbelpuntlemma optree. Wat Gouws (1990: 275) bepleit, word nou in WAT IX werklikheid. Vergelyk:

langbome: (na die) langbome (langboompies of soms lang bome) toe gaan, gedra word..., of by die langbome gaan bly idiomatiese uitdr. (*geselstaal*) Dood wees of gaan ...

Dat die idiomatiese uitdrukking in **ligte laaie** op twee plekke onder verskillende soekwoorde / dubbelpuntlemmas opgeneem word, toon weer eens WAT IX se gebruikersingesteldheid. Vergelyk:

ligte: in **ligte laaie** (**ligtelaaie**) of in **ligte laaie** (**ligtelaaie**) staan idiomatiese uitdr. ...

ligtelaaie: in **ligtelaaie** (**ligte laaie**) of in **ligtelaaie** (**ligte laaie**) staan idiomatiese uitdr. ...

Die gebruik van die dubbelpuntlemma by **langelaas:** en **langerlaas:** is leksikografies heeltemal regverdigbaar, maar wat vreemd opval, is dat **langelaas** en **langerlaas** selfstandig ook as lemmas opgeneem word, terwyl daar 'n definitiewe semantiese verband met **op langelaas** bestaan. Die WAT IX handel hier dus teen sy beleid vir die gebruik van dubbelpuntlemmas: hier is nou wel 'n leksikale item in die bepaalde uitdrukking wat selfstandig as lemma kan optree.

langelaas ... bw. (*ongewoon*) Meer dikw. *langerlaas*. Uiteindelik: ...

langelaas: op (selde na en ten) **langelaas** (**lange laas**) of op **langerlaas** en selde op **langlaas** idiomatiese uitdr. Uiteindelik: ...

langerlaas ... bw. Selde ook *langelaas*. Uiteindelik: ...

langerlaas: op langerlaas of op (selde na en ten) langelaas (soms lange laas) en selde op langlaas idiomatiese uitdr. Uiteindelik: ...

Die optrede van langelaas as gewone lemma naas langelaas: as dubbelpuntlemma laat die nielemmatiese optrede van loep nog meer opval. Alleen vier dubbelpuntlemmas van loep is opgeneem, naamlik:

loep: onder die loep bring idiomatiese uitdr. (*ongewoon*) Aan noukeurige ondersoek blootstel: ...

loep: onder die loep kom idiomatiese uitdr. (*minder gebruiklik*) Aan noukeurige ondersoek blootgestel word: ...

loep: onder die loep neem idiomatiese uitdr. Noukeurig ondersoek: ...

loep: onder die loep wees idiomatiese uitdr. (*ongewoon*) Aan noukeurige ondersoek onderworpe wees: ...

Loep as selfstandige lemma verdien wel opname in WAT IX met sy selfstandige betekenis van "vergrootglas". Bogenoemde vier dubbelpuntlemmas kon selfs binne die artikel van die leksikale lemma loep as idiomatiese uitdrukkinge gersorteer het: die semantiese verband met die vergrootglas is nog duidelik genoeg. Vergelyk hier bv. die hantering van die idioom vir langnekkaatjie soen in die artikel van die leksikale lemma langnekkaatjie. In 'n nie-omvattende woordeboek soos HAT (1994: 624) word loep wel as selfstandige lemma opgeneem. Hoeveel te meer hoort dit dan nie in WAT met hoogstens 'n etiket soos verouderend daarnaas nie?

Net soos streeplemmas nie in 'n homonimiese verhouding tot leksikale lemmas mag staan nie, kan dubbelpuntlemmas ook nie homonimies ten opsigte van leksikale lemmas gehanteer word nie: die dubbelpunt skep 'n vormlike verskil met die leksikale lemma. WAT IX tree hier nie heeltemal so konsekwent leksikografies korrek op soos by die hantering van die streeplemmas nie. Vergelyk die korrekte hantering van laas en laas: en die verkeerde homonimiese aanduiding van luisbos en luisbos:

laas [la:s]

I bw.

1 Die laaste keer of by 'n vorige geleentheid: ...

laas: op laas (ook opplaas) idiomatiese uitdr. Uiteindelik: ...

laas: vir laas idiomatiese uitdr. Vir die laaste keer: ...

¹**luisbos** ... s.nw. (*minder gebruiklik*)

1 Sien LUISIESBOS: ...

²**luisbos:** iemand met die luisbos eet, gooi, lek, werk, ... idiomatiese uitdr. (*streektaal*) Iem. m.d. kop stamp tydens 'n bakleiery ...

Die mikrostrukturele bewerking van multileksikale lemmas wyk net soos by subleksikale lemmas af van dié van leksikale lemmas. Hulle word dus steeds tot 'n mate as ondergeskikte lemmas beskou. Waar subleksikale lemmas deurgaans wel van uitspraakleiding voorsien is, word geen uitspraakleiding by meerleksikale uitdrukkings aangetref nie. Multileksikale lemmas soos leenwoordgroepe word wel van uitspraakleiding voorsien. Vergelyk *Lectori Salu-tem*. Afgesien hiervan word idiomatiese uitdrukkings ook nie grammaties benoem nie, behalwe dat daar deurgaans aangetoon word dat hulle idiomatiese uitdrukkings is. Anders as by die subleksikale lemmas verskyn die grammatiese aanduiding nie as deel van die definiens nie: dit word leksikografies heel korrek voor die definiens geplaas. Volgens Gouws (1990: 271) is dit in Afrikaanse woordeboeke grammaties onvanpas om multileksikale lemmas woordsoortelik te benoem: Afrikaans het 'n woordgebaseerde benoemingstelsel met terme soos **selfstandige naamwoord**, **byvoeglike naamwoord**, en **bywoord**. Dit behoort egter nie vir die leksikograaf 'n onoorkomelike probleem te wees nie, want multileksikale lemmas sou byvoorbeeld d.m.v. terme benoem kon word wat sinstukke beskryf, soos **naamwoordstuk**, **klaus**, of **adjektiefstuk**. Hier sou ook met vrug gefokus kon word op die funksie eerder as op die kategorie van die multileksikale lemmas deur bv. terme soos **bywoordelik**, **byvoeglik**, of **naamwoordelik** te gebruik. Laasgenoemde sou die gebruiker meer ter wille wees, want sodoende kry hy / sy meer onmiddellik toepasbare leiding. Vergelyk hier weer **langelaas** en **langelaaas**:

langelaas [lan'həla:s]... bw. (*ongewoon*) Meer dikw. *langerlaas*. Uiteindelik: ...

langelaaas: **op** (selde na en ten) **langelaas** (*lange laas*) of **op langerlaas** en selde **op langlaas** idiomatiese uitdr. Uiteindelik: ...

Volgens my waarneming word **op langelaas** frekwenter gebruik as **langelaaas**; die multileksikale lemma verteenwoordig dus die meer gebruiklike vorm. Tog word die uitspraakleiding en die woordsoort aanduiding alleen by die leksikale lemma aangetref. Uitspraakleiding kon net sowel by die multileksikale lemma verskyn, en die grammatikale inligting kan sodanig uitgebrei word dat daar naas "idiomatiese uitdr." ook bv. "bywoordelik" verskyn. Dikwels is uitspraakleiding baie nodig, veral indien daar geen leksikale lemma naas die multileksikale lemma verskyn nie. Vergelyk bv. die dubbelpuntlemma waar selfs moedertaalsprekers onseker is van die uitspraak.

laverent: in 'n **laverent** (ook *lawerent*) of in die **laverente** (ook *lawerente*) idiomatiese uitdr. (*streektaal; verouderend*) ...

Wat hier bo oor multileksikale lemmas gesê is, wat as selfstandige dubbelpuntlemmas optree, geld mutatis mutandis vir multileksikale lemmas wat binne die artikels van leksikale lemmas opgeneem is. Die verskil is dat hierdie uitdruk-

kings aan die einde van die betrokke artikel saam gegroeper word onder die opskrif UITDR. IDIOMATIES of GESPECIALISEERD. Hier veral is grammatikale leiding by die uitdrukking noodsaaklik, want die sintaktiese kategorie en funksie van die uitdrukking kan hemelsbreed verskil van dié van die betrokke lemma waarvan die woordsoort wel gegee word. Vergelyk die leksikale lemma *leef* met sy uitspraakleiding, woordsoortaanduiding, en byvoeglike en bywoordelike verbuigings:

leef [le:f] ww., het geleef; *byv.* en *byw.* lewende, lewend ...

Een van die idiomatiese uitdrukkings binne *leef* se artikel is:

leef: al wat leef en beef *Almal: Al wat leef en beef (is) daar om te sien wat die jag opgelewer het* (W.O. Kühne in *Huisg.*, 24 Jul. 1953, 11). *Al wat geleef en gebeef het, (het) elke middag na ete eers 'n bietjie gaan dut* (*Burg.*, 1 Mrt. 1954, 2).

Om agter te kom dat hierdie uitdrukking 'n naamwoordstuk is, en as onderwerp of voorwerp optree, moet die gebruiker die definiens en die voorbeeldmateriaal raadpleeg. Die sitate is ongelukkig so gekies dat albei slegs die onderwerpgebruik van die betrokke uitdrukking weergee.

Waar multileksikale lemmas nie afgeskeep word nie, is by die gebruik van etikette: sowel uitdrukkings binne die artikel van 'n leksikale lemma, as alleenstaande uitdrukkings in 'n eie artikel met 'n dubbelpuntsoekwoord word frekwent en korrek van etikette voorsien. Sien bv. die dubbelpuntlemma *laverent*: in 'n *laverent* hier bo met sy etikette *streektaal* en *verouderend*. Onder die leksikale lemma ¹*leeg* word die volgende etikette by die gespesialiseerde uitdrukkings aangetref:

boulbeurt: leë boulbeurt (*krieket*)

gewig: leë gewig (*ongewoon*)

versameling: leë versameling (*wiskunde*)

Wat mikrostruktureel veel beter is by die multileksikale as by die subleksikale lemmas is die aanbod van voorbeeldmateriaal by die uitdrukkings: hier word voldoende sitate verskaf sodat die gebruiker genoeg leiding kry oor die gebruik van die betrokke uitdrukking. Die gebruiker word deur middel van die sitate gelei hoe om die uitdrukking te gebruik, maar hy / sy sien ook waar dit gepas sou wees, want heelwat van die voorbeeldsinne dui ter staving van die etiket byvoorbeeld die gepaste stylvlak aan. Vergelyk byvoorbeeld die idiomatiese uitdrukkings onder lyn waar die poëme en sitate wel in die register van geselstaal hoort:

sny: 'n lyn sny (*geselstaal*) Haastig padgee of iewers heen gaan: *Ek sal nou-nou terug wees, ek sny net gou 'n lyn kafee toe.* ≈ "Laai op julle 'kit' en sny 'n lyn", is die manne toegevoeg (Paratus, Febr. 1982, 23).

wyn: alles in die lyn van wyn (*geselstaal*) Dit gaan goed: *As jy hom vra hoe dit gaan, sê hy altyd: "Alles in die lyn van wyn."* ≈ *Neels van Dam (het) kordaat gesê: "Ja-nee, dis nog alles in die lyn van wyn!"* (E. Roelofse: Kleigruis, 1970, 19).

Omdat uitdrukkings heel dikwels semanties ondeursigtig is, het die gebruiker 'n behoefte aan etimologiese inligting by sommige uitdrukkings. Ongelukkig word in deel IX geen etimologiese inligting by enige tipe uitdrukking verskaf nie. Volgens Gouws (1990: 272) kan etimologiese inligting met groot vrug in die artikels van vaste uitdrukkings verstrekkend word. Aangesien die WAT in die verlede al groot kritiek moes verduur oor hulle nie-aangawe of inkonsekwente aangawe van etimologiese inligting (vergelyk Gouws en Ponelis 1992: 97 en 98, Raidt 1992: 7, en Feinauer 1993: 291) kan die WAT dit gerus oorweeg om etimologiese inligting nie net te beperk tot leksikale lemmas nie, maar om dit tot bepaalde multileksikale lemmas uit te brei. Sodoende sal die WAT die eerste algemene eentalige Afrikaanse woordeboek wees wat uitgebreid etimologiese inligting by uitdrukkings verskaf. Uitdrukkings in deel IX wat hulle leen tot etimologiese toeligting is bv.: **die lakens uitdeel, onder die loep neem, op die lappe kom (bring, wees), op jou louere rus.**

Etimologiese inligting word weliswaar verskaf by uitdrukkings wat aan 'n ander taal ontleen is, maar wat eg Afrikaanse uitdrukkings verdring. Vergelyk:

strooi: die laaste strooi (strooitjie) of ook die laaste strooi (strooitjie) wat die kameel se rug breek (knak) (*n.d. Eng. the last straw of the last straw which breaks the camel's back*) ...

gee: iemand 'n lui(tjie) gee (*n.d. Eng. to give someone a ring*) ...

Die skrywer kon geen idiomatiese uitdrukkings met etimologiese toeligting opspoor anders as dié wat uit Engels afkomstig is nie. Omdat daar slegs by hierdie gevalle toeligting gegee word, kan die gebruiker dit weer eens as onnodige purisme en onvanpaste normering ervaar. Het die gebruiker nie veel eerder etimologiese inligting nodig by die uitdrukking wat die laaste strooi direk voorafgaan nie, naamlik sien: (die) laaste sien of (die) laaste sien van die blikkantien ...?

Wanneer daar eiename in veral gespesialiseerde uitdrukkings voorkom, word die etimologie ook verskaf soos in:

Lichtenstein: Lichtenstein se hartbees (*minder gebruiklik*) (na M.H.K. Lichtenstein, D. dierkundige en reisbeskrywer) Mofhartbees: ...

Malpighi: liggaampie van Malpighi (*na M. Malpighi, It. anatoom, histoloog en embrioloog*) Klein struktuur i.d. nier van 'n werweldier, ...

Leenwoordgroepe is nog 'n tipe multileksikale lemma wat bevoorreg is om etimologiese toeligting te kry. Hier word deurgaans die taal aangedui waaruit geleen is, bv.:

locus standi [lo'koesstan'di] leenwoordgroep. (*regswetenskap*) (*Lat.*) Reg van 'n party om in 'n hof te verskyn ...

laissez faire ... leenwoordgroep. Ook *laissez-faire*. (*Fr.*) ...

Let daarop dat daar by die leenwoordgroepe uitspraakleiding en etikette voorkom, maar die grammatikale leiding is maar yl in dié sin dat daar geen aanduiding is van sintaktiese kategorie of funksie nie. Deurgaans word slegs leenwoordgroep aangedui wat opsigself eerder by die etimologiese inligting tuis hoort. Die beste bewys dat WAT IX multileksikale lemmas steeds anders as leksikale lemmas hanteer, is dat in die artikel van *laissez-faire* as leksikale lemma, die etimologiese toeligting en die aanduiding dat dit 'n leenwoord is verkeerdelik verval, maar die sintaktiese kategorie nou wel aangetoon word, nl. s.nw.

laissez-faire ... s.nw. Sien LAISSEZ FAIRE

3.1.4 Afkortings

Nog 'n waardevolle makrostrukturele vernuwung in WAT IX is dat afkortings op hulle alfabetiese plek geplaas word, en nie soos in die vroeëre dele as eerste inskrywings by die betrokke alfabetletter nie. Volgens Gouws (1994) kan die vindbaarheid van sulke afkortinglemmas soms lastig wees, veral waar die uitspraak van die afkorting nie met sy alfabetiese plasing ooreenstem nie. Gouws noem hier die teken £ as afkorting van die Latyns *libra* ter aanduiding van die Britse pond as geldeenheid wat onder L gelemmatiseer is. Die skrywer stem saam met Gouws dat die voordele wat hierdie plasingstelsel vir die gebruiker inhou, veel groter is as die enkele nadele. Vergelyk byvoorbeeld die gebruiker wat hom tot deel V vir die begin van alfabetletter K moet wend indien hy / sy agterkom dat daar geen afkortings op hulle alfabetiese plek in deel VIII is nie.

Alhoewel die Opnamebeleid op pp. i-ii geen melding maak van die opname van afkortings nie, spreek leksikografiese dade sterker as toeliggende woorde: afkortings word as volwaardige lemmas met 'n volledige mikrostrukturele bewerking opgeneem. Eerstens word uitspraakleiding tereg slegs aangeref by dié afkortings wat in die spreektaal gebruik word. Hier is die klemaan-

duiding veral van belang. By afkortings wat slegs in die skryftaal voorkom, ontbreek die uitspraakleiding. Vergelyk:

Luk. afk. (*Bybelboek*) (*in verwysings*) Lukas: ...

LUK (*ook met punte L.U.K.*) [*elyka:'*] afk. LUK's. Lid van die Uitvoerende Komitee: ...

Etimologiese inligting word by sommige afkortings aangetref, sommiges het volledige sitate as voorbeeldmateriaal, en etikette word frekwent aangetref. Heelwat etikette handel oor die skryfwyse (met of sonder punte) van afkortings. Vergelyk:

lbf afk. (*gewigseenheid*) (*voorafgegaan deur 'n syfer*) *libra* (*Lat.*) + *force* (*Eng.*), pondgewig: 'n Onelastiese tou van lengte 4 vt kan net 'n gewig van 6 lbf dra (W.F. Beezhold: *Meganikaprob.*, 1969, 79).

l.c. afk. (*in bronverwysings*) Ook *loc. cit.* *loco citato* (*Lat.*), op die aangehaalde plek; *sin. t.a.p.*

By laasgenoemde afkorting word selfs 'n sinoniemverwysing gegee.

Mikrostruktureel word afkortings afgeskeep wat die aanduiding van woordsoorte betref. Alhoewel dit heelwat eenvoudiger is as by multileksikale lemmas om woordsoortstatus aan afkortings toe te ken, word daar feitlik deurgaans net aangetoon dat dit 'n afkorting is. Vergelyk al die aangehaalde voorbeelde hier bo. By enkele afkortings word daar egter wel 'n woordsoort aangedui, maar die gebruiker sal nie die rasionaal daarvoor kan bepaal nie. Vergelyk:

LSD (*ook soms met punte L.S.D.*) [*el'esde:'*] s.nw. 'n Kleurlose wit, in water oplosbare, kristallyne verbinding, ...; *sin. lisergiensuurdiëtielamied* (minder gebruiklik); **LSD-25** (ongewoon) ...:

LSD-25 (*ook soms met punte L.S.D.-25*) ... afk. (ongewoon) **LSD**: ...

LSH (*ook soms met punte L.S.H.*) [*el'esha:'*] afk. leerstafhoof.

L.S.O.D. (*ook sonder punte LSOD*) [*eleso:de:'*] s.nw. L.S.O.D.'s. Afk. vir *Laer Sekondêre Onderwysdiploma* (verouderd).

Let op die inkonsekwensie by twee agtereenvolgende lemmas, met **LSD** wat as selfstandige naamwoord aangedui word, maar die wisselvorm **LSD-25** direk daaronder word slegs as afkorting benoem. Hieruit sou die gebruiker dus kon aflei dat **LSD** wel 'n afkorting is. Van die daaropvolgende twee lemmas word alleen **L.S.O.D.** uitgesonder vir selfstandigenaamwoordstatus. Die bewerking van die twee afkortinglemmas wat woordsoortaanduidings het, verskil ook. In die artikel van **LSD** klink "afkorting" glad nie. Die volvorm word as 'n sinoniem verstrekk. By **L.S.O.D.** daarenteen, word "afkorting" as deel van die definiens saam met die volvorm aangebied.

Verdere dele van die WAT moet voortgaan met die huidige wyse van opname van afkortings, maar aandag sal gegee moet word aan die konsekwentheid en gedetailleerdheid van die mikrostrukturele bewerking van afkortings.

3.2 Mikrostrukturele inhoud

WAT IX is mikrostruktureel meer gevorderd as die voorafgaande dele: die lemmas word oor die algemeen beter bewerk, en die spektrum van taalkundige inligtingstipes per artikel is meer uitgebrei. Uitspraakleiding word by feitlik alle lemmas met behulp van fonetiese transkripsie verskaf, en by onverklaarde lemmas word die woordsoorte sowel as die hoofklem vir die eerste maal aangetoon. Volgens die Hoofredakteur (p. iii) word onverklaarde lemmas dus net soos gewone lemmas gehanteer om hulle makrostrukturele status binne die mikrostruktuur aan te dui. Verder word die leksikale betrekkinge van lemmas beter aangetoon deurdat daar heel dikwels nie net sinonimie aangetoon word nie, maar ook semantiese opposisie. Vergelyk onderstaande inskrywings:

laagaf [la:x'af] bw. Ook *laag af*. Ver na onder; ook, ver weg: teenoor *hoogop*:

lemoen [ləmun'] s.n.w., lemoene; lemoentjie. ...

- 1 Vrug v.d. lemoenboom ...
- 2 Lemoenboom ...
- 3 (*ongewoon*) Oranje: ...

lemoen'agtig, lemoenag'tig b.n.w., **lemoen'blaar** s.n.w., **lemoen'boer** s.n.w., **lemoen'boerdery** s.n.w. ...

'n Belangrike bate van die WAT IX is dat waar die *Afrikaanse Woordelys en Spelreëls* (1991) spel- en skryfopsies aan die gebruiker bied, bv. ten opsigte van die los- en vasskryf van woorde, hierdie opsies ook in die lemmas weergegee word. Sien **laagaf** hier bo. Die WAT IX dui selfs aan waar die huidige *Afrikaanse Woordelys en Spelreëls* afwyk van die skryfwyse van die 1964-uitgawe:

lt.genl. (ook *vroeër met 'n koppelteken* lt.-genl.) afk. (*gew. voor 'n eiennaam gebruik*) luitenant-generaal ...

lt.goew. (ook *vroeër met 'n koppelteken* lt.-goew.) afk. (*gew. voor 'n eiennaam gebruik*) (*histories*) luitenant-goewerneur ...

Dit val daarom op dat die 1964-spelling van **óf** ... **óf** in 'n sitaat by die lemma **lugsлаг** behoue bly. Die gebruiker word sodoende onder die indruk gebring dat dít die korrekte spelling is. Aangesien daar genoeg ander sitate binne die genoemde artikel voorkom, kon hierdie sitaat eenvoudig weggelaat word. Aan-

gesien die sitaat wel opgeneem is, moes hier minstens 'n opmerking wees wat die gebruiker attent maak op die verouderde spelling.

lugslag ... s.nw.

1 Slag waarby lug verplaas word: ... 'n *Hewige lugslag (het) gevolg wat veroorsaak het dat geboue in 'n omtrek van tot bykans twee myl, vanaf die punt direk onder die ontploffing of vernietig, of onbruikbaar was* (TNW, Des. 1964, 168). ...

Deur die jare het o.a. die volgende besware teen WAT-artikels geklink: die gebrek aan uitspraakleiding, te veel onnodige voorbeeldmateriaal of 'n swak keuse daarvan, foutiewe polisemiese versus homonimiese leiding, swak betekenisleiding, en rassistiese (deesdae veral ook seksistiese) definiense of voorbeeldmateriaal. Die res van die artikel sal ondersoek instel of WAT IX enigszins 'n verbetering op sy voorgangers is wat bogenoemde betref.

3.2.1 Uitspraakleiding

Dat WAT IX hom wel deeglik aan kritiek (Gouws 1985: 15, Vos 1992: 86, Feinauer 1993: 287) ten opsigte van uitspraakleiding steur, blyk duidelik daaruit dat hoofklemaanduidings nou ook soos reeds gemeld, deurgaans by onverklaarde meersillabige lemmas gegee word. Verder word fonetiese transkripsie in deel IX by die meeste lemmas, en nie slegs by uitgesoektes soos in voorafgaande dele nie, aangetref. Die transkripsie self word taalkundig korrek tussen vierkantige hakies, en nie meer tussen ronde hakies nie, geplaas.

Die WAT IX is die eerste Afrikaanse woordeboek wat bykans elke verklaarde lemma van 'n fonetiese transkripsie en hoofklemaanduiding voorsien, en ook Otto (1994) en Gouws (1994) maak melding van die positiewe aspek hiervan. By enkele lemmas neem die uitspraakleiding die gebruiker in ag deur hoegenaamd nie voorskriftelik te transkribeer nie, maar die meer korrekte en die lekeuitspraak naas mekaar sonder enige etiket te plaas. Vergelyk hier byvoorbeeld die onderskeie transkripsies van **lamé** en **logo**:

lamé [lame:' / laməi'], **logo** [lo:'xo: / læu'gœu]

Dit is dan tog vreemd dat by **lycra** slegs een transkripsie verskaf word, naamlik [lai'kra], terwyl [li'kra] veel meer gebruik word. Vir alle samestellings en afleidings van **Luther** word deurgaans ook net een uitspraak verskaf, naamlik [ly'tər], terwyl heelwat sprekers dit as [lu'tər] uitspreek. By die lemma **landbou-ekonomie** word daar tereg vir elke variasie-uitspraak van die "e" in **ekonomie** voorsiening gemaak, en by **lamelproduk** hou die transkripsie rekening met die vokaalreduksie wat by **produk** voorkom:

landbou-ekonomie [lant'boeui:kənumi / lant'boeue:kənumi / lant'boeue:kənumi],
lamelprodukt [lamel'pruðək]

Tot hier toe dan die positiewe uitspraak oor die gehalte van die fonetiese transkripsie in WAT IX. Gouws (1994) se siening dat die uitspraakleiding hier die inskrywingstipe is wat op taalkundige grond die skerpste gekritiseer kan word, word heelhartig ondersteun. Die kritiek handel nie oor enkele geïsoleerde gevalle waar 'n toevallige transkripsiefout gemaak is nie, bv. **langwerkend** wat in die transkripsie 'n slot-[d] kry en **ligabsorpsie** en **ligabsorberend** waarvan die "b" as 'n [b] getranskribeer word nie (Gouws 1994). Die probleem is heelwat dieperliggend deurdat die Redaksie hulle nie steur aan transkriberingsvoor-skrifte nie.

Die eerste ernstige transkriberingsfout wat voorkom, is dat lemmas volgens hulle letteruitspraak getranskribeer word. So word die eerste "r" van die lemma **lugverkeer** bv. getranskribeer: [lœxfær:er]. Sulke uitspraakleiding is vir die gebruiker van geen waarde nie. Talle sulke gevalle is opgeteken waarvan enkeles genoem word. Die lemma **letteruitspraak** is toevallig nie met twee [t]'s langs mekaar getranskribeer nie, maar vergelyk die volgende:

langsskeeps [lɑŋs'ske:ps], **langssy** [lɑŋs'søi], **langsterttinktinkie** [lɑŋ'ste:rt-
 tɔŋtɔŋki], **landdag** [lant'dax], **landtong** [lant'tɔŋ], **landtogganger** [lant'-
 tɔxxɑŋhər], **landdier** [lant'di:r], **laaggradig** [la:xxra:dəx / la:x'ra:dəx],
laedrukkoker [la:'hədrekko:kər], **lurexstof** [ly'reksstɔf].

Al bogenoemde voorbeelde is 'n samestelling of samestellende afleiding van die een of ander aard. Terwyl **laggas** as samestelling onjuis met twee [x]'s getranskribeer word, word **lagge** as simpleks korrek weergegee. Sodoende word die gebruiker net meer mislei, want by hierdie gevalle is die uitspraak dan nie die bepalende faktor vir die transkripsie nie, maar wel die morfologie.

Volgens Gouws (1994) misken die fonetiese transkripsies wat in WAT IX aangetref word bepaalde klankpatrone en -prosesse van Afrikaans. Assimilasieprosesse waaronder stemassimilasie en nasaalassimilasie word gewoon geïgnoreer wanneer die Redaksie alleen een uitspraakvariant, naamlik die letteruitspraak aanbied. Voorbeelde waar stemassimilasie misken word, sluit in **lid-doring** [lɔt'do:rɔŋ], **liefde** [lif də] en **loopbaan** [lo:p'ba:n]. Vir nasaalassimilasie wat ontbreek, vergelyk onder andere: **lemoenkleur** [ləmun'klør], **loonkoste** [lo:n'kɔstə], **lyngooi** [lɔin'xo:i], **langbeksteenbras** [lɑŋ'bəkste:nbras], en **ligaanpassing** [lɔx'a:npasɔŋ]. Vokaalreduksie word soms aangedui soos in **lamelprodukt** wat reeds genoem is en **liederkomponis** [li:'dærkɔmpɔnəs / li:'dærkɔmpunəs]. Dit ontbreek ongelukkig te dikwels om die ware taalgebruik van moedertaalsprekers weer te gee. Vergelyk byvoorbeeld **lojaal** [lo:'ja:l], **liefdesgodin** [lif dəsxo:dən], **logistiek** [lɔxəstik'], **lekeapostolaat** [le:kəapɔstola:t], en **lektō-**

raat [lektora:t']. Otto (1994) merk ook op dat vokaalreduksie inkonsekwent aangedui word: by **Laagportugees** wel, maar nie by **laaglandgorilla** nie.

Waar die Redaksie hoegenaamd nie daarin slaag om die werklike uitspraak weer te gee nie, is by lemmas waar glyerinvloeging ter sprake kom. Die vernaamste probleem hier is dat daar selde variasie aangetoon word. Volgens Gouws (1985: 17) moet daar by uitspraakleiding in die vorm van fonetiese transkripsie juis volledige variasie aangetoon word. Die belangrikste fout wat begaan word, is dat daar slegs 'n [h] as glyer ingevoeg word, en dat daar dan nie 'n variant voorsien word waar die [h] deur 'n [j] of [w] of bloot 'n glottisslag vervang word nie. In Afrikaans word die [h] dikwels 'n [j] na die vokale [i], [æi], [o:i] of [œy], en 'n [w] na die vokale [u] en [œu]. Na die vokale [i] en [a:] word gewoonlik 'n [h] alleen ingevoeg, indien die spreker wel iets invoeg. Vergelyk onderstaande voorbeelde uit die WAT waar die [h] ook met 'n [j] vervang kon word:

laboratorium [labəra:to:'rihoem], **liniaal** [ləniha:l' / liniha:l'], **linieer** [lənihe:r' / linihe:r'], **looier** [lo:i'hər], **leier** [ləi'hər], **lisensieer** [ləsensie:r'], **luier** [lœyhər]

Hierdie fout word verswaar daardeur dat die WAT by die lemma **laagrelief** wel die glyer [j] heeltemal korrek as variant aantoon: [la:x'relihef / la:x'relijef]. Die WAT-redaksie is dus bewus van die variasie: hulle toon dit net selde aan.

In die volgende WAT-voorbeelde kan die [h] met 'n [w] vervang word:

landauer [landau'hər / lan'dauhər], **linguaal** [ləŋxuha:l'], **landbouer** [lant'bœuhər], **louere** [lœu'hərə], **Ligbloue** [ləx'blœuhə]

Oordrewe glyeruitspraak word by sommige lemmas deur die fonetiese transkripsies aangetoon. Hier word die gebruiker dus ook kommunikatief verkeerd gelei. Otto (1994) wys op die ongevarieerde transkripsie van **lae** en al die samestellings en afleidings waarin **lae** optree. Nêrens word aangetoon dat sommige sprekers **lae** as eensillabig uitspreek nie, want dit word deurgaans as [la:'hə] getranskribeer. Dieselfde kritiek geld ten opsigte van **leë** [le:'hə]. By die deursneespreker word daar geen definitiewe glyer gehoor by die uitspraak van die woord **honger** nie. Vergelyk egter die misleidende transkripsies van **landhonger** [lant'høŋhər] en **lughonger** [lœx'høŋhər] in die WAT IX.

Wat uiters teleurstellend is, is dat WAT IX steeds nie sillabeverdelings saam met hoofklemaanduidings en fonetiese transkripsies gee nie. Hier steur die WAT hom kennelik nie aan vroeëre kritiek nie. Vergelyk o.a. Feinauer (1993: 288) en Gouws (1989: 254). Die weglaat van die derde tipe uitspraakleiding maak die fonetiese transkripsie van onderstaande lemmas nutteloos:

LUK [ɛlyka:], **L.S.O.D.** [ɛleso:de:]

By bogenoemde moet die leksikograaf hom op die taalintuïsie van die gebruiker beroep om die uitspraakleiding enigsins te laat slaag. Dit is vanselfsprekend leksikografies totaal onaanvaarbaar. Nie net by afkortings sou sillabe-verdelings groter uitspraakeendigheid lewer nie. Oningeligte gebruikers sou nie by die volgende lemmas die uitspraakverskil uit die fonetiese transkripsie en die hoofklemaanduiding kon aflei nie. Die derde tipe uitspraak-inligting, naamlik sillabe-verdeling is hier onontbeerlik:

¹leweraar [le:'vəra:r] s.nw. ... Leweransier:

²leweraar [le:'vəra:r] s.nw. Enigeen v.d. are wat bloed wat i.d. lewer gesuiwer is, n.d. onderste holaar voer: ...

Verdere lemmas waar die sillabe-verdeling uitspraak-hulp sou verleen, is byvoorbeeld by lammeroes [la'mərus], lammerooi [la'məro:i] en lusernerd-vlooi [luse:r'ənertflo:i / ...]. Saam met Gouws (1994) kan slegs gehoop word dat 'n volgende deel van die WAT ook sillabe-verdeling bevat. Die Redaksie sal in verdere dele ook indringend aandag moet gee aan die fonetiese transkripsies sodat die natuurlike Afrikaanse uitspraak so getrou moontlik weergegee word.

3.2.2 Voorbeeldmateriaal

Die WAT IX sorg vir 'n uiters volledige en bruikbare versameling voorbeeldmateriaal. Dit is opmerklik dat hier heelwat meer van poëme gebruik gemaak word as in vorige dele wat die gevolg het dat hier nie soos in vorige dele, lemmas is wat geen voorbeeldmateriaal het nie. Voorbeeldmateriaal moet egter te alle tye funksioneel wees sodat die gebruiker optimale nut uit elke woordeboekartikel kry.

Volgens Gouws (1991: 275) moet daar kohesie binne 'n woordeboekartikel bestaan: die afsonderlike inskrywings moet mekaar dus op die een of ander manier aanvul. Voorbeeldmateriaal is een tipe inskrywing wat hierdie kohesie kan bewerkstellig deur byvoorbeeld definiense, etikette of grammatikale toeligting aan te vul. Die WAT IX illustreer dit uitstekend deur die sitate in die lang artikel van liefde heg te laat aansluit by die grammatikale toeligting wat heel aan die begin gegee word, en wat herhaal word by die betrokke polisemiese onderskeidings

liefde ... s.nw. (met mv. in bet.1b I en ii, en 4)

1 ...

b I ... *Hy kyk op (na haar) met oë waarin nuwe belangstellings ou liefdes verdring* (E. Leroux: Silbersteins, 1962, 78).

ii mv. *liefdes*. ... *Vergeefse liefdes en ongekonsumeerde verhoudings* (H. Aucamp: Enkelvlug, 1978, 27).

4 mv. *liefdes; vkw. liefdetjie*. ... *Dit (is) miskien logies dat my vrou een*

van sy ou liefdes was (J. Meintjes: Gister, 1963, 145). *Sy gedagtes vér daarvandaan by sy eerste blonde liefdetjie* (A.P. Brink: Rapport, 1976, 150).

WAT IX wyk met bogenoemde weldeurdagte sitate ook af van die vroeëre beleid dat slegs die lemma sonder meer in die voorbeeldmateriaal vergestalt word. Die pragmatiese funksie van die voorbeeldmateriaal kom hier sterk op die voorgrond deurdat die gebruiker 'n praktiese toepassing van die grammatikale toeligting sien. Ongelukkig word hierdie goeie gebruik van voorbeeldmateriaal nie konsekwent genoeg toegepas nie. Vergelyk die sitate in die artikel van ²lui II 2 b:

II s.nw., lui; luitjie. (*t.o.v. 'n klok of soortgelyke apparaat*)

1 ...

2 a ...

b (*dikw. i.d. vkw.*) Telefoonoproep: *Hy staan op om ... die lui te beantwoord* (TL, Nov. 1969, 30). *Nou goed dan, Lenia, en dankie vir die lui* (H. Aucamp in Aucamp - Bakkes: Baksel, 1973, 24).

Nie een van die twee sitate toon aansluiting by die etiket, naamlik dat hierdie vorm dikwels as verkleinwoord optree nie. Hier word geen kohesie tussen etiket en voorbeeldmateriaal bewerkstellig nie. Dit het tot gevolg dat die leksikografiese waarde van die artikel laag is, en die gebruiker word benadeel deurdat hy nie die genoemde frekwente optrede van luitjie geïllustreer sien nie.

Nog 'n manier waarop die aanbod van voorbeeldmateriaal tot hegte kohesie binne die woordeboekartikel kan lei, is dat die sitate chronologies aangebied kan word met bv. die oudste optekening eerste en die modernstes laaste. Indien daar enige etikette aanwesig is wat op gebruiksfrekwensie dui, toon die sitate aansluiting daarby deur die tydvak aan te toon waarbinne die betrokke lemma aangetref is. Reeds in (1993: 293) wys Feinauer daarop dat daar in WAT VIII nêrens genoem word volgens watter maatstaf die sitate georden word nie. Die ouderdom en register van lemmas kan steeds in WAT IX nie nagegaan word aan die hand van die sitate nie. Dit lyk steeds of sitate lukraak opgeneem en gerangskik word. Afgesien van die kohesie en historiese inligting wat binne die artikel verlore gaan, gebeur dit dat sitate dikwels uit 'n baie beperkte tydvak geneem is. Volgens Gouws (1994) is hierdie jaartalverspreidings dan nie verteenwoordigend van die betrokke lemmas se werklike gebruik nie.

Die sitate wat by die lemma **lamlendig** aangetref word, illustreer presies bogenoemde probleem. Onder die eerste polisemiese inskrywing **1a** van **lamlendig** volg die jaartalle van die sitate soos volg op mekaar: 1964, 1959, 1975, 1971, 1953, 1940, 1929, 1949, 1962. Hierdie opname toon wel verspreiding wat die tydvak betref, maar die rangskikking van die sitate is heeltemal lukraak, en die modernste optekening van **lamlendig** is in 1975, terwyl hierdie lemma nog aktief is in sowel die skryf- as spreektaal. Die laatste sitaat in die hele artikel

van **lamlendig** is uit 1979. Hierdie keuse van sitate weerspieël dus nie die werklike gebruik van die woord **lamlendig** nie.

Daar is reeds gesê dat *WAT IX* makrostruktureel erkenning probeer gee aan die Suid-Afrikaanse werklikheid. Dit lei tot die opname van vorme wat frekwent in Kaapse Afrikaans aangetref word, soos **lamoen** en **lekka**. Die mikrostruktuur in die vorm van voorbeeldmateriaal wat by hierdie lemmas aangetref word, maak hierdie poging egter ongedaan, want dit weerspieël nie die sosiale realiteit nie. Vergelyk:

lamoen [lamun'] s.n.w. lamoene; lamoentjie. (*geselstaal; verouderd*) Sien LEMOEN: *Daar (word) lamoene en nartjies verkoop* (Brandw., 15 Des. 1912, 418).

lekka ... b.n.w. (*geselstaal; ongewoon*) Sien ¹LEKKER 1: *Man, was daai tamaties vir jou lekka! Lekka, lekka ywe* (I.D. du Plessis: Fez², 1944, 27).
Ook byw.: *Sy kan treurig sing en lekka sing* (Transv., 31 Jan. 1966, 2).

By **lamoen** is daar geen sitaat wat die Kaapse *geselstaal* illustreer nie; alleen 'n sitaat uit 'n ou bron wat terugwys na die etiket *verouderd*. Dit terwyl **lamoen** frekwent gehoor word. By **lekka** word die *geselstaal* wel heeltemal tereg in die poëem weergegee, maar die twee sitate uit ou bronne dui daarop dat **lekka** ook 'n verouderde vorm is, terwyl dit net so min ongewoon of verouderd soos **lamoen** is.

Hierdie swak gekose en min voorbeeldmateriaal by bogenoemde en ander soortgelyke lemmas dui weer eens op die tekortkominge in die databasis. Die databasis bevat te min spreektaalvorme om sinvolle voorbeeldmateriaal by *geselstaal* te lewer. Waar die redaksie hierdie probleem oorkom by ander frekwent gebruikte spreektaalvorme soos **laken**, **lekker**, **lepel**, **loop**, ens. deur wyd van poëme gebruik te maak, moet poëme konsekwent ook by frekwent gebruikte Kaapse vorme verskaf word. Gouws (1994) wys daarop dat waar 'n woordeboek oor 'n databasis beskik, veral sitate as voorbeeldmateriaal verstrekkend word. Dat daar by gewone woorde uit die spreek- en skryftaal, soos hier bo genoem, hoofsaaklik poëme en relatief min sitate gegee word, dui vir hom daarop dat algemene woorde nie so goed in die databasis verteenwoordig is nie. Dit verlaag op sy beurt die waarde van die *WAT* as 'n aanduiding van gebruiksfrekwensie.

Volgens Lombard (1992: 150) moet daar in woordeboeke baie meer klem geplaas word op die werklike en natuurlike taalgebruik. Rademeyer (1992: 188) merk op dat daar in die *WAT* nie genoegsaam uit die alledaagse skryf- en spreektaal gesitater word nie. Sy merk ook op dat praktiese leiding oor die alledaagse taal vir die woordeboekgebruiker van veel groter waarde sal wees as die talle literêre voorbeelde vol eienaardighede. In *WAT IX* geld hierdie aantyging van Rademeyer steeds deurdat daar oneweredige leksikale bewerking van lemmas voorkom: die leksikale lading van sommige gewone lemmas is heelwat

ligter as dié van ongewone lemmas. Sodoende word die gebruiker weer eens mislei t.o.v. die frekwensie van lemmas: die afwykende kry meer aandag wat groter frekwensie impliseer, terwyl die afskeep van tipiese lemmas ook minder frekwensie impliseer. Vergelyk byvoorbeeld die heelwat langer inskrywings van vreemde woorde soos *ladinoklawer*, *landauer*, *lemma*, *liewerlee*, *linguaal*, en *luglokasie* teenoor die inskrywings van gewone woorde soos *ladybird*, *landaf*, *leuning*, *liewers*, *liniaal*, en *lugmag*.

Hierdie probleem spruit uit die Prerenaisance-metode van ekserpering deur medewerkers, wat steeds in die WAT gevolg word: vreemde woorde word juis opgeteken om te bewys dat hulle bestaan, en heelwat van hierdie optekeninge word as voorbeeldmateriaal geplaas. Lombard (1992: 154) wys daarop dat ekserpeerders eerder die ongewone as die gebruiklike opteken. Hy haal ook Landau (1984: 80) in dié verband aan: "... citation readers all too often ignore common usages and give disproportionate attention to uncommon ones". Die redaksie van die WAT moet bedag wees hierop, en nie in verdere dele voortgaan met hierdie misleidende beeld t.o.v. die frekwensie van ongewone lemmas nie, of die WAT moet dit in sy toeligting baie duidelik stel dat sy sinate nie as absoluut verteenwoordigend van die taal beskou word nie, veral nie wat die frekwensie van woordgebruik betref nie.

In WAT IX is hier steeds te veel artikels waar voorbeeldmateriaal nie deurdink of gekeur word nje. Dit lei tot onnodige aanhalings of poëme wat op die ou end niksseggend is; m.a.w. dit dra niks nuuts by tot die gebruik of verstaan van die betrokke lemma nie. Vergelyk die ongewone lemma *luilappie* se voorbeeldmateriaal waar dieselfde inligting keer op keer herhaal word, bv. dat dit saam met die werkwoord *bak* optree, en dat dit 'n klein tipe pannekoekie is. Een sitaat en een poëem sou voldoende wees.

luilappie ... s.nw., *luilappies*. Selde ook *leilappie*. Klein pannekoekie: ook soms, plat vetkoekie of ander plat koekie in vet gebak: *Gaan ons plaatkoekies bak, of luilappies? As jy te lui is om groot pannekoekie te maak, dan bak jy sommer gou-gou luilappies* (Robertson). *Die susters gaan hierdie keer nie pannekoek nie, maar luilappies op die bazaar bedien* (Krugersdorp). *'n Luilappie is 'n klein pannekoekie. Gewoonlik word drie of vier tegelyk in 'n pan gebak* (Oos-Londen). *Dit is "luilappies" genoem omdat dit so gou en maklik is om te maak. Dit is 'n soort pannekoekie. Die resep kan effens gewysig word en stywer soos vetkoek gebak word* (Burg., 16 Jul. 1958, 5 — brief uit Clanwilliam).

Vergelyk ook die voorbeeldmateriaal by:

²lui [lœy]

I ww., het gelui; *byv.* en *byw.* luidende of soms luidende, ...

1 (t.o.v. 'n klok of soortgelyke apparaat)

a 'n Gew. helder en weerklinkende of weergalmende geluid voort-

bring: *Die kerkklokke lui galmend. Die roomyskar se klokkie lui klingelend. Deurklokkies lui rinkelend of gonsend. Die elektriese alarm lui skril. "Tieng-tong" lui die klokkie toe ons die winkel binnestap. Het die telefoon gelui terwyl ek weg was? ≈ Ons hoor die Paasfeesklokke / wat oor die wêreld lui* (Ges. 147: 3; 1983). *Die aandklokke lui in die verte / In tone wat saggies kom streel* (F.A.K.-sangb., 1961, 400). *'n Klokboei klink / oor die see, skommel en lui* (D.J. Opperman: Jorik, 1949, 60). *Hy* (die fietsklokkie) *lui nie soos 'n skoolklok en 'n voordeurklok en 'n gewone fiets se klok nie* (F. v. Zyl in Louw - Schumann: Afr. Kersboek, 1963, 135). *Die telefoon het gister feitlik onophoudelik by die pastorie gelui* (Volksbl., 3 Mrt. 1989, 1). *Jy hoor die ongewone geluid van luiende klokke* (E. Baker in Stad, 1977, 49).

Hierdie voorbeeldmateriaal is heelwat meer deurdag gekies as dié van luilapie. Verskillende tipes klokke word by die poëme aangetref. Die alarm en telefoon verteenwoordig die "soortgelyke apparaat" wat in die etiket genoem word. Wat die sitate betref, 'n verledetydsvorm word aangetref sowel as 'n illustrasie van die byvoeglike gebruik, nl. luiend. Die Redaksie kon egter die sitate uit die Gesangboek, die sangbundel, en uit Opperman gerus gelaat het. Dit verswaar die artikel onnodig, en is van weinig nut vir die gebruiker aangesien dit nie net verhewe literêre taal bevat nie, maar ook niks nuuts aangaande die lemma illustreer nie.

Heelwat aandag sal aan die opname en ordening van voorbeeldmateriaal in komende dele van die WAT geskenk moet word.

3.2.3 Rassisme en Seksisme

Dat dit vir die WAT erns is om rassistiese en seksistiese leksikale items so sensitief moontlik te hanteer, blyk uit die internasionale kongres oor die hantering van beledigende en sensitiewe leksikale items in die WAT wat die Buro van die WAT vanaf 8 - 10 Februarie 1994 op Stellenbosch aangebied het. In 'n beleidstuk wat na die kongres verskyn (Beleidstuk vir die hantering van beledigende en sensitiewe leksikale items in die *Woordeboek van die Afrikaanse Taal* 1994: 3), word gesê dat onder andere uitdrukkings waarin 'n rassistiese komponent voorkom, slegs in die elektroniese manuskrip gehanteer sal word. "Die uitdrukking *'n los hotnot* sal byvoorbeeld in die artikel van die lemma *hotnot* ... tot elektroniese manuskrip verwerk word."

In WAT IX is *loshotnot* egter as gewone leksikale lemma opgeneem beide as adjektief en as naamwoord. Daarmee saam verskyn die volgende etiket (*oorspr. wsk. rassisties, en tans mntl. deur sommige so aangevoel; geselstaal*). Alhoewel die etiket dalk ietwat sterker gestel kon word, is daar in die definiensgedeeltes en in die voorbeeldmateriaal geen rassistiese komponente nie. Die WAT het bepaald ver gevorder indien hierdie lemma vergelyk word met die gedeelte

in WAT IV waar 'n **Los Hotnot** inderdaad as uitdrukking binne die artikel van die rassitiese term aangetref word. In deel IV is die enigste etiket wat naas die lemma gegee word, (*geselstaal*). Die WAT moet gelukkigewens word met die verandering in perspektief en gesindheid wat duidelik blyk nie net uit bogenoemde etikettering nie, maar ook uit onder andere die etikettering by **lokasie** (*veral as toepassing v.d. apartheidsbeleid as diskriminerend en neerhalend aangevoel*), **losvolk** (*tans dikw. as neerhalend aangevoel; verouderend*), en **luglokasie** (*veral as toepassing v.d. apartheidsbeleid as diskriminerend en neerhalend aangevoel*). Ook nie in die definiense of voorbeeldmateriaal van bogenoemde lemmas is materiaal aangetref wat aanstoot sal gee nie. Een lemma is wel aangetref waar die definiens 'n onnodige verwysing na ras bevat, naamlik **lanie**:

lanie [la:ni] s.nw. lanies. (*geselstaal, ook neerhalend*) Ook soms *larnie*. Man of seun, veral blanke man of seun, dikw. in 'n gesagsposisie: ...

Wat wel uit die etiket by **lanie** blyk, is dat die WAT IX konsekwent etiketteer, en nie net t.o.v. lemmas wat op gekleurde persone betrekking het nie. Neerhalende benamings geld ook nie net vir verskillende rassegroepe nie, maar ook vir verskillende volkere. Ook hier word etikette ingespan:

lunsriem ... s.nw. .

1 ...

2 a ...

b ...

c ... (*dikw. met 'n hl. Lunsriem*) (*neerhalend; verouderd*) Engelsman: ...

Die WAT is ook bedag op sensitiwiteit rondom geslag, want soos Lombard (1992: 152) dit stel: "die tydsgewrig waarin die gebruikers hulle bevind, (moet) in ag geneem word." Volgens die Beleidstuk (1994: 4) word dit gestel dat geslag nie ophefbaar is nie. Die vroulike vorms in geslagsopposisiepare word dus wel gelemmatiseer en volledig bewerk in WAT IX. Die Buro wil egter 'n rol speel in die gelykberegting van geslagte. Die artikels van die volgende geslagsopposisiepare is nagegaan om te sien of daar wel gelykberegting plaasvind. Vergelyk eerstens die puik hantering van die geslagsopposisiepare **lektor** en **lektrise**:

lektor ... s.nw., *lektore of selde lektors; lektortjie*. Manlike of vroulike persoon wat as dosent aan 'n tersiêre onderwysinrigting, veral 'n universiteit, verbonde is en wat nie die status van 'n professor het nie; ook soms, manlike of vroulike persoon wat ander gevorderde of gespesialiseerde onderrig gee: ... Vgl. LEKTRISE.

lektrise ... s.nw., *lektrises; lektrisetjie*. Vroulike persoon wat as dosent aan 'n tersiêre onderwysinrigting, veral 'n universiteit, verbonde is en

wat nie die status van 'n professor het nie; ook soms, vroulike persoon wat ander gevorderde of gespesialiseerde onderrig gee: ... Vgl. LEKTOR.

WAT IX hou tred met die tydsgegewig waarin die gebruikers hulle bevind deur *lektor* as die neutrale variant in die definieëns aan te toon: heelwat vroue verkies deesdae om woorde soos *lektrise* te vermy wat vroulik geslagtelik gemerk is. Die *WAT* sou die neutraliteit verder kon verhoog deur in plaas van "Manlike of vroulike persoon" bloot "persoon" te gebruik. Verder word die vroulike vorm nie gedefinieer deur blote verwysing na die neutrale vorm nie, maar *lektrise* word selfstandig gedefinieer. Dit sal Beylefeld en Van Jaarsveld (1994: 46) tevrede stel wat beswaar maak teen die teenstellingsdefinieëns (*Vroulike vorm van ...*) wat gewoonlik in Afrikaanse verklarende woordeboeke vir vroulike vorme gebruik word. "Die woordeboekgebruiker word telkens vooruit- of terugverwys na die manlike vorm van betekenisverklaring." Wat die voorbeeldsinne betref, tree die *WAT IX* ook onberispelik op deur die neutraliteit van die term te illustreer met voorbeelde waar *lektor* sowel manlike as vroulike referente het. Met die kruisverwysing by sowel *lektor* na *lektrise* en by *lektrise* na *lektor* lewer die *WAT IX* baanbrekerswerk wat geslagtelikheid betref. Volgens Beyer (1995: 25) het die *HAT* (1994) in ongeveer 150 geslagsopposisiepare net een maal 'n kruisverwysing vanaf die neutrale na die vroulike vorm, naamlik by *sekretaris* en *sekretaresse*. By die res is daar net eenrigtingverwysing vanaf die vroulike na die neutrale vorm. Met die kruisverwysings tussen die neutrale en vroulike vorme dui die *WAT* hier ook die semantiese verhouding van opposisie aan. In par. 3.2 is reeds melding gemaak van hierdie wins in *WAT IX*. Volgens Gouws (1989: 158) "(is) die aanduiding van leksikale items wat in opposisie tot 'n bepaalde lemma staan, nie 'n verpligte woordeboek-inskrywingstipe nie. Dit kan eerder as 'n leksikografiese bonus beskou word".

By die geslagsopposisiepare *lyer* en *lyeres* is die bewerking ook bevredigend:

lyer ... s.nw., *lyers*; *lyertjie*. ... Iem. wat ly of aan iets ly: ...

lyeres ... s.nw. *lyeresse*; *lyeresse*. ... Vroulike persoon wat ly of aan iets ly: ...

Die neutraliteit van *lyer* word onmiddellik bevestig deur die gebruik van "iem.", alhoewel "persoon" groter konsekwensie met die direk daaropvolgende *lyeres* en bogenoemde geslagsopposisiepaar sou bewerkstellig. Die voorbeeldmateriaal by *lyer* dek ook vroulike *lyers*, en *lyeres* het 'n volledige definieëns. Die enigste beswaar is dat hier geen kruisverwysing van watter aard ook al tussen *lyer* en *lyeres* is nie. Weer eens is hier dus inkonsekwensie met *lektor* en *lektrise* as maatstaf.

Enkele artikels is wel gevind waar die *WAT IX* die manlike geslag onnodig begunstig:

¹leidster ... s.nw., leidsters; leidstertjie. Vroulike persoon wat lei: ... Vgl. LEIER.

leier ... s.nw., leiers; leiertjie.

1 a lem., dikw. 'n man, wat lei, of iets wat lei: ...

Vgl. ¹LEIDSTER.

Die gedeelte "dikw. 'n man" hoort nie in die definieus van **leier** nie, en dit versterk Beyliefeld en Van Jaarsveld (1994) se standpunt van geslagsongelykheid in woordeboeke. Afgesien daarvan word die gebruiker mislei, want hierdie inligting is onjuis; **leidster** word vandag selde gebruik. Die mees onlangse sitaat in die artikel van **leidster** is bv. uit 1972. Die Redaksie het by die keuse van die sitate ook 'n oordeelsfout begaan: in die enigste sitaat waar 'n vroulike leier figureer, tree die verkleiningsvorm van **leier** op, naamlik: "Sy het 'n baie sterk persoonlikheid, sy's 'n gebore leiertjie en die ander is min of meer haar slawe (E. Steenberg in Sarie, 12 Aug. 1970, 83)." Ander opmerkings oor bogenoemde: "iem." moet verkieslik deur "persoon" vervang word in die definieus van **leier**, en hier is weer kruisverwysings tussen **leidster** en **leier**.

Ook by die lemmas **leierdiaken** en **leierouderling** het die WAT IX nie rekening gehou met die veranderende maatskappy nie: vroue kan reeds 'n geruime tyd binne sommige gereformeerde kerke as diakens en ouderlinge diens doen. Dit word egter nie in die definieus van genoemde lemmas weer- spieël nie:

leierdiaken ... s.nw. Diaken deur sy kollegas gekies om namens hulle op te tree en sake waarby hulle as diakens betrokke is, te reël: ...

leierouderling ... s.nw. Ouderling deur sy kollegas gekies om namens hulle op te tree en sake waarby hulle as ouderlinge betrokke is, te reël: ...

Deur gewoon "sy" by die onderskeie definieus te skrap, kan geslagsneutraliteit aangetoon word. In verdere dele van die WAT moet definieus en voorbeeldmateriaal goed deurdink word sodat die Redaksie nie van geslagsvoorkeur beskuldig word nie. Ter wille van beter leksikografie moet geslagsopposisie- pare ook konsekwent bewerk word. Beyer (1995) kan byvoorbeeld geraadpleeg word vir die leksikografiese hantering van morfologies gemerkte geslagsopposisiepare.

3.2.4 Polisemie en homonimie

Wat betekenisverklaring as sodanig betref, word volstaan met 'n bespreking van die hantering van polisemie en homonimie in WAT IX: dit is een aspek wat

reeds heelwat kritiek ontlok het. Vergelyk bv. Gouws (1985), Gouws (1985a), Gouws (1986), en Feinauer (1993).

Wat die makrostrukturele ordening van die lemmas as homonieme of poliseme betref, word hier goeie leksikografie in WAT IX aangetref. Vergelyk byvoorbeeld die lemmas **lemmetjie** en **lensie**, waar die verkleiningsvorme van **lem**, en **lens** as homonieme onderskei word van die sitrusvrug en peulsaad. Soortgelyke homonimiese ordenings word aangetref waar meervoudsvorme as homonieme onderskei word, bv. **lappies** en **lansiers**. Die meervoudsvorm van **liggie** word egter nie as homoniem aangebied van **liggies** (besonder **lig**) en **liggies** (liggaamsoefening) nie. By die lemma **logger** word die vergrotende trap van **log** ook heeltemal tereg as homoniem van die kusvaartuig aangedui. So word **lewenswyse** as adj. ook homonimies onderskei van die naamwoord **lewenswyse**. Goeie leksikografie word ook bedryf deurdat die algemeenste lemma eerste geplaas word. By **lemmetjie** lyk die homonimiese plasing soos volg:

¹lemmetjie ... 1 Vkw. van *lem*.

²lemmetjie ... 1a Vrug van die lemmetjieboom ...

³lemmetjie ... Lemmie.

Die enigste afwyking is by **lansiers** waar die dansvorm voor die meervoudsvorm geplaas is: die meervoudsvorm is algemener bekend. Hierdie tipe homoniemaanduiding word soms ietwat te ver gevoer wanneer bv. **laslappies** as twee homonieme opgeneem word, naamlik eerstens as die meervoud van **laslappie** en daarna as **laslapwerk**. Hier is tog 'n definitiewe semantiese verband tussen die betekenis van **laslappies**, en hulle sou eerder as poliseme onderskei moes word. Vergelyk ook die homonimiese hantering van **lompe**:

¹lompe ... Verslete of verflenterde klere: ...

²lompe ... Iem. of iets wat lomp is: ...

³lompe ... Verboë vorm van *lomp*.

²lompe en ³lompe moes slegs polisemies en meerfunksioneel van mekaar onderskei word. Let hier ook op die plasing van die vreemdste vorm van **lompe** as eerste homoniem. Vergelyk ook die onverstaanbare homonimiese aanbieding van ³log en ⁴log waar die betekenisoreenkoms uiters duidelik uit die definiens blyk.

³log ... afk. (in wiskundige berekening) logaritme: ...

⁴log ... s.n.w., logs. (gew. in minder formele taal) Sien LOGARITME: ...

Aan die ander kant is daar polisemiese lemmas aangetref waarvan die betekenis so ver verwyder is van mekaar dat hulle eerder as homonieme opgeneem

kon word. Vergelyk die foutiewe polisemiese opname van die subleksikale lemmas **-logie** en **-loog**:

-logie ...

- 1 Slotkomponent van selfst. komposita m.d. bet. "leerstelling, teorie of wetenskap", bv. *antropologie, biologie, ekologie, ...*
- 2 Slotkomponent van selfst. komposita m.d. bet. "woord, rede, spraak of gesprek", bv. *doksologie, tetralogie, toutologie, trilogie.*

-loog ...

- 1a Slotkomponent van selfst. komposita m.d. bet. "deskundige, student en (of) beoefenaar v.d. studieveld aangedui deur die eerste komponent", bv. *antropoloog, bioloog, dialektoloog, entomoloog, ...*
- 2 a Slotkomponent van selfst. komposita m.d. bet. "woord, rede, spraak of gesprek", bv. *dialog, epiloog, monoloog, proloog.*

By komplekse lemmas soos samestellings word daar in die WAT IX werklik probleme ondervind t.o.v. homonimiese of polisemiese ordening. Wat uit die oog verloor word, is dat die geheelbetekenis van samestellings, veral dié wat as verklaarde lemmas optree, nie sonder meer afgelei kan word uit die betekenis van die afsonderlike komponente nie. Daaruit spruit voort dat indien 'n komponent van 'n samestelling as simpleks dalk 'n polisemiese verhouding met 'n ander simpleks het, hierdie polisemiese verhouding kan verval indien dié simplekse binne samestellings optree. Vergelyk die hantering van **lugsteen** en **lugweg** as poliseme:

lugsteen ... s.nw.

- 1 Steen met gate of gleuwe wat voor 'n opening in 'n muur ingebou word om lugsirkulasie te bevorder: ...
- 2 ... Metoor: ...

lugweg ... s.nw.

- 1 ... Enigeen van 'n verskeidenheid weë waarlangs lug die lange bereik of verlaat: ...
- 2 ... Lugroete: ...

Alhoewel hier by albei samestellings polisemiese verhoudings tussen die afsonderlike komponente as simplekse kan bestaan, bv. **lug** (vir asemhaling, hemelruim), **steen** (mensgemaakte of natuurlike klip), of **weg** (figuurlike pad, pad), is daar tussen die samestellings geen sinchroniese of psigologiese verwantskap nie. Of enige taalgbruiker die kanaal waarlangs lug vir asemhaling en die roete waarlangs vliegtuie vlieg as verwant in betekenis sal ervaar, is te betwyfel.

Die samestelling **lugrooster** word wel heeltemal korrek as homonieme gehanteer alhoewel daar ook 'n duidelike polisemiese verhouding tussen die afsonderlike komponente bestaan:

- ¹**lugrooster** ... s.nw. Rooster, gew. i.d. muur of fondament van 'n gebou, waardeur lug sirkuleer: ...
²**lugrooster** ... s.nw. Vlugrooster: ...

Waar die gebruiker eerder 'n polisemiese aanbieding verwag op grond van die betekenisverwantskap by die onderskeie betekenis van die lemma **laaikas** het die *WAT IX* besluit om dit as homonieme aan te bied:

- ¹**laaikas** ... s.nw. Lae kas van ong. tafelhoogte met laaie; ...
²**laaikas** ... s.nw. Enigeen van verskeie kaste of toestelle waarin of waaruit bv. erts by 'n myn of ruyster by 'n oond gelaai word: ...

In Feinauer (1993: 293) word daar melding gemaak van die volgehoue korrekte verwysings wanneer daar na 'n polisemiese of homonimiese lemma verwys word. In *WAT IX* word hierdie tipe verwysing ook aangetref, maar dit word ongelukkig nie konsekwent gedoen nie. By die lemma **lansering** word daar wel na 'n homonimiese onderskeiding van **lanseer** verwys, maar by die lemma **lanseerder** nie:

lansering ... s.nw. Handeling of proses van iets te lanseer (¹**LANSEER**);

lanseerder ... s.nw.

- 1 Iets wat lanseer, veral 'n toestel of struktuur wat bv. 'n vuurpyl of geleide missiel ondersteun ...
 2 Iem. wat lanseer: ...

By die korter definies van die tweede poliseem van **lanseerder** is die verwysing noodsaaklik vir die gebruiker om die regte afleiding te maak. By sirkeldefinisies soos by **lanseerder** is volledige verwysing absoluut onontbeerlik, want die gebruiker word eksplisiet deur die tipe definies na 'n ander lemma verwys. Die ontsluiting van die betekenisinligting, en dus die toeganklikheid van die woordeboek, sal juis verhoog word indien die gebruiker onmiddellik weet watter betekenis of betekenisonderskeiding ter sake is. Ander sirkeldefinisies waar verwysing ontbreek, is bv. by **lasweerstand** en **lopietempo**:

lasweerstand ...s.nw. (*elektronika*)

- 1 Weerstand van, of gevorm deur 'n las; sin. *belastingweerstand*: ...

lopietempo ... s.nw. (*krieket*) Tempo waarteen lopies aangeteken word; sin. *kolftempo*; ...

Dit lyk asof die gebruiker se ontsluitingstaak ietwat vergemaklik word deur die verskaffing van etikette sowel as sinonieme, maar vir die geskikte betekenis van *las* sal die gebruiker moet soek tot by die tweede kolom van die tweede homoniem van *las*. Indien hy / sy die sinoniem wil naslaan, moet deel I geraadpleeg word, en daar verskyn **belastingweerstand** met 'n ander etiket, nl. (*tegn.*). Hierdie artikel het weer eens twee poliseme waartussen die gebruiker moet onderskei. Hoeveel makliker en vinniger sou die ontsluiting van **lasweerstand** se betekenis nie wees indien daar 'n verwysing in die definies naas *las* verskyn het nie!

Waar proefleesfoute onder geen omstandighede mag voorkom nie, is by die aangawe van verwysings. In WAT IX is twee sodanige foute aangetref, naamlik by ²ligting en by luier-luier:

²ligting ... Proses of resultaat van te lig (²LIG III 2): *In die ooste is daar die flou ligting van die naderende dag* (Huisg., 23 Febr. 1940, 53).

Die verwysing moes lui (¹LIG III 2): ²lig dui op die adjektief wat beteken "met 'n relatiewe ligte massa".

luier-luier ... bw. Iter. van *luier* (¹LUIER II 1). Op 'n luierende manier: ...

Daar bestaan hoegenaamd nie so 'n verwysing nie: ¹luier is die vergrotende trap van *lui*. ²luier het egter nie I- en II-onderskeidings nie. Die korrekte verwysing is (²LUIER 1 b) wat dui op "ontspanne of talmend voortbeweeg".

Waar hierdie tipe verwysing as oorbodig aan gevoel word, is by die onderskeie polisemiese onderskeidings van ²lig waar die sekondêre betekenisonderskeidinge telkens terugverwys na die primêre betekenis. Dit is tog net logies dat die primêre betekenis oorkoepelend geld vir die sekondêre betekenisonderskeidinge wat juis direk daaronder geplaas word om die noue polisemiese verband met die primêre betekenis aan te toon. Die verwysings verswaar net die definies sonder dat enige bykomende inligting ontsluit word:

²lig ...b.nw., ligte; ligter, ligste.

1 Met 'n relatiewe lae massa; teenoor *swaar*: ...

Meer bep:

a Van 'n ligte (²LIG 1) ontwerp om maklike en vinnige ...

b Lig (²LIG 1) bewapen en toegerus ...

c Ligter (²LIG 1) as die korrekte, normale, ...

d Lig (²LIG 1) relatief tot grootte of omvang: ...

2a Wat rats, maklik of met vaardigheid ...

Gouws (1986: 76) merk op dat "(a) compound with one or more polysemous components may not be entered unexplained if the dictionary states that unexplained compounds have no semantic problems because of their transparency".

Tog gaan WAT IX steeds voort om ongeveer 2 500 onverklaarde samestellings, waarvan die meeste polisemiese komponente het, op te neem aan die einde van 'n artikel sonder enige polisemiese verwysing. By bv. die artikel **liggaam** word twintig onverklaarde samestellings aangetref: **liggaam** word in vier primêre en in verskeie ander sekondêre polisemie verdeel wat oor ongeveer drie kolomme strek. By elke optrede van **liggaam** as deel van die samestelling moet die gebruiker self besluit watter betekenisonderskeiding van **liggaam** geaktiveer word. Indien die Redaksie verkies om nie by verdere dele van die WAT verwysings by onverklaarde samestellings te verskaf nie, moet daar sover moontlik probeer word om die onverklaarde samestellings te lys aan die einde van die betrokke polisemiese onderskeiding in die artikel, eerder as om almal aan die einde van die artikel te plaas.

Op (p. xv par. 2.9) in die **Toeligting** word die simbole genoem wat gebruik word om die verskillende betekenisonderskeidings en gespesialiseerde gebruike aan te dui en te orden. Steeds word daar egter geen inligting gegee oor die stelsel waarvolgens die verskillende betekenisonderskeidings georden word nie. Vergelyk hier Feinauer (1993: 292) en Gouws (1989: 132) wat sê dat dit juis die leksikograaf se plig is om in sy woordeboek 'n duidelike uiteensetting te gee van die ordeningstipe wat in die woordeboek gevolg word, asook van die leksikografiese implikasies daarvan. Volgens watter maatstaf is bv. besluit om die twee primêre betekenisonderskeidings by die lemma **landloper** soos volg te plaas:

landloper ... s.nw.

- 1 Persoon wat geen vaste woonplek of gereelde werk het nie en van plek tot plek beweeg, ...
- 2 Atleet wat landloop (LANDLOOP 1) beoefen; ...

Is bogenoemde plasing op grond van die databasisfrekwensie of diachronies chronologies? Die sinchronies algemeenste betekenisonderskeiding word tweede geplaas. Alle ander lemmas waarin **landloop** voorkom, behalwe **landloper**, handel ook uitsluitlik oor die atletieksoort. Die verwysing by **landloper 2** na **landloop 1** is dus onnodig: daar is net een primêre betekenis in die artikel van **landloop**. By **landloper** egter is daar geen verwysing nie, en daar moet uit die sitate afgelei word dat dit om **landloper 1** gaan:

landloper ... s.nw. Handeling, daad of toestand van 'n landloper te wees:

...

Wat opval in WAT IX is dat daar heelwat lemmas is waar die sekondêre betekenisonderskeidings (klein letters) aangedui word sonder dat daar enige hiërargiese ordening (syfers) blyk: die primêre of oorkoepelende betekenisonderskeiding ontbreek naamlik. Dit lyk dan na 'n lukrake groepering van betekenisonderskeidings. By die lemma ²lig 1 gaan die WAT baie goed te werk met die

oorkoepelende betekenisonderskeiding wat eers gelys word, en daaronder die sekondêre betekenisonderskeidings wat dikwels meer verskillende gebruike aandui. Sien die lemma hier bo. Die inskrywings onder ²lig 2 toon egter geen sekondêre betekenisverwantskap t.o.v. mekaar nie, en daar word ook geen primêre betekenisonderskeiding gegee nie:

²lig b.nw. ...

- 2 a** Wat rats, maklik of met vaardigheid beweeg: ...
- b** Met min momentum, druk, fermheid, of krag: ...
- c** Nie dik, warm of baie formeel nie: ...
- d** Wat min is in volume of t.o.v. substansie: ...
- e** Los en poreus en maklik bewerkbaar: ...
- f** Met of van 'n ligte digtheid of dikte: ...
- g** Met 'n klein kaliber: ...

Vir dieselfde probleem vergelyk o.a. ²laai I 3.

Wat tog verblydend is, is dat daar in WAT IX minder blyke is van skynbetekenis, gebruike en saakbeskrywing wat as betekenisonderskeidings aangebied word. Vergelyk bv. die goeie hantering van lem I:

lem ...

I snw. ...

- 1 a** Gedeelte van 'n stuk gereedskap, implement of wapen wat 'n snykant, punt of tande bevat waarmee daar veral gesny of gesteek word; ... *Die lem van 'n mes, skalpel, beitel, skaaf, freesmasjien, skaapskêr, sekel, graaf, skoffel, padskraper, byl, kapmes, bandsaag, vyl, vleismeul, assegaai, dolk, lans, swaard. ...*

Dit is goeie leksikografie: die onderskeie sake waar 'n lem voorkom, word nie as betekenis aangebied nie, maar heeltemal korrek as voorbeeldmateriaal gelys. Hoe die saak lyk, en waarvoor dit gebruik word, word ook nie as afsonderlike polisemiese waardes aangebied nie, maar binne een onderskeiding saamgevat. Dieselfde goeie betekenisleiding word as primêre betekenis by die lemma *liefde* (bet. 1) aangebied. Ongelukkig word hier nog te veel gevalle aangetref waar kontekstuele inligting as primêre polisemiese onderskeidings aangebied word. Volgens Gouws (1987: 92 en 1989: 196-199) is dit semanties foutief om lemmas soos kort en lank te definieer as sou hulle twee polisemiese onderskeidings, naamlik tyd en liniariteit hê. Lank dui semanties bloot op 'n afmeting van groot uitgestrektheid. Of hierdie afmeting temporeel of lineêr is, is deel van lank se kontekstuele gebruik; nie van sy leksikale betekenis nie. Vergelyk nou die WAT IX se hantering van lang:

lang

I b.nw. ...

- 1 a** (*m.b.t. mense, diere, liggaamsdele, voorwerpe, lyne, afstande*) Van

groot of genoegsame afmeting in lengterigting, gemeet langs die grootse dimensie, van kop tot voete, van kop tot stert of van punt tot punt; ...

2 a (m.b.t. 'n tydsuur, die duur van 'n handeling of toestand, die verloop van 'n gebeurtenis, e.d.) Wat nie gou verbygaan of verby is nie of baie tyd in beslag neem of geneem het; ...

3 a (m.b.t. menslike uitinge in woord, musiek, geskrif, e.d.) Van groot lengte of omvang en (of) van geruime duur; ...

Let daarop dat liniariteit en tyd as twee aparte primêre betekenisonderskeidings aangebied word, met as derde primêre poliseem een waarin die twee kontekstuele toepassings van lang, n.l. tyd en liniariteit gekombineer word. Die derde polisemiese waarde word dus vasgestel t.o.v. die toepassings van die kontekstuele gebruike van lang. Hier is die gebruiker reeds baie ver verwyder vanaf die leksikale betekenis van lang. Poliseme 1a en 2a moes in een primêre onderskeiding 1 verenig word met die lineêre en temporele onderskeidings as sekondêre onderskeidings (a en b), terwyl 3 as deel van verdere sekondêre verfyning kon optree (i, ii en iii).

Veral wat die mikrostrukturele bewerking van polisemiese lemmas betref, is hier nog heelwat waarop komende dele van die WAT kan verbeter.

4. Slot

Alhoewel hier heelwat kritiek op sommige bewerkings deur die WAT IX uitgespreek is, is dit onteenseglik 'n groot verbetering op alle vorige dele. Wat veral verblydend is, is dat die Redaksie duidelik kennis neem, nie net van kritiek wat op vorige dele uitgespreek is nie, maar ook van die nuutste leksikografieteorie. Dit lei dan oor die algemeen tot die beter optekening en bewerking van lemmas. Die heelwat hoër leksikografiese standaard van hierdie eerste WAT onder die hersiene redaksionele stelsel, hou vir die taalgebruiker sowel as vir taalkundige die belofte in van verdere konsekwent uitstekende leksikografiese werk.

Aantekening

- * Van Schalkwyk, D.J. (Hoofredakteur). 1994. *Woordeboek van die Afrikaanse Taal*, Negende deel. Stellenbosch: Buro van die WAT. Prys: plastiekband R89,80; leerband R158,20.

Verwysings

- Beyer, H.L. 1995. *Die leksikografiese hantering van morfologies gemerkte geslagsopposisiepare in Afrikaanse woordeboeke, met spesifieke verwysing na die Verklarende Handwoordeboek van die Afrikaanse Taal*. Ongepubliseerde M.A.-verhandeling. Universiteit van Stellenbosch.

- Beylefeld, Adri en Gert van Jaarsveld.** 1994. Is die wêreld manlik, tensy anders vermeld? *Suid-Afrikaanse Tydskrif vir Taalkunde* 12 (2): 43-49.
- Buro van die WAT.** 1994. *Beleid vir die hantering van beledigende en sensitiewe leksikale items in die Woordeboek van die Afrikaanse Taal.* Ongepubliseerde verslag. Stellenbosch: Buro van die WAT.
- Combrink, J.G.H.** 1990. *Afrikaanse morfologie: capita exemplaria.* Pretoria: Academica.
- Feinauer, A.E.** 1993. Die agste deel van die *Woordeboek van die Afrikaanse Taal.* *Tydskrif vir Geesteswetenskappe* 33 (4): 286-298.
- Feinauer, Ilse.** 1994. An Afrikaans Learner's Dictionary. *South African Journal of Linguistics* 12(2): 60-67.
- Geeraerts, Dirk.** 1991. Principles of Monolingual Lexicography. Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta (Reds.). 1989-1991: 287-296.
- Gouws, R.H.** 1985. Die sewende deel van die *Woordeboek van die Afrikaanse Taal.* *Standpunte* 178: 13-25.
- Gouws, R.H.** 1985a. Die omvattende sinchroniese woordeboek as taalkundige handleiding. *Suid-Afrikaanse Tydskrif vir Taalkunde* 3(3): 21-47.
- Gouws, R.H.** 1986. Afrikaans Lexicography. Hartmann, R.R.K. (Red.). 1986: 69-77.
- Gouws, R.H.** 1987. Lexical Meaning versus Contextual Evidence in Dictionary Articles. *Dictionaries* 9: 87-96.
- Gouws, R.H.** 1989. *Leksikografie.* Kaapstad: Academica.
- Gouws, R.H.** 1990. Vaste uitdrukkings as multileksikale lemmas in Verklarende Afrikaanse Woordeboeke. *Tydskrif vir Geesteswetenskappe* 30(4): 265-283.
- Gouws, R.H.** 1991. Reviews: Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta (Reds.): Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography. Berlin: Walter de Gruyter, 1989-1991. *Lexicographica* 1991(7): 268-290.
- Gouws, R.H.** 1994. *Woordeboek van die Afrikaanse Taal, negende deel.* SAUK, Skrywers en Boeke.
- Gouws, Rufus H.** 1995. Dictionaries and the Dynamics of Language Change. Kachru, Braj B. en Henry Kahane (Reds.). 1995: 297-313.
- Gouws, R.H. en F.A. Poneelis.** 1992. The Development of Afrikaans and the Lexicographical Tradition. Zgusta, L. (Red.). 1992: 77-104.
- Hartmann, R.R.K. (Red.).** 1986. *The History of Lexicography.* Amsterdam: John Benjamins.
- Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta (Reds.).** 1989-1991. *Wörterbücher. Dictionaries. Dictionnaires. An International Encyclopedia of Lexicography.* Berlyn: Walter de Gruyter.
- Hausmann, Franz Josef en Herbert Ernst Wiegand.** 1991. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta (Reds.). 1989-1991: 328-360.
- Kachru, Braj B. en Henry Kahane (Reds.).** 1995. *Cultures, Ideologies, and the Dictionary. Studies in Honor of Ladislav Zgusta.* Tübingen: Max Niemeyer.
- Landau, S.I.** 1984. *Dictionaries: The Art and Craft of Lexicography.* New York: Scribner Press.
- Lombard, F.J.** 1992. Voorbeeldmateriaal in woordeboeke. *Lexikos* 2: 148-164.
- Odendal, F.F., P.C. Schoonees, C.J. Swanepoel, S.J. du Toit en C.M. Booysen.** 1994. *Verklarende Handwoordeboek van die Afrikaanse Taal.* Midrand: Perskor.
- Otto, Anel.** 1994. Rekenaar bespoedig WAT. *Die Burger*, 8 Junie 1994: 5.

- Rademeyer, L. 1992. *Die funksies van verbale voorbeeldmateriaal in eentalige woordeboeke*. Ongepubliseerde M.A.-verhandeling. Universiteit van Stellenbosch.
- Raidt, E.H. 1992. Agste deel van WAT toon vooruitgang, maar leemtes nog nie besweer. *Die Burger*, 23 Junie 1992: 7.
- Taalkommissie. 1964. *Afrikaanse Woordelys en Spelreëls*. Kaapstad: Tafelberg.
- Taalkommissie. 1991. *Afrikaanse Woordelys en Spelreëls*. Kaapstad: Tafelberg.
- Vos, I.L. 1992. *Semantiese norme in Afrikaanse verklarende woordeboeke*. Ongepubliseerde M.A.-verhandeling. Universiteit van Stellenbosch.
- Zgusta, L. (Red.). 1992. *History, Languages, and Lexicographers*. Tübingen: Max Niemeyer.

Resensies / Reviews

Woordeboeke en Woordelyste / Dictionaries and Word-lists

William Branford. *The South African Pocket Oxford Dictionary of Current English*, 2nd South African edition 1994, 2nd impression 1995, xxiii + 1138 pp. ISBN 0-19-570760-5. Cape Town: Oxford University Press. Price R54,95.

Australian, New Zealand and South African pocket dictionaries of English are relatively recent additions to the burgeoning proliferation of Oxford dictionaries. This second edition of *The South African Pocket Oxford Dictionary of Current English* (SAPOD), which is based on the eighth edition of *The Pocket Oxford Dictionary of Current English* (1992) edited by Della Thompson, has benefited greatly by the advances in modern technology by being able to draw on the latest updated computerized databases of the Oxford English dictionaries. This has enabled the compilers to include such topical entries as **ethnic cleansing, karaoke, pixel, toy boy and virtual reality.**

Comparing this 'pocket' dictionary with my old 1964 edition of the *Concise Oxford Dictionary* (which is about the same size), I was struck by how user-friendly this new version is. Some of the features making this dictionary more accessible to the user are set out below.

Headwords are set in bold type. A conscious policy of 'denesting' has been followed resulting in words being easier to find. For instance, compound words such as **French leave** are written as two words instead of being 'nested' at the end of entries. The various senses of a word are numbered instead of being separated by semicolons. Explanations of etymologies are set out in a more straightforward way than in many earlier dictionaries. Specific information about the forms of words has been given, such as difficult plural forms (e.g. **syllabuses / syllabi**). Pronunciation of words is given in a phonetic transcription (using IPA symbols).

The notes on usage are one of the most helpful features of this dictionary. These notes include guidance on aspects of traditional grammar such as noun phrases, transitive verbs, adverbs and the like, as well as old chestnuts such as the difference between *shall* and *will*, *due to* and *owing to*, *affect* and *effect* etc. There are also helpful comments on the use of noun class markers when referring to the names of Suntu languages or ethnic groups such as *amaXhosa*. Some of the most helpful notes are those on the South African usage of (often loaded) terms such as *baas*, *Bantu*, *boy*, *Bushman*, *comrade*, *girl*, *man*, *master* and *Springbok*. The note on *girl*, quoted below, gives a good idea of the nature of these notes:

Many S. Africans still say *girl*, *housegirl*, *washgirl* etc. of a Black woman servant without intending offence, but these terms are deeply resented by others.

The special distinguishing feature of *SAPOD* is, of course, its South African English (SAE) component which has drawn heavily on Jean and William Bransford's *A Dictionary of South African English*⁴ (1991) (*DSAE*). Entries of SAE origin constitute about 2500 of the 65000 entries in *SAPOD*. About half of *DSAE*'s corpus plus a number of new words appear in *SAPOD*. Generally the words not included in *SAPOD* were marginal items in *DSAE*. A quick comparative test-drive through some of the entries under the letter B of the two dictionaries helps to illustrate this point:

The following words appear in both dictionaries: **-baai, baas, baasskap, baba, babalaas, backvelder, baie dankie, bakgat, balie, bandiet.**

The following words in *DSAE* have not been included in *SAPOD*: **babbie-shop, bafta, bakkis, bakoond, bakore, baliestoel, ballasmandjie.**

The selection of SAE words in *SAPOD* is fairly extensive, including not only common SAE words such as **fundi, indaba, laatlammetjie, lekker, lobola, lucky bean, pap, wors**, but also less commonly used words which are of special socio-political significance: **imbizo, inkululeko, ubuntu**, or socio-cultural terms: **imbongi, isanusi, liretlo, manyano**. There are also numerous entries of specialized vocabularies relating to various subcultures such as troepie-talk: **bokkie, bossies, leopard crawl, min dae**; prison slang: **fransman, impimpi, lightie**; shebeens: **gweva, mahewu, mailer, mbamba, shebeen queen, skokiaan**; township music: **malombo, mbaqanga**.

Many common South African expressions have also been included: **banana boy, moenie panic nie, monkeys' wedding, môre is nog 'n dag, Natal fever, skop, skiet en donder.**

Many current SAE abbreviations also appear in the dictionary: **ANCYL, ASB, AWB, CCB, CFE, ESKOM, MK, SACP.**

The main languages of origin of the SAE entries are Dutch/Afrikaans, English and Sintu. There is a significant drop in the Afrikaans entries from 52% (in the 1987 edition) to 40% (in this 1994 edition), indicating that many of the marginal Afrikaans items have been dropped, and a number of new items mainly of (South African) English origin have been included, making the composition of the entries more representative of current SAE.

Being a pocket dictionary, there are naturally certain limitations on its scope. Proper nouns (except those of local South African relevance) and 'foreign' words (e.g. *babouche*) are generally excluded. Definitions of words are concise, but sometimes lacking in important details, and examples of usage are limited. Nevertheless, on the whole, this dictionary covers a good range of current Standard English and SAE. Given its limitations as a pocket dictionary, it is of use to anyone who would like to have current general and South African usage of English at their fingertips.

L.A. Barnes
 Department of Linguistics
 University of South Africa
 Pretoria
 South Africa

J. Combrink en J. Spies. *SARA. Sakboek van regte Afrikaans*, 2de uitgawe, 1ste druk 1994, 214 pp. ISBN 0 624 03252 3. Kaapstad: Tafelberg-Uitgewers. Prys R49,95.

1. Inleidend

Alle taalpraktisyns wat hulle sout werd is, doen gedurig en gereeld wat die skrywers van *SARA*² herhaaldelik by gebruikers van die boek aanbeveel: versamel vir jou 'n klompie betroubare standaardbronne en slaan maar altyd weer na. *SARA*² is een van dié soort bronne (wat geriefshalwe hier "taalhandleidings" genoem word) wat 'n taalwerker kan gebruik, maar die skrywers het klaarblyklik ook hulle boek gemik op almal wat in die loop van hulle dagtaak in Afrikaans moet skryf of vir wie keurige Afrikaans van belang is. Die vraag is nou in watter mate *SARA*² daarin slaag om te voldoen aan die behoeftes en verwagtinge van so 'n uiteenlopende gebruikersgroep.

Daar is verskeie soortgelyke taalhandleidings waarmee *SARA*² vergelyk kan word. Onder dié wat hierdie resensent op sy tafel het, tel die eerbiedwaardige *Die Juiste Woord* van L.W. Hiemstra (Tafelberg-Uitgewers, Kaapstad, 1980), 'n baie behoudende en nuttige bron wat plek-plek ly onder Hiemstra se eksentrieke beskouinge en natuurlik nou al sterk verouderend is, en dus oordeelkundig gebruik moet word. H.J.J.M. van der Merwe en F.A. Ponelis se *Die Korrekte Woord*⁶ (J.L. van Schaik, Pretoria, 1982) is 'n eweneens nuttige hulpmiddel, veral as 'n mens goeie, onbesproke Standaardafrikaans wil kontroleer en jy Ponelis se soms sonderlinge inskrywings kan uitken en ignoreer (om welke rede ek die vyfde uitgawe, 1975, verkies). *Die Kennis-gids tot Moderne Afrikaans*² van E.P. du Plessis (Human & Rousseau, Kaapstad, 1981) kan ook met vrug gebruik word, veral vanweë die ongekompliseerde inskryfstelsel, maar hy is ook pre-1991-AWS en verouderend. Moderner is A. Coetser se *Nuwe Taalgids* (Maskew Miller Longman, Kaapstad, 1992), wat ek besonder ontoeganklik vind en waarin ek selde kry wat ek soek.

Hoe vaar *SARA*² in hierdie geselskap?

2. Verwysingsmetodes

*SARA*² kom, soos haar ouer sussie *SARA*¹, in 'n duursame plastiekomslag, dié keer helder geel. Dit is 'n besonder smal boekie, so asof die "Sakboek"-deel van die titel heel letterlik opgeneem is. Die boekie is ongemaklik op 'n boekrak of lessenaarboekstander, maar sal inderdaad gerieflik in 'n baadjie- of handsak pas. En dit is miskien waar *SARA*² hoort: Dit bevat 'n magdom taal- en ensiklopediese inligting waarvan die gebruiker slegs sal weet as sy toevallig daarop afkom. As die boekie dus inderdaad in die sak gedra en van voor af deurgelees word, byvoorbeeld deur 'n pendelaar-gebruiker op die trein of bus, kan dit 'n

baie nuttige inligtingsbron word. Onder die onwaarskynlike lemma **klassieke woorde** word byvoorbeeld allerlei interessante en waardevolle inligting gegee, maar watter gebruiker van 'n taalhandleiding gaan so 'n lemma soek? Hy moet toevallig daarop afkom, en dan bly die probleem nog steeds latere heropsporing. 'n Mens weet wel jy het die inligting iewers teëgekrom, maar bitter min het die soort geheue wat jou weer direk by so 'n vreemde vindplek sal uitbring. Dit gee 'n gesoek af waarvoor 'n taalwerker selde tyd het.

SARA² maak wel van 'n stelsel van kruisverwysing gebruik. Die gebruiker sal byvoorbeeld aan die einde van die artikel oor **klassieke woorde** lees "Sien kondisie". By **kondisie** kry die gebruiker dan nie net inligting oor die gebruiksveld van **kondisie** nie, maar, as sy oopkop lees, ook oor nuanse- en gebruiksverskille tussen woorde van Romeanse en Germaanse herkoms. Dit sluit aan by **klassieke woorde**, waarna daar weer 'n verwysing is.

Naas die gewone alfabetiese gelyste lemmas het SARA² ook 'n Register waarin sleutelwoorde gelys word wat in die artikels voorkom maar nie deel van die hoof alfabetiese reeks uitmaak nie. Dit maak dit oor die algemeen makliker vir die gebruiker om uit te kom by wat hy soek, hoewel dit nie al die probleme oplos nie.

In die stelsel van kruisverwysing en in die Register kom daar glipse soos die volgende voor: By **advies** is byvoorbeeld een van die kruisverwysings **ekstra**. Laasgenoemde hoort m.i. egter by die vorige lemma (**adjektiewe, onverboë**); by **ekstra** self is daar geen korrelerende kruisverwysing nie. In die Register is die bladsyverwysing by *Roose* byvoorbeeld na dié van SARA¹ (p. 117) in plaas van SARA² (p. 130). By die lemma **aanstel, benoem, kies** is daar 'n kruisverwysing na **tot**, maar daar is nie so 'n lemma nie. Daar is wel 'n lemma **tot sy dood**, maar dié artikel het m.i. weinig te make met wat by **aanstel, benoem, kies** gesê word. Is die kruisverwysing foutief of nie? Daarteenoor is daar in die alfabetiese reeks die lemmas **benoem tot en kies as**, waarby telkens staan "Sien **aanstel**". Weer eens is daar nie 'n lemma **aanstel** nie (slegs **aanstel, benoem, kies**) maar in dié geval sal die gebruiker darem direk kan aflei dat **aanstel, benoem, kies** inderdaad die bedoelde lemma is.

Daar is ook inkonsekwentheid in die verwysingstelsel. Op p. 29 is daar twee baie nuttige artikels wat die verskille verduidelik tussen woorde wat dikwels verwar word, t.w. **blyk, lyk, skyn** en **blykbaar, skynbaar**. Die trefwoorde van dié twee lemmas kom net hier voor: nie as kruisverwysings in die alfabetiese reeks nie (vgl. **aanstel, benoem, kies** hierbo) en ook nie in die Register nie. As die gebruiker dus nie kan raai dat die skrywers dié woorde saamgegroeper het en onder **B** ingeskryf het nie, gaan sy nie daarby uitkom nie.

Daarteenoor is daar lemmas soos **bly** (ook op p. 29), waar die gebruiker daar en dan verwys word na **woon, bly** in die alfabetiese reeks, en soos **lamp, lig** (p. 97), waar dieselfde stelsel nie ten opsigte van **lig** gevolg word nie. Die gebruiker kan egter via die Register by die artikel **lamp, lig** uitkom. Eienaardig genoeg is daar wel 'n artikel **lig** (wat oor iets heeltemal anders handel) maar geen kruisverwysing na **lamp, lig** nie.

Dan is daar ook die lemma **helfte van die jaar**, waarby eenvoudig staan "Sien **laaste**". Dit is onduidelik hoekom — daar is niks by **laaste** wat met "helfte van die jaar" te doen het nie.

3. Lemmas

Iets anders wat vreemd voorkom en waarvoor daar nie 'n voor die hand liggende verklaring is nie, is die verskillende style wat by lemma-inskrywing gevolg word: **anachronisme**, **anakronisme** word só ingeskryf terwyl die twee bloot wisselvorme is. Eweneens word **feitelik**, **feitlik** só ingeskryf, maar hier word die betekenisverskil bespreek. Dan is daar nog lemmas soos **f of v**, **links of regs**, **enkelvoud of meervoud** en **laat of dat** (die "of" nie vet druk nie), sodat dit op die ou end moeilik is om te weet wat die verhouding tussen die trefwoorde van die lemma is.

Soms is die lemma deel van die (eerste) verklarende sin, hoewel meestal nie. 'n Tipiese voorbeeld van hierdie tipe inkonsekwentheid kom voor op p. 138: "**permissief** 'n Modewoord wat dikwels ...", maar "**persepsie** het skielik 'n beminde modewoord geword". Gelukkig is streng konsekwentheid of patroonmatigheid in 'n taalhandleiding soos SARA² nie van sulke groot belang as in byvoorbeeld 'n woordeboek nie. In laasgenoemde geval moet die gebruiker dikwels juis uit die *manier* van inskryf inligting aflei, terwyl sulke inligting in 'n taalhandleiding meestal vir die gebruiker in die artikel uiteengesit word. Nietemin vind ek dit soms lastig dat ek moet sukkel om te onderskei tussen verduideliking, kommentaar, redenering, voorbeeldmateriaal en beklemtoning.

Soms is lemmas alfabeties gelys onder die "bepaler" en soms weer onder die "kern". So kry die gebruiker byvoorbeeld **Afrikaans**, **goeie** en **sinne**, krom teenoor **regte Afrikaans** en **lang sinne**. Dit dwing die gebruiker om weer te raai waar hy moet soek: byvoorbeeld by **Engelse meervoude** en **sterk werkwoorde** dan wel by **meervoude**, **Engels** of **werkwoorde**, **sterk** (dié twee, terloops, is alfabeties by onderskeidelik **E** en **W** ingeskryf). En dan is daar ook nog 'n lemma soos **agter geslote deure** wat onder **D** (p. 40) ingeskryf is.

4. Vindbaarheid

Een van die probleme wat gebruikers van boeke soos SARA², ensiklopedieë e.d. ondervind, is "naslaanbaarheid" of "vindbaarheid": Waar kry 'n mens wat jy soek? SARA² gee baie, en baie nuttige, leiding oor allerlei eiename: plekname, buitelandse geografiese name, klassieke persoonsname en dies meer. Hier kom naslaanbaarheid ter sprake, want SARA² maak soms die lemma die Engelse naam of vorm en soms die Afrikaanse. So word **Pliny** en **Plutarch** byvoorbeeld die lemma, maar wel **Pompeji** en **Regensburg**. Gelukkig maak dit heel dikwels nie veel (alfabetiese) verskil nie: die meeste gebruikers sal waar-

skynlik ewe goed by *Pliny* as by *Plinius* uitkom, of by *Justinian* as by *Justinianus*. In gevalle soos *Ratisbon* x *Regensburg* en *Cologne* x *Keulen* is dit egter 'n ander saak. Weer eens: as die gebruiker toevallig aanvanklik 'n keer of wat Engelse vorme as lemmas teëgekrom het, sal sy waarskynlik gereeld verwag om haar antwoord op die "Engelse plek" te kry.

Ook hier is die verwysingstelsel inkonsekwent of onvolledig. Die Register help die gebruiker om byvoorbeeld van *Sorlings* by *Scilly Islands* en van *Munich* by *München* te kom, maar nie om van *Ratisbon* by *Regensburg* of van *Cologne* by *Keulen* te kom nie. SARA² bly nietemin een van die min bronne wat oor eiename leiding gee.

Van plekname gepraat: Op pp. 141-143 onder plekname se skryfwyse sit die skrywers die "nuwe" skryfwysebeleid van die Nasionale Pleknamekomitee heel uitvoerig uiteen. Hoewel dit inligting is waartoe die breë publiek selde toegang het, en dit dus by die deursigtige nuwe Suid-Afrika pas, is dit m.i. hier heeltal onvanpas. Dit verbaas 'n mens in die eerste plek dat die NPNK 'n beleid kon aanvaar, en amptelike gesag daaraan kon laat gee, wat in verskeie opsigte in stryd is met die 1991-AWS (8ste uitgawe, 2de druk, Tafelberg-Uitgewers, Kaapstad, 1991) (vgl. byvoorbeeld AWS, pp. 41 en 42, parr. 3.10.1 en 3.10.3, en p. 52, par. 6(1), (2) en (3), met SARA², p. 142, parr. 4, 8 en 9). In die tweede plek — en belangriker vir 'n taalhandleiding — bring dit die gebruiker niks in die sak om dié inligting te hê nie, want dit help hom hoegenaamd glad nie om die korrekte skryfwyse van 'n "ou" óf 'n "nuwe" binnelandse (of trouens buitelandse) plek- of geografiese naam te voorspel of af te lei nie. Dié artikel bring m.i. meer verwarring as hulp.

5. Anglisismes

Soos reeds genoem, roer die skrywers onderwerpe van die mees uiteenlopende aard aan, en soos 'n mens in 'n Afrikaanse taalhandleiding kan verwag, ook die kwessie van Anglisismes. Ek vind die skrywers se beskouing van Anglisismes vir doeleindes van SARA² (p. 15: "In dié boek word alles wat uit Engels in Afrikaans gekom het Anglisismes genoem") besonder ongelukkig. Dit klop nie met wat in byvoorbeeld die eerste paragraaf van die artikel *klassieke woorde* (p. 88) gesê word nie en is in stryd met die heersende menings (bv. p. 8 van die 1991-AWS) oor wat Anglisismes is. Taalwerkers het juis 'n ewige stryd om dit by kliënte, navraers, onderwysers ens. tuis te bring dat NIE alles wat uit Engels kom of na Engels lyk, Anglisismes is nie. Die groot teikengroep op wie SARA² volgens die Voorwoord gemik is en aan wie die skrywers wil "wys hoe oor taal gedink moet word", het darem 'n meer genuanseerde benadering verdien. 'n Mens wonder hoe die artikel *geboortemerk* (p. 60: "geboortemerk word in Afrikaans nie erken nie; *moedervlek* of *huidvlek*") deur dié simplistiese benadering van die skrywers beïnvloed is. Hoe ook al, in die lig van SARA² se Anglisisme-definisie is die *geboortemerk*-tipe uitsprake onder verdenking.

6. Subjektiviteit

Taalhandleidings loop uiteraard gevaar om die versamelplek te word van die skrywer(s) se persoonlike siening en gelofies waarvoor daar dikwels min steun in ander bronne is. SARA² het ook 'n goeie skoot hiervan weg, soos blyk uit onder andere die skrywers se m.i. ongeregverdig skerp verwerping van *kyk* i.p.v. *sien* as verwysingswoord (*sien*, p.158), *aan diens* i.p.v. *op diens* (slotparaagraaf van *op*, pp. 132-133) en dergelike. SARA² is egter uitsonderlik in die sin dat die skrywers darem meestal probeer redeneer hóékom hulle sê wat hulle sê en, veral, omdat hulle herhaaldelik in die boek gebruikers verwys na allerlei ander bronne en hulle aanraai om dié gereeld te gebruik. Veral in die geval van skrywers wat hoë aansien of bekendheid geniet onder diegene wat in taal belangstel, is dit besonder belangrik, want dit bring balans en perspektief, en dit voorkom dat uitsprake as evangelies beskou word.

Die skrywers se redenering is nie altyd ewe geslaagd nie, soos ek hierbo reeds te kenne gegee het. Soms is die inskrywing of artikel eenvoudig te kripties. Vir my verklaar die artikel *ydellik* (p. 207) byvoorbeeld bitter min. En vir die gebruiker wat leiding wil hê (en dit is per slot van rekening waarom iemand 'n taalhandleiding aankoop), help 'n artikel soos *aanhouding*, *arres*, *hegtenis* (p. 2) wat met 'n vraag eindig, ook min of niks.

7. Opsommend

Ten slotte: Wat is die doel van 'n taalhandleiding? Wat verwag die gebruiker daarvan? Volgens my ondervinding wil mense wat sulke boeke gebruik, sinvolle en duidelike leiding hê oor minder of meer kontensieuse taalkwessies en ander taalprobleme wat nie geredelik uit 'n woordeboek of 'n skoolhandboek of 'n grammatika opgelos kan word nie. Van die taalhandleidings wat hierdie gebruiker tans op sy lessenaar het, is SARA² van die bestes: sy is nuut, modern en byderwets sonder om te ver oor te hel na die gedagte dat enigiets wat "die mense sê", ook maar goed genoeg is; sy is by met die nuwe spel- en skryfreëls van die Akademie soos vervat in die 1991-AWS; en sy erken darem dat sy nie die enigste, geïnspireerde waarheid oor Afrikaans verkondig nie.

SARA² het egter ook tekortkominge. Daar is 'n massa nuttige, stimulerende en insiggewende taal- en ander inligting wat ongelukkig dikwels ontoeganklik verskuil lê onder die onwaarskynlikste lemmas en op die onwaarskynlikste plekke. Skrywers van taalhandleidings het natuurlik die probleem dat die kwessies wat hulle moet of wil bespreek, heel dikwels nie maklik gealfabetiseer kan word nie en van gebruiker tot gebruiker van gedaante kan verwissel. Hiervoor behoort 'n verbeterde stelsel van kruisverwysing in sowel die alfabetiese reeks as aan die einde van artikels al baie te help, veral as daar dan nog 'n goeie register ook is.

Gebruikers van taalhandleidings kan baie hulp, raad en leiding uit SARA²

kry, veral as hulle hulle deeglik vertrou maak met die boekie as geheel en miskien hulle eie aantekeninge maak en verwysings inskryf. Ewe belangrik is egter dat die gewone liefhebber van Afrikaans ook heelwat genot kan put uit *SARA*², veral omdat dit in 'n pittige, lewendige styl geskryf is. En al stem 'n mens nie met alles saam nie, dit sal jou minstens laat dink.

J. du T. McLachlan
Taalpraktisyn
Pretoria
Suid-Afrika

Jonathan Crowther (Editor), Kathryn Kavanagh (Assistant Editor) and Michael Ashby (Phonetics Editor). *Oxford Advanced Learner's Dictionary of Current English*, 5th edition (New International Students' Edition) 1995, x + 1428 pp. ISBN 0-19-431423-5. Oxford: Oxford University Press. Price R54,95.

Learner's dictionaries of English are mushrooming nowadays. What does the New International Students' Edition of the *Oxford Advanced Learner's Dictionary* (henceforth *OALD*) offer, apart from the very reasonable price, to make the potential buyer decide in favour of this specific dictionary? On the outer back page, five innovations are listed, of which two feature in the preface written by the editor. They are the use of the British National Corpus, "a massive and carefully balanced computer databank of modern written and spoken English" and the introduction of a 3 500 word vocabulary used for the writing of the definitia. "The words in the list (given in full as Appendix 10) were chosen principally according to their frequency in the language, as revealed by the corpus, ..." Apart from the British National Corpus, the Oxford American English Corpus was also used to update this dictionary. The fact that the macrostructure as well as the microstructure (a defining vocabulary of 3 500 words and over 90 000 examples) are corpus-based would be a major decisive influence in favour of this dictionary for this user.

If the user decides on a generating, and not merely a diagnostic aid, *OALD* would be your first choice, since the first promise listed on the outer back page, is that "this exciting new edition will help you: write and speak better English". These processes are assisted by a simplified defining vocabulary and "easy-to-understand definitions", coded clause patterns, idiomatic examples, a very explicit indication of labels, and an update of the pronunciation usage. Regarding the latter, one of the most outstanding features of the *OALD*, apart from the thorough update of the phonetics, is the indication of main and secondary stress in phrasal verbs and idioms, since this is always an extremely difficult area for learners.

For the writing of better English, two of the ten appendices should be singled out, namely the appendix on **Punctuation and writing**, where the use of punctuation in written English is shown, as well as the way in which a formal or informal letter is set out. The other appendix is **Notes on usage** where all the entries are listed at which notes can be found, e.g. on clarifying points of grammar, explaining differences between British and American usage, or showing variations between groups of words with similar meanings. The list not only gives the entry, but also the words or language points covered by each.

The diagnostic side however, is not neglected, since the second promise listed on the outer back page, is that *OALD* will help the user to "understand more easily and more quickly". Apart from the 65 000 definitions, which according to the editor are "both accurate and easy to understand", 1 700 words

are illustrated to complement their verbal definienda. Illustrations are effectively used mostly to discern between entries such as **shade** and **shadow**, **spade** and **shovel**, and **frog** and **toad**, or to disambiguate the definiens of entries such as **siphon**, **spider**, and **pillory**. The illustration is for instance only found at the entry of **frog**, but **toad** has a cross-reference to the illustration at **frog**.

The retrieving of information in *OALD* is explicated very thoroughly by means of different types of user information: on the front inner page the abbreviations, symbols, and labels, and on the back inner page, the phonetic symbols used in the dictionary are found. A very explicit key to dictionary entries, based upon real dictionary entries, is given on p. vii, where each part used within these articles is indicated. In addition to that, pp. ix and x consist of extra information on "using your dictionary".

Typologically *OALD* is extremely user-friendly in that a running bottom consisting of the new verb coding scheme is applied: the users do not need to remember or figure out the coding system themselves. The text-layout itself contributes to the easier retrieving of information from an article: the various polysemous values of a lemma are separated by bold numerals, idioms and phrasal verbs are marked clearly by bold blocks, and the idiom or phrasal verb itself is printed in bold. The one segment that should stand out the most in articles of learner's dictionaries, namely the special notes on usage, does so in *OALD*: it is placed right at the end of an article between two horizontal lines. The note is given only at one lemma, but cross-references to that lemma occur in the articles of other lemmas mentioned in the note.

Idioms are listed alphabetically, and not according to search words as most other dictionaries do. This is semantically a very sound microstructural treatment of idioms since one lexical item within the idiom is not selected as being semantically the most salient. Derivatives and compounds of a lemma are distinguished within the article by the use of bold symbols, and they form subarticles within the main article with their own linguistic information, polysemous distinctions, examples and idioms. The compounds of the adjective **dead** e.g. each has its own article and each subarticle forms a new paragraph. The following sublemmas for **dead** are alphabetically listed as compounds: **dead-beat**, **dead beat**, **dead end**, **dead heat**, **dead letter**, **dead loss**, **dead weight**, **dead wood**.

Another macrostructural aspect that should be mentioned in helping the user to retrieve entries effortlessly, is that abbreviations are listed alphabetically, and not as in many dictionaries, as the first entries of a specific letter. If the abbreviation shows no resemblance to the full form, the abbreviation is repeated in the article of the full form, e.g. **lb** is listed between the lemmas **lazybones** and **lbw**, but it also occurs as part of the article of the lemma **pound**. The noncanonical forms of verbs e.g. are lemmatized alphabetically as well, although almost no information is found except for cross-references to the canonical form: "gave *pt* of GIVE¹."

The entries in general reflect "an up-to-date picture of today's English" with lemmas such as **desktop publishing**, **Filofax**, **laptop** and **notebook computer**, **windsurf**, **mountain bike**, and even **rave** (a large party for young people, with dancing to fast electronic music, ...). The lexical items **roller blade** and **internet** are now more conspicuous by their absence.

Semantic information is one of the most important microstructural components of dictionaries. *OALD* gives a first-rate treatment of lemmas regarding the explanation of meaning, the indication of synonymy and opposition, and the homonymy and polysemy. What should however be emphasized here, is the way in which the polysemy of semantically empty verbs like **do**, **go** and **make** is handled. Because these verbs are so difficult to define, many lexicographers list the various contexts in which these verbs operate randomly as their different polysemous values. In *OALD* the various contexts of these verbs are grouped under a thematic heading. The user could immediately go to a specific heading under which specific contexts are listed, and concentrate only on its definitia and exemplary material. The themes for **go** e.g. comprise **movement**, **position**, **activity**, **state**, **sound**, **coming to an end**, **commands**, and **other meanings**. This type of organizational defining could also be used for the defining of prepositions or other functional items.

The only aspect of *OALD* that left a slightly negative impression, is the strong encyclopedic character with the inclusion of maps, cultural information, amongst which features information on the Commonwealth and the legal system in the United States, as well as illustrations of flowers common in Britain and the United States (p. 452) and breeds of dog (p. 343). These facts and illustrative material definitely belong in the encyclopedic and not in the linguistic version of *OALD*.

All in all, *OALD* is a more than worthy member of Oxford's series of dictionaries and of the range of other learner's dictionaries. It will positively enhance the encoding and decoding abilities of students of modern English.

Ilse Feinauer
Department of Afrikaans and Dutch
University of Stellenbosch
South Africa

L.G. de Stadler en Amanda de Stadler (Medewerker). *Groot Tesourus van Afrikaans*, 1ste uitgawe, 1ste druk 1994, xxxiii + 863 pp. ISBN 1 86812 478 9. Halfweghuis: Southern Boekuitgewers. Prys R89,99.

Abstract: De Stadler's *Groot Tesourus van Afrikaans* is reviewed. Critical questions are examined such as: What is a thesaurus? What is the function of a thesaurus? Who is the typical user of a thesaurus? What is it good for? How should it be compiled? What should be included in it? Does *Groot Tesourus van Afrikaans* satisfy these criteria?

Keywords: SYNONYMS, ANTONYMS, HYPONYMS, SUPERORDINATES, LEXICOGRAPHY, THESAURUS, CONCEPT CATEGORIES, NEED DETERMINATION, USER-FRIENDLY, SEMANTIC STRUCTURE, USER GUIDE, COMPUTER TECHNOLOGY, WORD PROCESSOR, CONCEPTUAL FRAMEWORK, RACIST, INDEX

Opsomming: De Stadler se *Groot Tesourus van Afrikaans* word in oënskou geneem. Kwessies wat aan die bod gestel word, is vrae soos: Wat is 'n tesourus? Wat is die funksie van 'n tesourus? Wie is die tipiese tesourusgebruiker? Waarvoor is dit geskik? Hoe moet dit saamgestel word? Wat moet daarin opgeneem word? Voldoen *Groot Tesourus van Afrikaans* aan hierdie kriteria?

Sleutelwoorde: SINONIEME, ANTONIEME, HIPONIEME, SUPERORDINATE, LEKSIKOGRAFIE, TESOURUS, KONSEPKATEGORIEË, BEHOEFTEBEPALING, GEBRUIKERSVRIENDELIK, SEMANTIESE STRUKTUUR, GEBRUIKSLEIDING, REKENAARTEGNOLOGIE, WOORDVERWERKER, KONSEPTUELE RAAMWERK, RASSISTIES, INDEKS

0. Inleiding

Na die verskyning van Eksteen (1981) se *Afrikaanse Sinoniemwoordeboek met antonieme* en Hartevelt e.a. (1992) se *Woordkeusegids* is L.G de Stadler (1994) se *Groot Tesourus van Afrikaans* (saamgestel met die medewerking van Amanda de Stadler) vir Afrikaans leksikografies gesien 'n mylpaal in die sin dat dit die eerste volwaardige tesourus van Afrikaans is. Waar eersgenoemde twee publikasies inderwaarheid sinoniemwoordeboeke is, moet *Groot Tesourus van Afrikaans* getoets word aan die meer beperkte definisie van die begrip *tesourus*.

1. Tesourus gedefinieer

De Stadler (1994: xxviii) definieer 'n tesourus as

"... 'n woordeboek waarin die leksikale items gerangskik word met inagneming van hulle verhouding tot bepaalde konsepte of konsepkategorieë, maar ook met inagneming van hulle onderlinge betekenisverhoudings."

Hy wys daarop (bl. xxviii) dat "(d)ie tesourus ... 'n besondere plek in die woordeboekfamilie (het)."

1.1 Ordening van leksikografiese materiaal in tesourusse

'n Tesourus is 'n "woorde-boek", maar volgens De Stadler (1994: xxvii) met die verskil dat woordeboeke die leksikografiese materiaal van 'n taal in die reël alfabeties orden sonder dat aangedui word hoe hierdie leksikale items op grond van hulle betekenis saamhang.

Hy onderskei veral vier soorte verhoudings binne hierdie struktuur, naamlik:

1.1.1 Sinonimie, die verhouding tussen twee woorde waarvan die betekenis min of meer dieselfde is, soos in die geval van *gierigheid*, *behaagsug* en *hebsug*.

1.1.2 Teenoorgesteldheid, die verhouding wat bestaan tussen twee woorde waarvan die betekenis in opposisie tot mekaar staan, soos in die geval van *liefde* en *haat*.

1.1.3 Hiponimie, die verhouding tussen 'n superordinate term (oorkoepelende term) en daardie terme wat as subordinate (of hiponieme) dien. Dit is die verhouding wat bestaan tussen 'n woord soos *dier* en sy subordinate of hiponieme *bees*, *skaap*, *koedoe*, *springbok*, *leeu*, *renoster*, *olifant*, ens.

1.1.4 Die deel-geheel-verhouding (meronimie), 'n verhouding wat bestaan tussen woorde soos *voertuig* en *stuurwiel*.

En dit is juis hierdie verhoudings wat nie tot hulle reg kom in gewone een- en tweetalige woordeboeke nie.

2. Soorte tesourusse

De Stadler (1994: xxviii) onderskei die volgende soorte tesourusse:

2.1 Kumulatiewe tesourusse

Hierdie tesourus is 'n voorbeeld van 'n kumulatiewe tesourus. Vir 'n kumulatiewe tesourus om tot sy reg te kom en omdat hierdie soort tesourus betekenis nie verklaar nie, moet dit idealiter saam met 'n goeie verklarende woordeboek gebruik word. Hierdie beskouing geniet algemene konsensus.

2.2 Verklarende tesourusse

'n Verklarende tesourus verskil van 'n kumulatiewe tesourus in dié opsig dat daar betekenisverklarings vir die woorde of uitdrukkings aangebied word. Die nadeel van die verklarende tesouri is dat hulle om praktiese redes slegs 'n gedeelte van die woordeskat kan behandel. Indien die woordeskat van die betrokke taal in sy geheel gedek moet word, sou so 'n tesourus in omvang onhanteerbaar en selfverdelend word.

2.3 Vertalende tesourusse

In plaas van 'n betekenisbeskrywing verskaf die vertalende tesourus 'n vertaal-ekwivalent vir 'n woord binne sy konsepkategorie. Ook hierdie soort tesourus kan uiteraard ook net 'n gedeelte van die woordeskat behandel.

('n Praktiese uitvloeisel van die vertalende tesourus is die twee- of meertalige gidse wat aan toeriste beskikbaar gestel word.)

3. Waarna soek 'n woordeboekgebruiker in 'n tesourus?

De Stadler (1994: xxviii) sê in dié verband:

"In 'n tesourus word die struktuur van die woordeskat en die besondere verhoudings wat bestaan tussen die leksikale items in daardie woordeskat (in teenstelling met die gewone een- en tweetalige woordeboeke — CvS) egter wel op die voorgrond geplaas en moet die woordeboekgebruiker dus verwag om inligting te kry oor die sinonieme, teenoorgesteldes, hiponieme of superordinate van 'n bepaalde woord."

4. Wie is hierdie woordeboekgebruiker?

Dié vraag dring hom as 't ware aan ons op. Die vir die tesourussamesteller ontstellende antwoord daarop is dat die gebruiksfrekwensie van tesourusse (in teenstelling met verklarende en vertalende woordeboeke en woordelyste) baie laag is.

Daar dien in dié verband gelet te word op Alberts (1990, 1992) se werk in verband met behoeftebepaling in die leksikografie, naamlik om vas te stel wat die diskrepansies is tussen die heersende en verlangde situasie (vgl. Alberts 1992: 20).

'n Beperkte rondvraag onder taalbedrywige akademici bevestig die persepsie dat tesourusse slegs by rare uitsondering en dan enkel deur taalkundige fynproewers en taalliefhebbers geraadpleeg word. Die behoefte aan en die

gebruik van 'n tesourus vooronderstel 'n gesofistikeerde moedertaalkennis van die semantiese struktuur van die betrokke taal. Die gebruiksfrekwensie onder die lekepubliek sal dus om dié rede in alle waarskynlikheid baie laag wees.

'n Rondvraag onder boekhandelaars het aan die lig gebring dat, benewens instansies soos biblioteke, tesourusse (in teenstelling met verklarende en vertalende woordeboeke) gekoop word deur 'n opvallend groot aantal persone wat nie taalgeleerdes of taalpraktisyns is nie, maar gewoonweg taal- en blokraaisel-liefhebbers is of persone wat hulle om 'n verskeidenheid redes ter wille van die verstrooiing met taalkwessies doenig hou.

'n Persoon se gesofistikeerde moedertaalkennis van die semantiese struktuur van 'n taal se talige intuïesies sal hom feilloos na die regte woordkeuse (sinoniem, antoniem, hponiem of superordinaat) lei. Dit is egter nie waar van 'n persoon met 'n beperkte of gebrekkige insig in die semantiese struktuur van die betrokke taal nie.

5. Diskrepanisie tussen heersende en verlangde situasie

Wat uit hierdie persepsie voortvloei, is die vraag: Watter doel en nut het die ontsaglike hoeveelheid werk wat die saamstel van 'n tesourus soos *Groot Tesourus van Afrikaans* onderstel en hoe kan die diskrepanisie tussen die heersende en die verlangde situasie verklein of uitgeskakel word?

Laird (1980: vii) sê in dié verband:

"If we are to have good reference books, however, we must ask fundamental questions like the following: What is a thesaurus? Who uses it? What is it good for? How should it be made?"

6. Onderrig in die gebruik van tesourusse

Waarop die voorafgaande afstuur, is die wenslikheid, nee, noodsaaklikheid dat die bestaan en waardevolle gebruike van die tesourus as besondere soort woordeboek dieselfde gebruiksfrekwensie verkry as verklarende en vertalende (tweetalige) woordeboeke. Verder moet die woordeboekgebruiker (kleintyd) geskool word in die woordeboekkultuur en in die gebruik van die tesourus. Dit sal die voordeel inhou dat 'n groter woordsensitiwiteit en woord- en taalbewustheid ontwikkel word by die taalpraktisyn (in vakverband) sowel as by die veel groter lekepubliek (in alledaagse taalgebruik). Hierdie argument kom daarop neer dat Alberts se genoemde diskrepanisie tussen die heersende en verlangde situasie wat betref die algemene gebruik van een- en tweetalige en verklarende woordeboeke (in teenstelling met die seldsame gebruik van tesourusse) verklein of uitgewis word. Dit veronderstel die dwingende vereiste dat veel meer gedoen moet word om studente in die gebruik en waarde van tesourusse op skool en aan universiteite en teknikons te onderrig.

7. Jan en alleman en die tesourus

Mairé Weir Kay (1976: 9a) verduidelik dat die voortbestaan van Roget se *The Thesaurus of English Words and Phrases* toegeskryf kan word aan die feit

"... that in spite of its complexity and elaborate structural plan the sheer wealth of relevant material that it offers a *sophisticated* (kursivering van my — CvS) user is the explanation for its survival."

Daarteenoor stel Laird (1980: vii) dit dat 'n tesourus 'n praktiese boek is en niks te make het met die klassifikasie van woorde asof hulle botaniese spesimens is nie. Dit is veronderstel, sê hy, om aan die skrywer of spreker spesifieke hulp wat hy voel hy nodig het, te verleen. En, sê hy verder:

"(a) thesaurus should be developed to satisfy the needs of the user, not to elucidate the nature of synonymy ... the editor of such a work ... should ask himself, "When the speaker or writer rejects this word, what word or combination of words is he likely to want?"

Die onvermydelike implikasie hier is dat die tesourus (in teenstelling met Roget se tesourus wat, in die woorde van Kay, bedoel is vir "gesofistikeerde" gebruikers en dus buite die bereik van Jan en alleman is) eintlik bedoel moet word vir Jan en alleman. En dit sal alleen kan geskied indien die "gewone man op straat" toereikend onderrig is in die gebruik van die tesourus.

8. Wat hoort in die tesourus opgeneem te word?

Die vraag is dan: Hoe toeganklik en gebruiksvriendelik is *Groot Tesourus* vir Jan en alleman? Is dit ingestel op Jan en alleman of op die ingewyde, gesofistikeerde woordeboekgebruiker?

Laird (1980: vii) sê:

"A thesaurus needs to enter a locution only if it is the kind or word or phrase that people recall readily. Nobody is troubled because he can think of no synonym for fanforan or latitant. Most people have never heard of these words and if they can think of fanforan they will have no trouble in calling up boaster, braggart, bully ..."

In die *Groot Tesourus* se indeks verskyn inskrywings soos *lateriet*, *bisbigliando*, *kollenchiem en mandragora* (om maar enkeles te noem) wat, volgens Laird, nie in 'n tesourus tuis hoort nie en gevoeglik in 'n verklarende of tegniese woordeboek tuis hoort en daar nageslaan kan word.

Sulke gespesialiseerde en geleende woorde vorm in die reël nie deel van

die ongesofistikeerde taalgebruiker se woordeskat nie en het 'n te lae gebruiksfrekwensie om nuttig te wees en kan gewoonlik in 'n verklarende woordeboek nageslaan word. 'n Tesourus hoef dus, wat dit betref, nie volledigheid na te streef nie. En dit geld *Groot Tesourus* van Afrikaans in gelyke mate.

Roget se tesourus maak hom, op die oog af, nie hieraan skuldig nie.

9. Funksies van die tesourus

Wat betref die funksies van die tesourus sê Dutch (1968: vii) met verwysing na Roget's *Thesaurus* dat hierdie tesourus

"... is a vocabulary on a large scale, categorized by topics. With such an arrangement it offers a choice of words to fit any given context. This, indeed, is its main function, and is the reason why it is so often alluded to as an aid in *composition*." (kursivering van my — CvS)

De Stadler (1994: xxviii) sê aansluitend dat

"... die belangrikste funksie van 'n tesourus is om die regte woord vir die regte konteks te suggereer."

of, soos Van Schalkwyk (1992) dit stel:

"Dit maak die woord los wat nog maar net op die tong lê. Dit stel die leerling, die student, die sekretaris, die politikus, die leraar, die invuller van blokkiesraaisels, die skrywer en elke ander taalgebruiker in staat om die korrekte woord op die regte plek in te span."

En dit is inderdaad waarop Laird aandring as hy sê:

"A thesaurus should be developed to satisfy the needs of the user ..."

10. Doelmatige gebruik van die tesourus

Dit lê voor die hand (Laird 1980: vii, De Stadler 1994: xxviii) dat die tesourus slegs ten volle tot sy reg kan kom indien dit (soos hierbo aangedui) saam met 'n goeie verklarende woordeboek gebruik word aangesien

- die tesourus betekenis gewoonlik nie uitdruklik verklaar nie en
- die gebruiker (veral die lekegebruiker) nie ten volle op die hoogte is van wat die tesourus hom bied en hoe om hierdie inligting te ontsluit nie.

11. Gebruiksleiding by die tesourus

As daarop aanspraak gemaak word dat die tesourus die tesourusgebruiker help om die regte woord vir die regte konteks te vind, soos Roget (en ander) dit stel, is dit belangrik vir die gebruik en verstaan van die tesourus dat die gebruiker, en dan veral die minder gesofistikeerde gebruiker, hom deeglik sal kan vergewis van die inleidende opmerkings en die gebruiksleiding deur die skrywer(s), m.a.w. die woordeboek- en tesourusgebruiker moet presies kan vasstel wat hy in die betrokke woordeboek of tesourus sal kan vind en hoe om dit te vind. Min woordeboekgebruikers (en dit geld in veel groter mate die tesourusgebruiker) getroos hulle die moeite om hulle op die hoogte te stel van die gebruiksaanwysings in woordeboeke of tesourusse. 'n Intieme kennis van die nut en gebruike van die tesourus sal moontlik bydra tot 'n verhoging in die gebruik daarvan — 'n taak wat aan die deur van die skool en teknikon of universiteit se onderrig in die leksikografie gelê kan word.

12. Tesourusse en die rekenaar

'n Gedagte wat hieruit voortvloei, is die rol wat die moderne rekenaartegnologie, insonderheid die woordverwerker, gespeel het in die bewusmaking van die bestaan en gebruik van woordeboeke en tesourusse. Die gebruik van die tesourus word as 't ware aan die gebruiker van die woordverwerker opgedring, en waar dit gewoonlik heelwat lyflike ongerief meebring indien daar tydrowend na woorde en woordbetekenisse in die gedrukte medium gesoek moet word, het die rekenaartegnologie dit vir die rekenaargeletterde moontlik gemaak om die benodigde inligting met eenvingerbedrewendheid te ontsluit. Wat daaruit voortvloei, is die opvallende feit dat — gesien die feit dat *homo grammaticus* al hoe meer rekenaargeletterd raak — die saamstel van tesourusse en woordeboeke op hierdie mark ingestel moet word en dat gebruikers van die woordverwerker bewus gemaak word van hierdie waardevolle hulpmiddel en geheuehulp. Dit het vanselfsprekende praktiese implikasies vir die leksikografie en die onderrig daarvan op skool en aan tersiêre inrigtings soos teknikons en universiteite.

Dat De Stadler ook die elektroniese mark met *Groot Tesourus* betree het, belooft veel vir die hierdie reusetree in die Afrikaanse leksikografie.

13. Die organisasie en gebruiksleiding van *Groot Tesourus van Afrikaans*

Groot Tesourus van Afrikaans kan in die lig van die opmerkings wat hoër op gemaak is, beoordeel word. Aangaande die onkunde rakende die bestaan en gebruik van tesourusse, is dit noodsaaklik dat die organisasie en gebruiksleiding van die tesourus vir ook die leek (deursigtig) duidelik moet wees.

In sy inleidende verduideliking van die aard en gebruik van 'n tesourus op bl. xxvii sê De Stadler:

"(Gebruiksleiding word aangebied in paragraaf 4 op bl. ix.)"

Geen sodanige inligting verskyn op bl. ix nie, maar wel op bl. xxx. Hierdie soort onnoukeurigheid is (ook en veral vir die leek) nie net verwarrend en irriterend nie, maar loop tydrowend uit op 'n gesoek na die inligting waarna verwys word — 'n faktor wat die potensiele tesourusgebruiker 'n renons in tesourusse kan laat kry.

Wat betref die soorte verhoudings wat binne die struktuur van die leksikon onderskei kan word (bl. xxvii) en die organisasie en gebruiksleiding van 'n kumulatiewe tesourus soos hierdie (bl. xxx), voldoen *Groot Tesourus van Afrikaans* vormlik aan hierdie vereiste, naamlik dat dit gewoonlik drie opvallende dele bevat:

13.1 'n Uiteensetting (klassifikasieplan? — CvS) van die konseptuele raamwerk waarin die gebruiker 'n opsomming kry van die verskillende konsepkategorieë wat in die tesourus behandel word. Hierdie uiteensetting dien dan ook as 'n inhoudsopgawe van hierdie tesourus. Boaan elke bladsy word, naas die bladsynommer, ook 'n verwysing gegee na die kategorie of kategorieë wat op daardie bladsy voorkom sodat die gebruiker maklik 'n kategorie kan naslaan. Dan word daar ook 'n verkorte uiteensetting (inhoud? — CvS) van die konseptuele raamwerk aangebied om 'n oorsig oor die sisteem te gee — aldus De Stadler.

13.2 Die tweede deel van die woordeboek, wat De Stadler die "hart van die woordeboek" beskou, is die tesourusgedeelte self, waar die woordeskat onder bepaalde konsephoofde of die tesourusartikels georden word.

Die woorde of uitdrukkings in 'n tesourusartikel word (soos gebruiklik) woordsoortlik ingedeel. As die konsep waarom dit gaan as 'n selfstandige (naamwoordelike) konsep aangevoel word, word die selfstandige naamwoorde eerste aangebied, gevolg deur byvoeglike naamwoorde, werkwoorde, ens. Die selfstandige naamwoorde, werkwoorde, byvoeglike naamwoorde en bywoorde kan beskou word as die hoofkategorieë in die woordeskat.

13.2.1 Rangskikking van inskrywings

'n Kwessie wat hom tog aan die gebruiker van *Groot Tesourus* opdring en wat ook uitdruklik in die inleiding gestel kon word, is die basis waarop of die ordeningsprinsipe waarvolgens die "sinonieme" onder die woordsoortlike indeling soos hierbo uiteengesit, gerangskik is. Aangesien die "sinonieme" in *Groot Tesourus* nie alfabeties georden is nie, moes dit volgens 'n ander prinsipe of

heeltemal lukraak geskied het.

Laird (1980: viii) sê in dié verband in sy voorwoord tot *Collins New World Thesaurus*:

"... usually the commonest of a set of synonyms carries the main entry; uses of a word are arranged roughly by frequency of use, from common to uncommon, and the same sequence is maintained within the synonymies ..."

Laird verduidelik nie die grondslag waarop die gebruiksfrekwensie van sinonieme in *Collins* bepaal is nie.

Om konsensus oor so 'n kwessie, naamlik "the commonest of a set of synonyms" te verkry, is iets wat gedink maar nie gedoen kan word nie en sou 'n onderneming van bomenslike dimensies wees wat ver buite die bestek van die leksikograaf of die samesteller van 'n tesourus se opdrag lê. Dit kan De Stadler of enige tesourussamesteller gevolglik nie ten laste gelê word nie, maar dit kan nut hê om die kritiese tesourusgebruiker oor vrae van dié aard gerus te stel.

De Stadler (xxxi) wys daarop dat daar ook binne elke woordsoortartikel 'n besondere organisasie is. So 'n paragraaf kan 'n aantal kleiner paragrawe bevat wat betrekking het op die feit dat 'n bepaalde konsep semantiese onderverdelings het. Onder *Begeerte* kan onderskei word *begeerte*, *die vervulling van die begeerte*, *veroorsaking van die begeerte*, *die persoon wat begeer*. Hierdie subparagrawe word duidelik gemerk deur die eerste woorde *vet kursief* te druk.

Hierdie soort begeleiding kan alleen bydra tot 'n beter insig in die organisasie en die groter "bruikbaarheid" van die inligting in die tesourus.

13.2.2 Gebruiksaanduidings

In die tesourusartikel word daar dikwels inligting tussen hakies aangebied in die vorm van gebruiksaanduidings. Woorde soos *gatvol* s.v. paragraaf 775. *Weersin* word aangedui as (*plat*). Ander etikette wat gebruik word, is (*veroud.*), (*geselst.*), (*ong.*), (*kwetsend*) vir verouderd, geselstaal, ongewoon en kwetsend.

Hierdie inligting is, in die lig van sosiaal-politieke eise wat aan tydgenootlike leksikograwe deur politieke, godsdienstige en ander drukgroepe gestel word, 'n noodsaaklike vereiste omdat dit, met inagneming van die sensitiwiteit van die betrokke groepe, gebruikers waarsku teen, o.a., die kwetsende aard van sulke leksikale items of uitdrukkings.

13.3 Indeks

Die derde gedeelte van die tesourus is die indeks waarin 'n beduidende persentasie van die materiaal wat in die tesourusgedeelte voorkom, alfabeties gerang-

skik word met verwysing na die konsepkategorie waar die betrokke woord of uitdrukking aangetref kan word. Hierdie indeks is 'n belangrike hulpmiddel wanneer die gebruiker 'n bepaalde woord vinnig wil naslaan ten einde 'n sinoniem of ander aanverwante woord op te spoor.

13.3.1 Leemtes in die indeks

Sommige afleidings verskyn nie in die indeks nie. Hoewel dié feit genoem word, verduidelik De Stadler nie die rasionaal agter die besluit om bepaalde afleidings en samestellings in te sluit nie terwyl ander weggelaat word. Hoewel daar begrip is vir die feit dat die tesourus uiteraard onhanteerbaar groot sou geword het indien dit na volledigheid gestreef het, wil dit tog lyk of die besluit om bepaalde samestellings in die indeks in te sluit terwyl ander as inskrywings onder die kernwoord verskyn, enigsins willekeurig geneem is. *Saalknop* kom nie as aparte inskrywing voor nie, maar wel na daardie woorde wat verwys na die onderdeel van 'n saal, maar *wieldop* word wel in die indeks aangegee, maar verskyn nie onder *wiel* as onderdeel van 'n wiel nie. Om *saalknop* dus op te spoor gaan vir die oningeligte 'n tydrawende oefening wees.

Dit wil voorkom of die besluit watter samestellings in die indeks opgeneem moet word en watter nie, enigermate arbitrêr is — 'n oorweging wat vir iemand wat daarvan bewus word, enigsins hinderlik aandoen.

Dieselfde geld woorde wat as *rassisties* geëtiketteer sou kon word. In die indeks kom 'n woord soos *hottentot* voor wat in die artikel, tereg, as *kwetsend* geëtiketteer word. Verwante woorde soos *bantoe*, *boesman*, *kaffer*, *lokasie*, *meid*, *swartman*, *witman* kom egter nie in die indeks voor nie, terwyl 'n item soos *boer* wel in die indeks opgeneem is.

Die vraag wat hom hier aan die gebruiker van *Groot Tesourus* opdring, is of die samestellers hulle as gevolg van oorgevoeligheid / oorversigtigheid nie skuldig gemaak het aan die skets van 'n skewe beeld van die werklikheid nie. Die items kon, soos *hottentot*, as (*kwetsend*) of (*verouderd*) geëtiketteer word.

Hoe dit ook al sy, 'n verduideliking van die rede waarom sulke woorde weggelaat is uit die indeks sou tot groter deursigtigheid bygedra het.

13.4 Beknopte gebruikersgids

Die beknopte gebruikersgids op die voor- en agterbinneblad is nuttig en behoort die lekegebruiker van *Groot Tesourus* oor sy ergste probleme met die interpretasie van die gebruikersgids heen te help.

14. Algemeen

Groot Tesourus bevat (volgens die *Inleiding*) nagenoeg 120 000 inskrywings, maar op die agterblad word die getal aangegee as 150 000 wat in 855 konsep-

tuele kategorieë georden is. Dit dui op nalatige redigering.

Dit is (vir my — CvS) nie sonder meer duidelik waarom die verskil tussen De Stadler se 855 konseptuele kategorieë in teenstelling met Brouwers en Roget se 1000 konseptuele kategorieë noodwendig 'n plusfaktor is nie. Daar sou gere-deneer kon word dat hoe meer domeine afgebaken en hoe meer toegangs-poorte daar vir die gebruiker beskikbaar is, hoe bruikbaarder die tesourus behoort te wees.

15. Slotopmerking

Hierdie eerste uitgawe van *Groot Tesourus van Afrikaans* is, soos inleidend beskryf, 'n mylpaal op die gebied van die Afrikaanse leksikografie. Die feit dat De Stadler dit ook vir die elektroniese medium toeganklik gemaak het, kan net bydra tot die groter gebruik en wyer bekendstelling van hierdie waardevolle werk.

Verwysings

- Alberts, Mariëtta. 1990. *'n Bepaling van Afrikaanse vakleksikografiese behoeftes*. D.Litt. et Phil.-proefskrif. Pretoria: Universiteit van Suid-Afrika.
- Alberts, Mariëtta. 1992. Behoeftebepaling in die leksikografie. *Lexikos* 2: 1-27.
- Dutch, Robert A. 1968. *Roget's Thesaurus of English Words and Phrases*. Reprint. Harmondsworth: Penguin Books.
- Kay, Mairé Weir (Ed.). 1976. *Webster's Collegiate Thesaurus*. Springfield: G. & C. Merriam Company.
- Laird, Charlton. 1980. *Collins New World Thesaurus*. Reprint. London and Glasgow: Collins.
- Van Schalkwyk, D.J. 1992. Voorwoord. Hartevelt, P. (Red.), L.G. de Stadler en D.C. Hauptfleisch (Medewerkers). 1992. *Woordkeusegids. 'n Kerntesourus van Afrikaans*. Eerste uitgawe. Halfweghuis: Southern Boekuitgewers.

C. van Schalkwyk
Emeritusprofessor
Departement Linguistiek (Vertaalkunde)
Universiteit van Suid-Afrika
Pretoria
Suid-Afrika

Margaret Doyle. *the A—Z of Non-Sexist Language*, 1st edition 1995, 112 pp. ISBN 0 7043 4430 0. London: The Women's Press Ltd. Price £6,99.

The race and gender *insensitivity* of much common English usage was memorably expressed by the Zimbabwean writer Dambudzo Marechera:

For a black writer the language is very racist; you have to have harrowing fights and hair-raising panga duels with the language before you can make it do all that you want it to do. It is so for the feminists. English is very male. Hence feminist writers also adopt the same tactics (Veit-Wild 1992: 4).

Margaret Doyle, described as "a freelance writer and editor" in the book blurb, put together a small book called *the A-Z of Non-Sexist Language*. The trendy 'decapitalisation' of the first word of the title is not a misleading impression. No lexicographer myself, I am not convinced of the scholarly thoroughness of this publication, even though its earnestness is much in evidence. On p. 74 one is warned against using the term **prehistoric man**, presumably because it 'erases' the women of that period, but what the note tells one is that "*Prehistoric* is ambiguous as well as sexist (it can refer to any of a number of distinct time periods)" and then four lines down in the same entry Doyle lists the term "missing link" as one of the preferable alternatives! A similar impression of inconsistency and even incongruity is created when the reader is reassured that **female** is "Not a sexist term" because it "derives from *femella*, Latin for 'little woman'" (30). Surely the etymology suggests how ancient the (bad) habit of infantilizing women is. In the last sentence of this entry Doyle adds: "As a noun, though, in most contexts it retains its sense of non-humanness and is used mostly in relation to animals or statistics" (31). Here, too, Doyle sends mixed signals to the well-intentioned researcher after gender-sensitive language use.

To cite these examples is not to dismiss the worthiness of the undertaking in itself. Doyle's introduction makes entirely sensible points: that "sexist language is unclear and inaccurate, ... excludes more than half the population ... encourages destructive stereotypes" (3), also that "[u]sing inclusive language does not have to be clumsy" (2) and that "inclusive language [should not be] confused with euphemism" or caricatured "to discredit ... legitimate aspirations" (5). In the thesaurus-style listing of terms which follows, many entries make it (enlighteningly) clear how easy it is (where there's a will) to avoid bad (sexist) habits in English usage and contrarily reminds its readers how readily people slip into using such (avoidable) expressions. I cite one such entry:

career woman/girl There is no parallel for men. The term relies on a sexist distinction between 'real' work (i.e. paid and outside the home)

and home-based work (such as homemaking) — a distinction that is not made in relation to men. The term is ambiguous, as it is also sometimes used to distinguish between 'professional' and 'unskilled' or low-paid employment. In both cases, the subtext is that a 'career woman' sacrifices family for job. If career is important in the context, be specific — e.g. accountant. Otherwise, avoid (16).

This entry includes and illustrates the point made elsewhere — that in many supposed gender parallels women are called "girls" and men "men". Valid, too, are the warnings against using a word like "whore", crisply labelled "Offensive, and applied only to women" and the advice given as an "OPTION: [to use the word] prostitute (for male or female)" (106). As Doyle illustrates, sexuality tends to be used 'against' women and 'for' men. She writes: "*Nymphomania*, for example, is widely used ... The masculine 'equivalent', *satyriasis*, is rarely, if ever, heard" (6 — see also her "Topic Note" on "Sex Words", 94-95). As a general rule, she sensibly advises readers to "use the root form" (27) of career names, since "-ess and -ette suffixes" imply "that the female is somehow less important than the male" (26). But when Doyle, having warned readers against using "effeminate [as] A sexist term that associates femininity with weakness and passivity", gives us "flowery" and "chintzy" (25) as two of the supposedly acceptable alternatives, her advice seems snobbish and pointless.

In some comments, Doyle seems quite unreliable to this reviewer. It may be true, as we read under the entry *frigid*, that this term is "A sexist and insulting label used for women who cannot achieve orgasm", but that "its male 'equivalent', *impotent*, does not carry any connotations of neurosis" (33) is decidedly false and (I think) sexist in its own way. Nor can I go along with Doyle's suggestion that a commonly used word like *headmistress* (for example) "suffers from the illicit and sexual connotations of MISTRESS as used today" (37)!

A general principle operative in Doyle's line of work seems to be the admirable ideal of avoiding type-casting gender roles through language. She quotes from a British manual *Guidelines on Countering Sexism in Schools* which says that the three main ways in which language can be sexist are "'by stereotyping females and males, by excluding women and girls, and by classing women and girls as inferior'" (99). This is quite right. How deft and elegant non-sexist language may be can be seen from Doyle's entry under "Scouts No longer called Boy Scouts; one percent are girls ... 'A Scout is a brother to all Scouts' has become 'A Scout belongs to the worldwide family of Scouts'. Similarly, 'A Scout has respect for himself and for others' has become 'A Scout has self-respect and respect for others'" (89). I don't find the recognition which Doyle extends to a ludicrous alternative ("*ovarimony*"!) to the word *testimony* (having made the point that the term is "not sexist"), appropriate (96-97). Such uneven standards are to my mind one of the chief weaknesses of this book. Doyle does attempt to be consistent in advising the avoidance of *all* gender-

specific terms and her 'neutering' advice would have us cut out both **fatherland** (30) and **motherland** (61). The question is, of course, whether some of these practices would not merely impoverish and 'blandify' the English language — or, of course, make it (in *some* cases, I hasten to add) imprecise or confusing. Doyle's judgement that among the "new forms of address" *Ms* is "[t]he enduring invention" (39) is surely valid, though.

I like the way in which some of the entries in the A-Z listing give one a sense of the large range of (perhaps far more precise) alternatives for a term deemed sexist. Hence, as "Options" (i.e. possible replacements) for **machismo**, "courage, strength, mettle, bravado, muscle, pride, swagger, self-confidence, over-confidence, aggressiveness, potency" (51) are listed. Some other entries exhibit a paucity in the face of the rich resources of the English language, though. If we are to be cautioned against the sexism implicit in the expression **brotherly love**, "familial love, affection, platonic love" (15) seem to me inadequate and even inaccurate as alternatives — expressions that come to mind are humanity; humaneness; benevolence; generosity; kindness; charity [the New Testamental term]; a sense of communion; concern [for others]; care; community consciousness; social conscience; universal love; public spiritedness; a caring disposition.

In conclusion it needs pointing out that Doyle indicates her awareness of the point "that discrimination in society will not change simply by ridding our language of sexism" (2-3) and that using non-sexist language perhaps pays mere "lip-service to reform rather than addressing the very real problems of sexism in society, including discrimination, harassment, violence against women, and economic inequality" (3). Indeed. But then, one wonders whether excision of such a term as **abominable snowman** from the English language in favour of (Doyle's **OPTIONS**) "yeti [or] abominable snow creature" (9 — the second entry in her alphabetical listing) will do much to redress gender imbalance in the world at large.

Reference

Veit-Wild, Flora. 1992. *Dambudzo Marechera. A Source Book on his Life and Work.* Harare: University of Zimbabwe Publications.

Annie Gagiano
Department of English
University of Stellenbosch
South Africa

I.E. du Toit. *Tweetalige Polisiewoordeboek. Bilingual Police Dictionary*, 1ste uitgawe / 1st edition 1994, 469 pp. ISBN 0 627 01903 X. Pretoria: JL van Schaik. Prys / Price R57,50.

Wat die struktuur en algemene voorkoms betref, maak die *Tweetalige Polisiewoordeboek* 'n mooi indruk. Uniek en bruikbaar is die lys van Dissiplinêre oortredings, Latynse regsterme en Wette wat bygevoeg is.

'n Mens sou die vertaling van *Dissiplinêre oortredings* as die meeromvattende en moontlik andersbetekenende *Offences against duty and discipline* kon bevreemte. By die alfabetiese lys van Wette ontbreek die nommering wat die naslaan van wette sou vergemaklik het. Verder loop die outeur die risiko dat van die Wette reeds verouderd (herroep) of in die proses om gewysig te word, is.

Essensieel gaan enige beoordeling om die vraag of dié woordeboek wel woordeboek in die ware sin van die woord is en verder of dit 'n *Polisiewoordeboek* is. Dan eers sou 'n mens kon vra of dit 'n suksesvolle *Polisiewoordeboek* is.

Die outeur stel kortweg die hoofdoel van die woordeboek: "om die leemte wat op hierdie gebied bestaan, aan te vul deur gepaste woorde en uitdrukkings te voorsien vir 'n polisie-omgewing". Wat die outeur met **gepaste** bedoel is onduidelik. Waarskynlik dit wat op die polisie-omgewing toepaslik is — of is dit woorde en uitdrukkings wat "ordentlik" is? Dit is bekend dat die SA Polisie op 'n gereelde grondslag 'n lys van sogenaamde weersinwekkende terme publiseer; woorde waarvan die gebruik daarvan deur 'n polisiebeampte hom / haar die gramskap van die burokrasie op die hals sal haal. Terme soos **klopjag** (i.p.v. huissoeking) en **vangwa** (i.p.v. patrollievoertuig) word as "weersinwekkend" gepersepsueer. Dit is nogal ironies omdat die woordeboek genoemde weersinwekkende terme bevat maar nie die sogenaamde gepaste alternatiewe nie.

Hoe dit ookal sy, groot gedeeltes van die woordeboek kan nuttig in die administratiewe omgewing van die polisie aangewend word. Die operasionele omgewing stel egter ander vereistes — vereistes met betrekking tot **volledigheid, andersheid en bruikbaarheid**. Leksikografiese oorwegings verminder ten gunste van pragmatiese verwagtinge.

Wat **volledigheid** betref kan die tydsberekening van die verskyning van die woordeboek nie slegter gewees het nie. Verandering in die vorm van strukturele amalgamasie, die aanname van die gemeenskapspolisieringsideologie en 'n strewe na demilitarisering (om slegs 'n paar te noem) het duidelike littekens op die polisiewoordeskaf gelaat: sekere terme het *verouder* (die stam distrik soos in **polisiedistrik, distrikskommissaris**, ens. het in die vergetelheid geraak, rangbenamings soos **adjutant-offisier, luitenant, majoor, kolonel, brigadier** en al die variante van **generaal** bestaan nie meer nie) en nuwe terme is geskep: **area** (soos in **areakommissaris**), **gemeenskapspolisiëring**, die aanspreek- en gebruiksvariante van **kommissaris** (soos **nasionale kommissaris, provinsiale kommissaris** en **stasiekommissaris**). Die woordeboek bied ook

nie erkende (gepaste!) polisiëringsterme soos die volgende aan nie: **polisiediens** (i.p.v. **polisiemag**), **proaktiewe polisiëring**, **regstoepassing** (nader aan die kol as **wetstoepassing**), **skokiaan**, **sindikaatmisdaad**, **geldwassery** ("money laundering"), **wisselruitery** ("kite flying") en vele ander. Tussen-departementele terme soos **korrektiewe dienste** en **strafregspiegling** verskyn nie. Verdere lastighede wat op minder goeie redigering dui, is dat daar nie altyd streng alfabeties gewerk is nie en ook dat herhaling voorkom (vgl. p. 41 en p. 42 in dié verband).

Die woordeboek maak aanspraak op **andersheid** daarom is dit 'n *Polisiewoordeboek*. Hier gaan dit sekerlik nie slegs om 'n lys van erkende terme en uitdrukkings waarvan die gebruiksfrekwensie lyk of dit hoër is in die polisieberoepsomgewing nie. Seker iets meer, soos byvoorbeeld die beroepstaal. Barnes in sy studie na die Transvaalse tronktaal (1990: 16) wys daarop dat waar persone in werksverband bymekaar kom of onder ander omstandighede daaglik met mekaar kommunikeer (soos bv. in 'n gevangenis of 'n opleidingskollege) daar dikwels 'n eiesoortige vaktaal of beroepstaal ontstaan. Dit skep groepsidentiteit en bevorder kohesie. Beroepskultureel beskou, versterk die polisiebeampte deur sy / haar gebruik die gespesifiseerde aard van die beroepskultuur. Die taal vorm 'n integrerende deel van die beroepskultuur en reflekteer dié kultuur se eie siening van solidariteit, relatiewe onafhanklikheid van die gemeenskap in die algemeen en 'n bepaalde houding teenoor gesag. Dit kom baie na aan Halliday (1976: 570) se sogenaamde anti-taal en Barnes (1990: 6) se misdadigertaal.

Alhoewel sommige van die terme en uitdrukkings sekerlik nie aan die vereistes soos gestel in die beleid vir die hantering van beledigende en sensitiewe leksikale items (Harteveld en Van Niekerk 1995: 232 e.v.) voldoen nie, is dit die taal van die straat en van operasionele polisiëring. Miskien lê die kern van die kritiek teen die *Polisiewoordeboek* daarin dat dit handel oor polisie eerder as oor polisiëring.

Terme en uitdrukkings mag nie altyd "gepas" wees nie, maar is kleurryk en interessant. Dit sou 'n ander dimensie en dinamika aan 'n polisiewoordeboek verleen het —

- "Bloues" en "stink-na-store" (nuwelinge)
- "Hoofkantoor", "Pretoria" en die "Engelse" (administratiewe gesag)
- "Likkewane", "stasieblompotte" en "bloubaadjies" (ander departemente)
- "haasbout (pistool / rewolwer); "pompie" (haelgeweer)
- "witduiwe" (onderhandelaars)
- "spykerbekke" (polisiehonde)
- "speakeasy" (kroeg — uit die Amerikaanse onderwêreld)

Die gevolgtrekking is dat die *Tweetalige Polisiewoordeboek* nóg volledig nóg anders is. Dit het 'n negatiewe invloed op die bruikbaarheid van die boek.

Bibliografie

- Barnes, L.A. 1990. Transvaalse tronktaal: 'n studie van 'n anti-taal. *Acta Criminologica* 3(1): 16-20.
- Hagen, G. 1995. Polisietaal as simbool van die polisie-subkultuur. *Praetor Vaktydskrif* 5(3): 5-8.
- Halliday, M.A.K. 1976. Anti-Languages. *American Anthropologist* 78(3): 570-584.
- Harteveld, P. en A.E van Niekerk. 1995. Beleid vir die hantering van beledigende en sensitiewe leksikale items in die *Woordeboek van die Afrikaanse Taal*. *Lexikos* 5: 232-248.

B.F. Smit
Departement Kriminologie
Universiteit van Suid-Afrika
Pretoria
Suid-Afrika

Longman Dictionary of Contemporary English, 3rd edition 1995, xxii + 1668 + B18 pp. ISBN 0-582-23750-5. Harlow: Longman Group Ltd. Price R49,95.

This is the third edition of the *Longman Dictionary of Contemporary English*, the first having appeared in 1978. The project continues to be guided by the sagacious, practical and indefatigable Professor Sir Randolph Quirk, who visited South Africa recently. He chairs the British National Corpus, as he did the previous survey of English Usage based at University College, London, and he points out in his Preface that the Spoken English Corpus "has been put to prominent use for the first time in this new edition".

The first edition introduced a number of valuable innovations designed to assist the learner of the language, since it was "intended primarily for the foreign student". Most significant among these was the usage in definitions of a core vocabulary of approximately 2 000 words based on frequency of usage, implementing the axiomatic principle that "the definitions are always written using simpler terms than the words they describe". It also used diagrams to illustrate technical items, such as the vocabulary of the castle, the laboratory and aircraft, as well as the varieties of fruit, birds, animals, etc. It had a fairly formidable introductory guide of 39 pages explaining the conventions of the typeface, abbreviations, pronunciation and a potted grammar. In all it consisted of 1303 pages containing 55 000 entries.

This third edition of the *LDOCE* (as it calls itself) uses the same devices and follows the same pattern, but the introduction is simplified and made more user-friendly, as the opening sentence (from the Director of Dictionaries) shows: "Welcome to the third, completely new, edition of the Longman Dictionary of Contemporary English!". The introduction is reduced to 22 pages, with certain sections moved to the Appendix, together with the core vocabulary and a few other items of general information. Overall the work has expanded to 1668 pages.

In keeping with the general shift in nomenclature, *syntax* is now used for grammatical inflections (which was the preferred term in the first edition). This change could be problematic for students who have learnt the terminology of grammar and syntax in their traditional senses. The grammar is now divided between the Introduction and the Appendix. The Introduction has sections headed "Syntax — verbs", "Syntax — nouns", etc., whereas the Appendix deals with "word formation", "irregular verbs" and so on, an arrangement which could also be confusing.

There are, as promised, various new features. The visual aspect has been greatly expanded. Now over 2 300 words are illustrated, and there are 24 pages in full colour, illustrating various categories. Some of these presuppose a fairly low level of competence: one has twenty pictorial representations of various sounds. For instance, "crunch" is exemplified by a girl crunching into an apple, "crackle" by a brightly burning fire, and "splash" by someone diving

into water. Such a strategy is only partly helpful, and could be misleading, since it necessarily can show only one example of the particular sound.

Other colour pictures show direction and position, varieties of some notion (e.g. "broken"), sporting terms from American football, baseball and basketball, soccer and golf, patterns and fabrics, the gadgets of the modern affluent kitchen, British and American terms for the components of the car/automobile and the varieties of fruit and vegetables (but not cuts of meat).

These examples demonstrate the stress on contemporary usage, which as Saussure reminded us, is the primary reality for most users of the language, and the dominant one for most learners. The obverse of this emphasis is that there are no etymologies, which naturally saddens lexicographers and academics, but is a necessary sacrifice.

Frequencies of usage in spoken and written usage are shown by means of graphs based on the British National Corpus and the Longman Lancaster Corpus. Many provide interesting information, for example, that setting *rubbish*, *garbage* and *trash* against each other. (The source-base for the American usage is not mentioned.) In addition, some 3 000 individual words are graded on a frequency scale of 1 to 3.

Pronunciation policy for the *LDOCE* has traditionally favoured the IPA system in a fairly narrow form, even to the extent of using double symbols to show, for example, variations in the last vowel sounded in *desperate*. As is well known, this alphabet can present a formidable challenge for learners, although the long-term benefits are great. Spelling accommodates American forms to a greater extent than is usual in standard English dictionaries, revealing the shift in the balance of power between British and American varieties.

Definitions, which form the heart of a dictionary, also have several innovations to help the learner. Key words are marked off by different formatting and the various principal meanings and idiomatic uses are covered within the boundary by means of signposts and menus. This is obviously preferable to the traditional practice of setting out the meanings seriatim. As Lord Quirk trenchantly insists in his Preface, "It will not do to regard language as comprising simply a host of separate items called 'words'".

Obviously the use of the core vocabulary has great practical advantages, but it can inhibit definition and occasionally make for undue simplicity. Thus a horse is "a large strong animal that people ride on and use for pulling heavy things", while irony is no more than "the use of words that are the opposite of what you really mean, in order to be amusing or to show that you are annoyed". There is no reference to *dramatic* or *Socratic irony*, as in the *COD*. However, the lexicographers have sensibly gone for a policy of copious instances illustrating usages. There are also extensive and helpful usage notes for problematic terms, such as *negro*.

Overall, the *Longman Dictionary of Contemporary English* is a useful, practical and appealing guide for the student of the language. For it provides not

only atomised information, but a good sense of how the building blocks of the language fit together to produce meanings.

Geoffrey Hughes
Department of English
University of the Witwatersrand
Johannesburg
South Africa

N.E. Osselton. *Chosen Words. Past and Present Problems for Dictionary Makers*, 1st edition 1995, x + 188 pp. ISBN 0 85989 419 3. Exeter: University of Exeter Press. Price £27,50.

Introductory remarks

This book is an informative collection of sixteen essays by one of the pioneers of modern-day academic Lexicography. It is the product of over forty years' scholarly work on prominent lexicographers of the seventeenth and eighteenth centuries like Cawdrey, Kersey, Bailey, Johnson and others. With the exception of two of the essays all the others have already been published before in various journals and books. They cover a wide variety of topics but still the author, a founder member and later President of The European Association for Lexicography, has managed to achieve a remarkable thematic cohesion. Although the original version of some of these essays were published as early as 1963 and 1964, they still are relevant to the current lexicographic discussions. This can already be predicted by merely looking at the subtitle of this book: *Past and Present Problems for Dictionary Makers*. In the *Preface* the author identifies the common theme running through the essays as that of a shared past in dictionary making. The choices confronting a modern compiler of dictionaries often worried early dictionary makers too. Although this book does focus on the practical problems of lexicographers it has a much wider scope than the lexicographic practice because this collection of essays also has to be regarded as a substantial metalexical contribution.

A remarkable aspect of this book is the way in which it displays the growth and development of lexicography as an academic industry and a fully-fledged discipline within the field of linguistics. While the original research for some of these essays was done in the 1950's when there was little scholarly interest in dictionaries in America and Europe, the work on some of the latest contributions was done as part of a flourishing focus on dictionary research.

The sixteen essays in this book are the products of research on closely related topics. Although, as a separate chapter, each essay forms an integral part of the book, it may still be read — and enjoyed — as an independent study. Immediately following the title of each chapter, the original date of publication is given. An added value of this collection is the *End notes* to chapters in which references are provided to later work in the same field. For the scholar using this book these references enhance the academic value of this collection.

Thematic cohesion in *Chosen Words*

A collection of essays written by one author can easily have an arbitrary character which leaves a reader unfulfilled. In the *Preface* to *Chosen Words* Osselton

says the bringing together of articles written over a long period of years has inevitably resulted in some duplication of material. This duplication has to be regarded as a positive aspect because it strengthens the thematic cohesion of this book.

The first chapter deals with the character of the earliest English dictionaries. Chapters two to six focus on various aspects of the macrostructural selection: the inclusion and treatment of figurative words, common words, dialect words, old words and literary words. Chapters seven, eight and sixteen discuss bilingual dictionaries. The use of sources is a theme dealt with in chapters ten and fourteen, while spelling stimulates an interesting discussion of fixing versus codifying in chapter nine. The compilation of a sixteenth century English dictionary and alphabetisation in early dictionaries of English are the topics of chapters eleven and twelve respectively. An innovative typological shift towards bilingualized learners' dictionaries is not only something of the present era but has been performed in the eighteenth century. This is discussed in chapter thirteen. Dictionary criticism is the theme of chapter fifteen, and although the discussion in this chapter is directed at three historical dictionaries, the general principles introduced in this discussion have a far more general applicability.

In spite of a variety of topics covered in these essays, there is one theme running through all of them: the historic perspective on (English) Lexicography. However, the way in which Osselton deals with the history of dictionaries makes this book just as interesting for a synchronic approach to the study of dictionaries because of the comprehensive picture it portrays of the development of English dictionaries.

The present-day study of dictionaries can best be done within a theoretical framework. One such model is the metalexicographical approach of Wiegand. Metalexicography can be divided into four components, i.e. the *History of Lexicography*, a *General Theory of Lexicography*, *Research on Dictionary Use* and *Criticism of Dictionaries*. The coherence of *Chosen Words* can best be illustrated by the way in which this book contains elements of all these components of a metalexicographical model. The most salient feature of this book is its portrayal of Historical Lexicography. In its focus on older dictionaries and the development of lexicographic activities *Chosen Words* gives a clear indication of the importance research on dictionary use has for the compilation of any new dictionary. Although the "sociology of the dictionary user" and the so-called "user perspective" are terms used in the present lexicographic discussions, Osselton gives a thorough account of the role these notions have played in the compilation of the earliest dictionaries. Today dictionary criticism does not always have a profound influence but quite often it does lead to substantial changes and improvement in dictionaries. Osselton's reflection on earlier criticism and its influence or lack thereof on specific dictionaries of the past, gives evidence of definite changes in the attitude of lexicographers towards linguistically motivated comments on their work. The formulation of a General Theory of

Lexicography is an ideal and ambition many lexicographers and linguists share today. Such a theory has to be based on sound principles. In more than one of the essays Osselton refers to lexicographic principles applied or not applied in the compilation of various dictionaries of the past.

The way in which the essays in this publication can be categorised within the different components of the Wiegand model emphasises the validity of this model but also the theoretical basis of Osselton's work. The need lexicographers have for an underlying set of theoretical principles is illustrated quite explicitly in this book by Osselton's sound interpretation of the principles governing Samuel Johnson's work, which Johnson himself expressed in no uncertain terms. Already in the first chapter Osselton refers to Johnson's semantic discipline and the *Preface* of his dictionary, foreshadowed by his *Plan of a Dictionary of the English Language* (1747), as an exposition of lexicographic theory. This date, Osselton argues in chapter ten, may be taken as the starting point for the science of lexicography, in England at least.

Central themes

Although this book focuses primarily on older dictionaries and early lexicographers, the themes and problems identified in the different chapters are just as relevant to present-day dictionaries and lexicographers. A distinctive feature of this book is the reoccurrence of certain themes in different chapters and their interactive relation. Dictionaries have to present their users with real language and not with the lexicographers' subjective interpretation thereof. Therefore the sources used by lexicographers to collect their material is of major importance for the success of any dictionary. Osselton repeatedly emphasises this aspect of the lexicographer's work. In historical linguistics secondary documentation will be of little value. However, dictionaries often rely on this form of documentation. The validity of this procedure is questioned by Osselton in his discussion of its application in a dictionary like the *OED*.

The presentation of real language compels the lexicographer to inform the dictionary users of the stylistic values and taboos of certain lexical items. This can be done by means of labelling — another topic discussed in various chapters but always relevant to both the specific dictionary dealt with in the respective chapter and dictionaries in general. Labelling is a topic discussed, e.g. with regards to figurative use, old words and extensively in the chapter on dialect words. This chapter also contains an interesting typology of dialect entries.

Samuel Johnson

Samuel Johnson is a central character in this book. One of the most interesting chapters (chapter ten) deals with his treatment of phrasal verbs. From a survey

of historical evidence, Osselton claims, it appears that the modern treatment of these lexical items in English dictionaries has been Johnson's invention. In this chapter Osselton not only discusses Johnson's treatment of phrasal verbs but he also gives the reader an interpretation of some of the passages in Johnson's Preface in which he identifies some crucial points lexicographers are still concerned with. The idiomatic nature of phrasal verbs, the semantic unpredictability of some and the self-explanatory nature of others have definite implications for their lexicographic treatment. Johnson's comments on these problems are interpreted in a sound manner by Osselton.

In chapter nine some aspects of Johnson's spelling conventions are discussed. Although many people claim that Johnson attempted to fix the spelling in his dictionary, Osselton shows how he actually embarked on an attempt to codify a recognised spelling. Various influences on Johnson's selection of spelling conventions are presented in an intriguing essay, dealing with an extraordinary lexicographer.

In conclusion

This book, a title published in the series *Exeter Studies in Linguistics* with R.R.K. Hartmann as General Editor, is another valuable contribution to the field of Lexicography, emphasising the quality of research done by Osselton in the fields of Linguistics and Lexicography. *Chosen Words* confirms the status of Lexicography as a linguistic discipline in its own right, as well as the status of the University of Exeter, and especially its Dictionary Research Centre, as one of the leading international institutions for lexicographic research.

Chosen Words has to be regarded as a fine example of academic Lexicography. However, this book can and should be appreciated by a much wider audience than practising and theoretical lexicographers. Any person interested in historical linguistics will benefit from this work. This book will also appeal to a more popular audience. Dictionaries are the most important containers of linguistic information utilised by ordinary language users. *Chosen Words* will be a stimulating experience on their reading list.

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

Rajend Mesthrie. *A Lexicon of South African Indian English*, 1st edition 1992, xxix + 148 pp. ISBN 0 94883 10 6. Leeds: Peepal Tree Press. Price £7,95.

Introduction

Mesthrie's statement on p. i that this *Lexicon* is intended to supplement the 1980 edition of *A Dictionary of South African English* (ed. Jean Branford) with reference to South African Indian English, the second largest (sub)dialect of English in South Africa, locates the book under review in the context of the study of English as a world language. First published in 1992, this book arises out of research commenced in 1985, and documents the state of South African Indian English from the earliest acquisition of English by the South African Indians (basilect forms of English) with focus on their contribution to the language. Mesthrie notes that the variety *South African Indian English* is a valid development of the English language, in the same category as Indian English or Caribbean English. This study pertains to the language in KwaZulu-Natal.

The persistence of features of South African Indian English in the English usage of South African Indians, and the appearance of new lexical items underline the significance of Mesthrie's contribution while pointing to the need for continued work in this field.

Method and format

The writer enumerates the sources, including personal experience, student projects and interviews with speakers and written sources such as newspapers with an Indian readership. He does not mention sources such as dramatic performances at weddings etc., which also captured South African Indian English usage; however, the *Lexicon* includes items used in such performances.

The format employed facilitates the distinction between various different usages: items used by most speakers irrespective of ancestral language, items peculiar to subgroups of a particular ancestral language, idioms and phrases, grammar, slang, pronunciation and general South African English items in South African Indian English. In this way Mesthrie has succeeded in giving a precise background and description of the item or usage. Furthermore, labels such as *rare*, *historical*, *rural* and *basilectal* help to categorize the items, e.g.:

COLD-TOUCH n.

A cold, a touch of cold. Bas. (p. 11)

DOSH n.

A small container for holding snuff, usually made of tin. Rural areas.

Now rare. [Zulu *idosha*]

Etymologies from the relevant Indian substrate language are provided. These often refer to more than one substrate language. The introductory part of the book also provides abbreviations as well as symbols and spelling conventions.

South African Indian English

In an introductory essay "South African Indian English", Mesthrie discusses, *inter alia*, the genesis of South African Indian English from around 1878, and explains why this dialect is not a pidgin or creole but rather a creoloid (after Trudgill, Platt and Todd). Mesthrie asserts that South African Indian English is a "model for the second language English of more and more young Zulu speakers in parts of rural Natal" (p. xvii). While this attests to the widespread influence of this form of English, the writer also expresses concern about the survival of many items of this language, with the loss of domains of usage consequent upon religio-cultural changes.

I do not unreservedly share this concern, in view of the dramatic changes wrought by the process of democracy in South Africa. The social, cultural and political milieux which generate forms and types of language, and which dictated the need for the development of South African Indian English, in the present period promise to empower the various languages in South Africa. Cultural and ethnic identity, liberated from apartheid's negative connotations, engender pride in one's heritage. South African Indian English, as well as South African English will be enriched by lexical items arising from the sociopolitical milieu, e.g. the *biryani bunch* (Indian politicians entertaining apartheid leaders with biryani) and the *House of Deli-goats* (House of Delegates). More positive contributions will emerge from politics, music, arts, culture, literature, diplomacy and tourism. Moreover, as Mesthrie concedes, "the early history of South African Indian English has yet to be fully uncovered" (p. xxiii). To this, one could add that a more comprehensive study of this subject, important to the study of the contributions of the Diaspora Indians, as well as the development of the English language, awaits the scholar.

The Lexicon

Mesthrie has applied the same painstaking scientific skills to this *Lexicon* as is evident in his other works. The *Lexicon*, a valuable source book for students of languages and linguistics, is also an appealing, readable book for non-specialists and laypersons. The largest section of the *Lexicon* deals with South African Indian English used by most speakers (Indians) irrespective of their ancestral language. The sources of these items are various: Hindi, Gujarati, Arabic, Sanskrit English, Tamil, Telugu, Zulu. It is notable that words from Afrikaans are not included here since they more properly constitute part

of South African English (SAE).

In the 50 pages making up *Section Ia*, Mesthrie has collected an interesting range of lexical items which can be easily identified by "native" speakers of South African Indian English. The wide variety of sources, including Zulu, attests to the assimilative nature of language generally, and the linguistic openness and ingenuity of the Indian people. This has contributed to the enrichment of the English language in South Africa; and would probably add to the treasury of English with exposure of South African writing internationally. The items refer to various aspects of the Indians' lives in South Africa; from the politically oppressive (**AFFECTED AREA**, p. 1), cultural practices (**BACK SARI**, p. 2) to religion and various culinary entries. Items such as **BIG-SISTER** (p. 6) and **COUSIN-BROTHER** (p. 13) (as well as **COUSIN-BIG-FATHER**, **COUSIN-SISTER**, **COUSIN-BROTHER-IN-LAW**) point to both the inadequacy of the English language regarding precise terms of relationships, as well as the Indians' adherence to traditional relationships, with emphasis on affection, respect and the extended family. It is also interesting that many lexical items, with Indian substrates, are applied in a context peculiar to South Africa. The word **GADHRA** derives from Hindi/Urdu and became attached to the mature (but not dried) bean seeds in Natal. **Gadhra-beans** seems to be the popular description for the seeds of beans (like fresh peas, lima beans) and is used in many chain supermarkets in the province of KwaZulu-Natal. Colloquialisms, swear words and some quaint expressions such as **GOAT-PRAYERS** (p. 19) persist in the speech of the South African Indians.

The basilectal components of South African Indian English, prevalent amongst the different linguistic groups, reflect on the source languages while indicating the Indians' determination to bend the English language to their use. Many examples of the basilectal expressions have been listed, e.g. **FUTURE** (p. 18) for fiancé or fiancée, **GOING-TIME** (p. 20) and **INDEPENDENT** (p. 24) for haughty, aloof. A more assiduous search and documentation of the basilect of South African Indian English would yield interesting results.

Section Ib lists items "limited to certain subgroups, dependent upon ancestral languages" (Indian languages). While this categorization is generally valid, there has been much interlinguistic contact and words such as **AKKA**, **AMMA** (p. 51) and many others are used across the linguistic spectrum. A linguistically interesting feature is the seeming "Anglicization" of Hindi (Bhojpuri) verb roots, e.g. **bhuj** (roast) becomes **bhujo** in usages such as "let us *bhujo* some meat", extended to **bhujo-ing** in participial forms. In South African Indian English the **-o** ending is added to many Hindi/Bhojpuri/Urdu verbs. In Caribbean (Trinidad) English the verb is given an **-e** ending to make it part of the spoken English, e.g. *bele* (from *bel* = roll dough), e.g. let us *bele* the dough.

Section II notes some idioms and popular phrases, many of these humorous and witty, e.g. **TO HAVE A DOUBLE-ENGINE**, p. 87 (i.e. to have a lover in addition to a spouse). I am confident that this list can be rewardingly expanded.

Section III, headed "Grammar", details a variety of departures from Standard English, on the analogy of Indian languages, in which usages for emphasis, auxiliaries and a wide range of grammatical features are recorded. Many of these "aberrations" such as the use of NEVER (*I never see him*) are found in other groups as well. Likewise, double negation (*I never tell him nothing*) is not peculiar to South African Indian English.

Section IV lists slangs used in South African English and South African Indian English. The use of slang, though prevalent, is eschewed by mature speakers of South African Indian English, for the simple reason that it is *slang*.

The *Fifth Section* deals with the pronunciation of English words in a way peculiar to Indian speakers. In my opinion these words do not necessarily constitute "South African Indian English" in the true sense. These belong more appropriately to the domain of "uneducated" usage of English words, although they may appear among speakers of even Standard English. The word *maintenance* (p. 123), often pronounced as *maintainance* by Indians, is also pronounced in this way by other groups. The same applies to NUINSANCE. Other items are speech idiosyncrasies, e.g. AKS for ASK (p. 121).

The *Lexicon* ends with a list of words common to South African English and South African Indian English. These words belong to various domains of South African life - from "Book of Life" and "conditional exemption" to "bottle store", "butchery", "dagga" and "donga".

Conclusion

Mesthrie's *A Lexicon of South African Indian English*, described by the author as a supplement (and also complement) to the 1980 edition of *A Dictionary of South African English* (ed. Jean Branford) is a valuable and interesting endeavour in its own right. It is a valid contribution to the study of language and should appeal to students of linguistics, sociologists, anthropologists and cultural historians. The *Lexicon* also adds to the growing body of works on the contributions of the Indian South Africans.

Rambhajan Sitaram
Department of Indian Languages
University of Durban-Westville
Durban
South Africa

Terminology Committee for Social Work. Vaktaalkomitee vir Maatskaplike Werk. *New Dictionary of Social Work. Nuwe Woordeboek vir Maatskaplike Werk*, 1995, xxiii + 140 pp. ISBN 0-621-16908-0. Pretoria: Government Printer. Die Staatsdrukker. Price/Prys R45,60.

Volgens die inleiding van die *Nuwe Woordeboek vir Maatskaplike Werk* is die primêre doel om praktiserende maatskaplike werkers en dosente en studente in Maatskaplike Werk te voorsien van die betekenis van die mees algemene woorde en konsepte wat in maatskaplike werk gebruik word.

Die verandering en vernuwing op politieke, maatskaplike en ekonomiese terreine wat vir maatskaplike werk besondere betekenis inhou, word ook in die *Woordeboek* gereflekteer deur insluiting van relevante begrippe wat daaglik gebruik (gaan) word. Die terminologie van die HOP (Heropbou- en Ontwikkelingsprogram) wat vir maatskaplike werk van waarde is, word goed gehanteer. Die *Woordeboek* verduidelik nie slegs die betekenis van woorde nie, maar sal ook die woordeskat van maatskaplike werkers uitbrei. Dit verskaf die 691 terme wat die hoogste gebruiksfrekwensie het in die praktyk en opleiding van maatskaplike werkers.

In die breë gee die *Woordeboek* ook 'n aanduiding van die aspekte waarmee maatskaplike werk daaglik gemoeid is. Sodoende verskaf die *Woordeboek* aan die belangstellende en belanghebbende nie-maatskaplike werker 'n waardevolle blik op die terrein van maatskaplike werk.

Die *Woordeboek* bestaan uit twee gedeeltes (Engels en Afrikaans) wat vooraf gegaan word deur 'n afdeling met die opskrif "Verduidelikende Aantekeninge" wat die gebruik vergemaklik. Verder verskyn 'n lys van medewerkers, 'n bibliografie, asook 'n lys van die toepaslike wette vir gebruik in maatskaplike werk.

Die eerste gedeelte van die *Woordeboek* (p. 1-68) bevat die alfabeties gerangskikte Engelse terme en die Afrikaanse ekwivalente. Die tweede gedeelte bevat die alfabeties gerangskikte Afrikaanse terme en die Engelse ekwivalente. Wanneer die *Woordeboek* as 'n bronverwysing gebruik word, sal die bladsyverwysing problematies wees omdat die Engels en Afrikaanse afdelings elkeen met bladsy 1 begin.

Die verduideliking van sommige terme sal nie sonder meer algemeen aanvaar word deur maatskaplike werkers nie. So, byvoorbeeld, word in die term **case study: gevallestudie** (p. 8 Engels, p. 23 Afrikaans) klem geplaas op die vasstelling van oorsake vir bepaalde verskynsels. Gevallestudies word egter dikwels gebruik om, benewens oorsaaklikheid, ook die dinamiek van hulpverlening / terapie te illustreer of om die kompleksiteit van 'n geval te beskryf.

Die verduideliking van die term **clarification: klarifisering** (p. 10 Engels, p. 33 Afrikaans) is eng, want dié tegniek behels ook hulp aan 'n kliënt om duidelikheid te kry oor die verwantskap tussen oënskynlik nie-verbandhoudende gebeure of ervarings.

Clinical social work: kliniese maatskaplike werk (p. 10 Engels, p. 33

Afrikaans) word nie duidelik genoeg onderskei van **gevallewerk: casework** of **maatskaplike werk: social work** nie. Geoordeel na die verduideliking van die onderskeie terme wonder 'n mens of daar regtig bestaansreg is vir kliniese maatskaplike werk naas gevallewerk of maatskaplike werk.

Die woord "suksesvol" in die verduideliking van die term **coping: hantering** (p. 15 Engels, p. 25 Afrikaans) is hinderlik. Hantering van lewenseise of die aanpak van probleme verloop nie altyd "suksesvol" nie. Mense kry dit wel soms reg om probleme te kan verdra en daarmee saam te leef. Die woord "suksesvol" is te sterk gerig op 'n eindproduk wat in hierdie konteks daarop dui dat die probleem nie meer bestaan of relevant is nie.

Wat betref die term **private adoption: privaat aanneming** (p. 47 Engels, p. 50 Afrikaans) word 'n eensydige verduideliking verskaf. Wanneer die biologiese moeder en die voornemende aanneemouers met mekaar in verbinding gebring word waarna die kind vir aanneming geplaas word, is 'n 'openbaarmakende aanneming' ter sprake. Hierdie soort aanneming is beslis nie die enigste tipe privaat aanneming soos wat die *Woordeboek* die term verduidelik nie. Privaat aanneming behels ook dat die biologiese moeder en die voornemende aanneemouers nooit met mekaar in verbinding gebring word nie, maar die baba tog wel by die aanneemouers geplaas word. In hierdie geval word gepraat van 'n 'nie-openbaarmakende aanneming'. Die *Woordeboek* stuur dus 'n verkeerde boodskap, naamlik dat privaat aanneming slegs openbaarmakende aannemings behels.

Die Engelse en Afrikaanse verduideliking van die woord **private practitioner: privaat praktisyn** (p. 47 Engels, p. 51 Afrikaans) is nie korrek nie. Die Engelse beskrywing moet verkieslik aandui dat die maatskaplike werker **onafhanklik van 'n welsynorganisasie of die staat praktiseer**. Die Afrikaanse gebruik van die woord "alleen" in die beskrywing van die term **privaat praktisyn** is nie dieselfde as die woord "independently" in die Engelse verduideliking nie. Trouens, daar is talle maatskaplike werkers wat alleen praktiseer in die eenmanskantore van welsynorganisasies. Hulle werk egter nie onafhanklik van die organisasies nie.

Vervolgens word gewys op enkele terme wat direk verband hou met omvattende teorieë en terapeutiese modelle. Dit sou sinvol gewees het om die 'tuisste' van die terme ook aan te dui vir die gebruiker wat verder oor die term wil nalees, soos trouens die geval is met die terme **id**, **ego** en **superego**.

- **kondisionering** (p. 34 Afrikaans): Tuiste is behaviorisme
- **teenoordrag** (p. 63 Afrikaans): Tuiste is psigo-analise
- **oordrag** (p. 47 Afrikaans): Tuiste is psigo-analise

Die stelling in die Inleiding rondom die eklektiese aard van maatskaplikewerk-kennis (p. viii) word aanvaar, maar juis daarom behoort meer erkenning gegee te word aan die teorieë en modelle waaruit maatskaplike werk put. Die psigo-analise, behaviorisme, eksistensialisme, lewensmodel, geïntegreerde benade-

ring en andere is hier ter sprake.

Die gebruiker sal ook nie noodwendig weet wat bedoel word met die "eenheidsbenadering" as die tweede uitgangspunt van die woordeboekopstellers nie (p. viii). Dit behoort verduidelik te word.

Verreweg die meeste van die 691 terme groepeer hulle rondom hulpverlenende handeling (individu-, groep- of gemeenskapsgerig), gevolg deur beskrywings van terme met betrekking tot:

- Gedragseienskappe (problematies en nieproblematies)
- Welsynbegrippe (welsynbeleid, -beplanning, -finansiering en bestaans-beveiliging en armoede)
- Statutêre maatskaplike werk en regsaspekte
- Maatskaplikewerkbestuur
- Maatskaplikewerkopleiding
- Maatskaplikewerknavorsing
- Kinders
- Huwelik en gesin
- Gesondheid- en gestremdesorg
- Groeibevorderende menslike, interaksionele vaardighede (empatie, refleksie)
- Menslike ontwikkelingsfases
- Maatskaplikewerketiek en professionele gedrag
- Unieke / eiesoortige maatskaplikewerkhandelinge
- Besondere welsynsdienste
- Modelle en benaderings in maatskaplike werk

Die Bibliografie en lys van medewerkers beïndruk en daar is duidelik daarna gestreef om soveel moontlik belanghebbendes te betrek by die taak om die *Woordeboek* op te stel.

Maatskaplike werkers het nou 'n bron in die hand wat met entoesiasme ontvang, gebruik, verbeter en uitgebrei moet word om sodoende die profesie op dinamiese wyse te dien. Daarom is dit goed om te verneem dat die Departement Maatskaplike Werk van die Universiteit van Stellenbosch voortaan die *Woordeboek* op datum gaan hou. Sodoende word die sinvolle werk van die Vaktaalkomitee op 'n stewige toekoms pad geplaas.

W.F. van Delft
Departement Maatskaplike Werk
Universiteit van Suid-Afrika
Pretoria
Suid-Afrika

Della Thompson (Editor). *The Concise Oxford Dictionary of Current English*, 9th edition 1995, xxi + 1673 pp. ISBN 0-19-861319-9. Oxford: Clarendon Press. Price R125,00.

Looking for clues as to why *The Concise Oxford Dictionary* is useful to us in South Africa today, I found the following in one of the preface sections titled 'English over Fifteen Centuries':

'... It is very important that dictionaries should take account of English overseas, especially as it affects usage in Britain. The process is a strengthening and enriching one, and is the mark of a living and flourishing language.' (p. xii)

On the one hand, this statement is probably designed to explain the inclusion of colloquialisms in what is historically-speaking a reference work designed for a relatively exclusive discourse community. This statement places *The Concise Oxford Dictionary* firmly in its place as relevant worldwide. It acknowledges that English is a language owned by everyone who uses it, legitimizes to some extent the existence of varieties of English which hitherto has been a somewhat thorny issue for linguistic purists. On the other hand, the quote above emphasizes the importance of influences from elsewhere which is only spoken about as being significant in regard to how it affects usage of English in England itself. It nevertheless prompts the reader to investigate the nature and extent of linguistic accommodation vs. purism inherent in the current edition of *The Concise Oxford Dictionary*. It also raises questions of ideological orientation in dictionary compilation.

It is clearly stated on the inside flap of the dust jacket that this edition reflects 'the English language as it is written and spoken in its varied styles and international varieties'.

Another disclaimer statement in 'English over Fifteen Centuries' is: 'with usage constantly changing the distinction between 'right' and 'wrong' is sometimes difficult to establish.' (p. xii) It would be interesting to investigate the nature of the distinction between prescription and description in *The Concise Oxford Dictionary*. Does it lend any guidance at all on desirability of specific lexical items in specific discourse contexts?

The appearance of this publication is well-timed for the political transformation in South Africa at present and may stimulate much discussion in terms of appropriate forms of linguistic accommodation which have become common practice in South African educational institutions. Now more than ever, the debate of 'standard' vs. officially recognized varieties of English is rearing its head. While moving towards fully fledged multilingualism in South Africa and while it is not clear how 'standard' English will be defined in South Africa in the 90's, this dictionary serves as an essential source of reference. Clearly the editors view language corpus extension to include borrowings and foreign lex-

emes as positive rather than negative. This indicates to what degree it is 'normal' for communities and individuals world wide to use conformity or diversity as touchstones for asserting their degree of proximity or resistance to diverse communities and thereby to social groups.

An interesting comparison might be whether the new Afrikaans dictionaries have been equally accommodating in recent years in terms of inclusion of Anglicisms than they used to be in the 60's and 70's when concerted efforts were still made to replace each Anglicism with 'pure', 'correct' words, even if this meant having to coin them at a rapid rate. I dare say my English teachers from that era would be shocked and amazed to see that slang such as 'gonna' (an American version of 'going to') which was unclassifiable as 'real' language, now appears in *The Concise Oxford Dictionary*!

C. Jeffrey (1993) in 'Standards in South African English' in the *English Academy Review* 10:14-25 discusses ways of looking at new Englishes and explains how it was documented in 1872 that the new vocabulary (under the strong influence of Dutch) was said to be 'incorrect' but 'useful' and that in future perhaps a new language might emerge. The same statement was made by Njabulo Ndebele in 1987 in his famous keynote address at the annual conference of the English Academy. A new variety of English has since emerged and is commonly known as 'Black English'. It is currently probably the most researched variety of English in Southern Africa. Prominent linguists such as D. Gough and Q. Buthelezi have focused much of their attention on its emergence. Jeffrey's article also refers to a document dated December 1909 in which is stated that from 1872 onwards a general openness to new features was prevalent, but that there was disapproval of 'corruption' of English standards.

Some comparisons between this dictionary and some of its predecessors and this dictionary and the new *Dictionary of South African English* that has been so long in the making and has appeared this year, might be useful. The first *Concise Oxford* appeared in 1911, while the first *Dictionary of South African English* appeared in 1996. This may be compared to the history of English dictionaries generally, the first of which appeared in 1604. The current *Concise Oxford* contains 140 000 meanings while the 1992 *Oxford Dictionary* and the *Webster's Third New International Dictionary* have round about 500 000 entries each. The current *Concise Oxford* reflects 15 centuries of recorded English language while the *Dictionary of South African English* reflects items recorded in the last three centuries only.

If it is true that the *COD* only contains about 15-20% of lexical items that exist, and that entries are based on selections by editors and not necessarily by frequency of degree of usage, then the magnitude of the English language is truly overwhelming. With regard to discrepancies in types of lexemes recorded in different dictionaries of the English language, D. Chrystal (1995:119) in his *Encyclopedia of the English Language*, says:

'Discrepancies are usually caused by differing editorial emphasis. The *Oxford* has far more historical references and British dialect items than does the *Webster*, which in turn has far more local American items.'

Historical links between English and South Africa probably account for the inclusion of much South African vocabulary in the current *Concise Oxford*, where a *Webster's* would probably not contain that many South Africanisms.

A lexicon reflects aspects of vocabulary of a language, how the word is formed, its origin and shift of meaning over time, its current usage, how meanings relate to other meanings. Chrystal (1995:118) defines a lexeme as follows: 'A lexeme is a unit of lexical meaning, which exists regardless of any inflectional endings it may have or the number of words it may contain.' For the purposes of this review, I randomly chose certain words common in South African English and will describe how they are circumscribed in the *Concise Oxford*. The words are 'donga', 'Bushman', 'apartheid', 'rooinek', 'bastard' and 'linguicism'.

It is stated that 'donga' is a common word in Australia as well as in South Africa but that it has Nguni roots. This is rather puzzling as it is difficult to understand how, if it is related to a similar word in Nguni, it could possibly have been transported into Australian vocabulary, unless there is a similar foreign word with equivalent meaning.

The word 'Bushman' is constantly under contestation. There is a debate raging among academics, historians and linguists around the political correctness of the term, which is not reflected in *The Concise Oxford Dictionary*. This is over and above the fact that many groups exist who have been classified in the typical South African way under one umbrella term 'Bushman'. Some of the groups are the Khoi-Khoi, the Khoisan and the Griquas, all of whom have similar but differing cultures and languages.

According to the *Concise Oxford* the word 'apartheid' which is uncontestedly a homegrown South African term, has become a general term used worldwide in reference to discriminatory policies.

'Rooinek' and 'bastard' are both words with overwhelmingly derogatory as well as racist connotations. *The Concise Oxford Dictionary* defines 'rooinek' as "a British or English speaking South African" and does not mention the extent to which the word has persisted throughout decades of political shifts in South Africa.

In contrast to this, three uses of the word 'bastard' are listed, one of which is its use of depicting 'an unpleasant or despicable person'. What is noteworthy here is the way in which the derogatory connotation of 'bastard' is explained, while that of 'rooinek' is not. One could assign the omission of the mention of the derogatory nature of 'rooinek' to a lack of access to knowledge about the historical origin of the term and its current colloquial usage. What I am highlighting here is the vastness of the lexicographer's task in a world where the cultural politics of English worldwide are unfathomable.

'Linguicism', a term that refers to linguistic prejudice, is a word that has become quite widely spread and is sprinkled liberally in the subject literature on the relationship between linguistics and human rights in South Africa. The term is used liberally by N. Alexander and K. Heugh from the Project for Alternative Education in South Africa and also by T. Skuttnabb-Kangas, a prominent writer on multilingualism and disadvantage, based at the University of Roskilde in Denmark. Whether it is common worldwide is not clear to me, but it is not listed in *The Concise Oxford Dictionary*.

This dictionary presents itself as a useful reference work for all teachers, linguists, students, scholars and all others as a tool for gauging overlap between common usage locally and usage elsewhere. This seems a useful exercise while the aspirations of South Africans generally is to be able to participate in discourse communities globally. I am not arguing for a fixed standard as such, but for the value of having knowledge of standards that are known to be influential. Informed choices, in terms of lexical items and syntax, cannot be made in the absence of good reference material. *The Concise Oxford Dictionary* provides the user with a specific section at the back of the publication, titled 'Style Guide'. South African students often ask to be taught rules of English in the hope that their writing might improve. A reference work such as this provides an excellent source for self-teaching in this regard. As an Oxford publication it is internationally recognized as a 'yardstick' of good and appropriate style and a 'must' for every shelf. Computer language adds another layer of lexemes constantly changing which cannot possibly be included in print fast enough.

The extensive inclusion of loan words and foreign borrowing in this edition reflect the extent to which such borrowing and loaning of lexemes is regarded as a normal or natural process in the nineties.

It is stated in the preface that the *COD* has been reviewed 9 times, this means that it was reviewed roughly every 10 years. The *COD*'s of the future will probably be reference works on current English usage and will appear on the World Wide Web in a constant state of review rather than reference works on 'standard, pure or proper English' as they might have been seen to be in the past.

Liesel Hibbert
Department of English
University of the Western Cape
Bellville
South Africa

G.W.R. Tobias and B.H.C. Turvey. *English-Kwanyama Dictionary*, 9th reprint 1991, viii + 199 pp. ISBN 0 85494 382 X. Johannesburg: Witwatersrand University Press. Price R27,00.

This dictionary was first published in 1954.

A note on Oshikwanyama: The prefix *oshi-* has been affixed to the stem Kwanyama so as to make it meaningful. If used without a prefix it renders no meaning to the native speakers.

Oshikwanyama is one of the Oshiwambo dialects used as a medium of instruction from Grade 1-3 and taught as a subject from Grade 4-12, and also at the teacher training-colleges.

Due to the extensive use of this language, there is a dire need for reference materials. However, the lexicographic work done so far with regard to Oshikwanyama is in the embryonic stage. This book only caters for learners of Oshikwanyama as English is used as the source language. The compilers have explained the abbreviations used in the dictionary but the explanations are not translated into Oshikwanyama. Besides, the abbreviations used would only be understood by people with the necessary linguistic background. In my opinion, a dictionary does not have to be overloaded with linguistic information and, should this occur, it should be made as simple as possible so as to ease the understanding of the target users.

In their introduction to this dictionary, the compilers have provided information on how to use the dictionary. The information provided includes the pronunciation of sounds in Oshikwanyama and their English equivalents, e.g. *tj* pronounced as *ch* in church, and how entries are entered, e.g. parts of speech such as nouns, verbs and adjectives. For the verbs information has been provided on how they operate in moods such as indicative, infinitive and imperative. Besides this, the question of verbal extension, which occurs quite extensively in Oshikwanyama, has also been treated with insufficient care because only causative has been referred to in the introduction.

With regard to adjectives, the compilers have given the stem preceded by the corresponding subject concord.

The concord given is unnecessary, e.g. p. 2 *acute* = *sa hongua* (read: *sha hongua*). *Sha* as a subject concord applies only to one noun class, namely class 7.

To avoid this difficult situation and indeed a more confusing one if not clearly defined, the compilers should have given the table of noun classes in the introductory remarks and should have illustrated the omission of a subject concord by means of a tilde.

The dictionary also lacks contextualisation. This means most of the lexical items have not been used in sentences to illustrate their meaning. It is a well-known fact that many words are not self-explanatory, and a mere translation of a word in a target language might tempt the users to misuse such words.

In the final analysis, the orthography of Oshikwanyama leaves much to be

desired. It is based on the old way of writing. Hence, most of the sounds used, do not appear in the existing Oshikwanyama texts, e.g. *s* in *okuhasiva* (not to know): the *s*-sound has become *sh*. In the same manner the *n* in *omunú* p. 112 should become *nh*, without a diacritic. On the basis of the above discussion, the dictionary needs to be updated if it is to serve its purpose.

While much still has to be done to update the lexicographic work under review, it should be noted that this is a great pioneering contribution to Oshiwambo in general and Oshikwanyama in particular. The commendable efforts by Tobias and Turvey should therefore be complemented by innovative work by prospective lexicographers.

P.A. Mbenzi
Department of African Languages
University of Namibia
Windhoek
Namibia

Annelize van Rooyen. *Die Afrikaanse Naamboek*, 1ste uitgawe, 2de druk 1995, 121 pp. ISBN 1 874901 18 X. Kaapstad: Queilleries-Uitgewers. Prys R44,95.

1. Die boek is gerig op die populêre mark. Dit merk mens onder andere uit die feit dat daar deurentyd in die boek sprake is van meisie- en seunsname, nie van mans- en vrouename nie. Dit merk mens ook uit die feit dat die meisiesname die eerste sowat 70 bladsye beslaan en gevolg word deur sowat 50 bladsye wat aan seunsname gewy is. (Vir die leek: Afrikaanse vrouename IS baie meer as mansname.)

'n Akademiese graad van hoë wetenskaplikheid moet mens dus nie verwag nie. As mens dit in ag neem, is Annelize van Rooyen (voortaan AvR) se *Die Afrikaanse Naamboek* 'n baie verdienstelike aanwinst op die boekrak.

2. Die boek bevat sowat 700 leksikografiese artikels oor voornaamgroepe, georden onder 'n hoofvorm. In hierdie artikels — of as alfabeties geordende verwysings na hierdie artikels — word daar meer as 4 500 Afrikaanse voorname en "variante en afleidings" aangegee. Dit is 'n groot oes.

Buitendien is kassies van sowat 'n halfbladsy gewy aan elk van die volgende onderwerpe: die registrasie van 'n geboorte, aanneming, (geboorte) buite die grense van die land, naamskat uit die Oudheid, die doop, hoe om 'n naam te verander, Germane vereer die natuur, modename, skuilname, die Griekse patroon van naamgewing, professionele name, patriotiese name, wat is jou nommer?, eienaardige name, so Afrikaans as kan kom, die Romeinse mitologie, byname, wat beteken beroemdes se name, die Griekse mitologie, die Germaanse mitologie ('n hele bladsy lank), Stefanus gaan op reis, en name wat iets gedenk.

Hierdie gekaste inligting maak die boek uniek in Afrikaans. Dit is inligting wat nêrens anders byeen is in Afrikaans nie.

3.1 Daar is bedenkinge teen *Die Afrikaanse Naamboek* wat geopper moet word. Maar die bedenkinge moet teen die agtergrond van die eerste paragrawe van my resensie gesien word.

3.2 In die *Voorwoord* sê AvR die Afrikaanse naamskat kom hoofsaaklik uit vier bronne: "die Bybel, die Germaanse, die Griekse en die Romaanse naamskat." Gelukkig word dié "Romaanse" darem "Latyn" in die kassie op p. 26. Dit is naamlik nie die Romaanse naamskat nie, maar die Romeinse.

Hierdie foutjie is tipies van die hele boek — daar is oral geringe mistastings.

3.3 Dit is vir my as Afrikaanse naamkundige glad nie duidelik waarom sekere name, "variante en afleidings", wel opgeneem is, en ander nie. Eksemplaries kyk ek na *Gaynor* en *Gail* in dié verband.

Gaynor is die trefwoord van 'n mooi 10-reël-artikel. Maar van *Gail* en sy "variante en afleidings" is geeneen in *Die Afrikaanse Naamboek* opgeneem nie. Hoe is die behandelde name uitgekies? En waarom is die weggelatenes wegge-laat?

Uit my studie van die frekwensie van voornamen in Afrikaans is dit baie duidelik dat *Gail* (in sy variante spellings) 'n baie hoër frekwensie as *Gaynor* het onder Afrikaanse mense jonger as 41.

Die keuse van behandelde eienamen skep 'n skewe beeld van die Afrikaanse voornaamskat.

3.4 Gevalle waarvan mens, wetenskaplik gesproke, maar ook populêr-wetenskaplik gesproke, sou verwag dat hulle op parallelle wyse behandel moet word, is nie altyd so behandel nie.

'n Voorbeeld. In die baie gevalle waar vrouename afleidings van grondliggende mansname is, word dit wel so aangedui. By *Flavia* — wat nou ook nie juis 'n frekwente vrouenaam in Afrikaans is nie — word geskryf: "Die vroulike vorm van *Flavius*. Sien *Flavius*." Onder *Johanna* word die manlike herkoms en die etimologie gegee. Maar onder *Frederika* word geen melding gemaak van die feit dat dit 'n vervrouliking van *Frederik* is nie (hoewel *Frederik* 'n inskrywing aan die seunsname-kant is). Onder *Salome* word geen melding gemaak van die feit dat dit 'n vervrouliking van *Salomo(n)* is nie.

3.5 Die keuse van hoofvorme van voornamen — in teenstelling tot "variante en afleidings" — skep 'n skewe beeld van die Afrikaanse naamskat.

Een voorbeeld: die artikel onder die trefwoord **Salmon**.

"**Salmon** 'n Hebreeuse naam afgelei van die woord *s'alôm* (vrede), dit wil sê 'vreedsame'.

Variante en afleidings: Salamon, Salie, Salman, Salomo, Salomon, Sol, Sallie, Solomon."

Op watter grond(e) is die trefwoord **Salmon** verkies? Waarom nie die meer frekwente *Salomon* nie? En waarom is die noemnaam *Sol* wat in Afrikaans byna totaal onbekend is, wel vermeld, maar die baie meer bekende *Sallie* nie?

Hierdie soort vraag kan mens onder byna elke naamartikel vra.

3.6 Hier en daar is daar flagrante misvattinge. Daar is byvoorbeeld onder die seunsname 'n artikel wat so lui:

"**Roan** Vermoedelik 'n baie ou Gotiese naam afgelei van die woord *runa*, wat 'geheim' beteken.

Variante en afleidings: Roann, Ruan."

Roan — wat weens uiters lae frekwensie nie opname verdien in 'n beperkte Afrikaanse voornaamwoordeboek nie — het weinig met Goties te make. Dit is

'n afwykende spelling van die Britse *Rowan*. En *Roann* is verdefliging van *Roan*. *Ruan*, aan die ander kant, word nie "róú-wín" uitgespreek nie, maar "róé-ann", en het niks met *Roan* of *Roann* te make nie. Dit is 'n afgekort- samegestelde neologisme uit *Rudolf* en *Johannes*.

3.7 'n Groot gemis in *Die Afrikaanse Naamboek* is dat dit nie die naamskat onder die Afrikaanse Moesliemmense behandel nie. Mens soek byvoorbeeld verniet na gewone name soos *Ayesha*, *Ismail*, *Sharifa*, *Nabuweya* en *Gabiba*.

Selfs wanneer hulle verwant is aan nie-Moesliem-name, makeer hulle totaal. Onder **Abraham** word die Moesliemnaam *Ibrahim* (en variante) byvoorbeeld nie vermeld nie en onder **Jacobus** nie die Moesliemnaam *Yakoob* (en variante) nie.

4. Ten slotte, net dit. Dit is goed dat daar na jare se relatiewe droogte weer 'n Afrikaanse voornaamboek op die mark is wat veel meer is as net "Wat noem ek my babatjie?" Ek beveel *Die Afrikaanse Naamboek* aan vir elke geïnteresseerde in name, en vir alle professionele mense wat naamgewende ouers moet bystaan.

Johan Combrink
Stellenbosch
Suid-Afrika

Problems in Swahili Lexicography¹

Albina R. Chuwa, *Institute of Kiswahili Research,
University of Dar es Salaam*

Abstract: Swahili dictionaries have been on the market since 1882. Most of these dictionaries were compiled by foreigners who were either missionaries or political administrators working in East Africa. Although the history of Swahili dictionary compilation seems to have survived a number of years, the field of lexicography is still a very new field in relation to what has been done by Tanzanians in this field. The first Swahili monolingual dictionary by a team of Tanzanian amateurs in dictionary compilation was published in 1981.

Due to inexperience in the field of lexicography the compilers have been facing various kinds of problems arising either from the technical level (whereby solutions for such problems are beyond the dictionary compilers) or from the practical level (whereby solutions depend more on the working team). When the two levels intertwine it makes the work of the dictionary compilers even more difficult, thus requiring support from outside the team. The sample problems discussed in this paper, though not exhaustive, pose some difficulties to Swahili dictionary compilers. Among those problems are:

1. Fiscal constraints which force compilers to embrace a big user target, making the dictionary objectives too wide to address any group satisfactorily.
2. The choice of language (standard versus non-standard varieties) inhibits the expansion of existing vocabulary in Swahili dictionaries.
3. The identification of lemmas and their grammatical categories still requires special attention from Swahili grammarians and structuralists.

Keywords: AMATEUR, CATEGORY, COMPILER, CORPUS, DATA BANK, DIALECT, DICTIONARY ENTRY, FREQUENCY, GRAMMAR, GRAMMATICAL CATEGORY, HEADWORD, LEMMA, LEXICOGRAPHY, MEANING, OBJECTIVE, PROBLEM, SELECTION (OF ENTRY), SOURCE, STANDARD DICTIONARY, SYNONYM, TARGET

Opsomming: Probleme in Swahili leksikografie. Swahili-woordeboeke is op die mark sedert 1882. Die meeste van hierdie woordeboeke is saamgestel deur buitelanders wat óf sendelinge óf politieke amptenare werksaam in Oos-Afrika was. Alhoewel die geskiedenis van die samestelling van Swahili woordeboeke daarop dui dat dit al 'n hele aantal jare bestaan, is die leksikografie steeds 'n baie nuwe veld betreffende werk gedoen deur Tanzaniërs. Die eerste Swahili eentalige woordeboek saamgestel deur 'n span Tanzaniese woordeboekamateurs is in 1981 gepubliseer.

Weens 'n gebrek aan ervaring in die leksikografieveld ervaar die samestellers verskillende soorte probleme wat spruit uit óf die tegniese vlak (waar oplossings buite die vermoë van die

1 This paper was presented at the First International Conference of the African Association for Lexicography, held at the Rand Afrikaans University, Johannesburg, 1st - 2nd July, 1996.

woordeboeksamestellers is) óf die praktiese vlak (waar oplossings meer by die werkspan berus). Wanneer die twee vlakke saamval, maak dit die taak van die woordeboeksamestellers selfs nog moeiliker, wat steun van buite die span vereis. Die voorbeeldprobleme wat in hierdie artikel bespreek word, alhoewel nie uitputtend nie, lewer probleme op vir die Swahili-woordeboeksamestellers. Onder hierdie probleme is:

1. Fiskale beperkings wat samestellers dwing om 'n omvattende doelmark in aanmerking te neem, wat die woordeboekdoelwitte te wyd maak om bevredigend in die behoefte van enige groep te voorsien.
2. Die keuse van taal (standaard- versus niestandaardvariëteite) beperk die uitbreiding van die bestaande woordeskat in Swahili-woordeboeke.
3. Die bepaling van lemmas en hul grammatikale kategorieë vereis steeds spesiale aandag van Swahili-grammatici en -strukuraliste.

Slutelwoorde: AMATEUR, KATEGORIE, SAMESTELLER, KORPUS, DATABANK, DIALEK, WOORDEBOEKINSKRYWING, FREKWENSIE, GRAMMATIKA, GRAMMATIKALE KATEGORIE, TREFWOORD, LEMMA, LEKSIKOGRAFIE, BETEKENIS, DOELSTELLING, PROBLEEM, SELEKSIE (VAN INSKRYWING), BRON, STANDAARDWOORDEBOEK, SINONIEM, DOELWIT

1. A Historical Perspective of Swahili Lexicography

The Swahili language, widely spoken in East Africa (Kenya, Uganda and Tanzania) and parts of the bordering countries (Rwanda, Burundi, Zaire and the Northern areas of Zambia, Malawi and Mozambique), has mainly depended on dictionaries compiled by foreigners who came to East Africa either as missionaries, colonial administrators or tourists.

Swahili dictionaries can be grouped into three publishing phases. In the first phase we find dictionaries published in the last decade of the 19th century and these include the works of the following personalities: Krapf, *Swahili-English Dictionary* (1882); Sacleux, *Dictionnaire Français-Swahili* (1891); Madan, *English-Swahili Dictionary* (1894) and *Swahili-English Dictionary* (1903). The second phase covers the period from 1939 to 1958. In this second phase there are works of Sacleux, *Dictionnaire Swahili-Français* (1939), Johnson F., *Kamusi ya Kiswahili yaani Kitabu cha maneno ya Kiswahili (A Swahili monolingual dictionary)* (1935), Johnson F., *A Standard Swahili-English Dictionary* (1939a), Johnson F., *A Standard English-Swahili Dictionary* (1939b) and Snoxall, *A Concise English-Swahili Dictionary* (1958). The third phase runs from 1981 to the present day. In this last phase a number of dictionaries have been published, such as the Institute of Kiswahili Research's² (henceforth IKR) *Kamusi ya Kiswahili Sanifu* (1981) (*A Standard Swahili Dictionary* — henceforth KKS), Lenselaer, *Dictionnaire Swahili-français* (1983), Miachina, *Kamusi ya Kiswahili-Kirusi* (1987)

2 In the reference section this Institute appears in its Swahili name 'Taasisi ya Uchunguzi wa Kiswahili'.

(*Swahili-Russian Dictionary*), Legere, *Wörterbuch Deutsch-Swahili* (1990), Baba Malaika, *The Friendly Modern Swahili-Modern English Dictionary* (1991) and Bakhressa, *Kamusi ya Maana na Matumizi* (1992) (*A Dictionary of Meaning and Usage*). Members of the lexicographical section of IKR have been involved in different capacities either as compilers or as reporters and the following has been realized through their joint efforts: In 1991 the IKR published *Kamusi ya Isimu na Lugha* (A linguistic dictionary), plus *Kamusi ya Biolojia, Fizikia na Kemia* (A dictionary of Biology, Physics and Chemistry). In 1992 IKR was co-author of the *French-Swahili Dictionary*, and in 1996 co-author of the *Swahili-French Dictionary*. Other lexicographical projects of IKR include the 2nd edition of the *Kamusi ya Kiswahili Sanifu* (in press); and the *English-Swahili Dictionary* (in press).

As seen from the above summary Swahili lexicography has focussed both on monolingual and bilingual dictionaries. The most used dictionaries, particularly in Tanzania, have been those of Johnson and the IKR.

2. The Art of Dictionary Making

Looking back from the time when the first Swahili dictionary by Tanzanians was published, and in relation to what has been done by Tanzanians in the field of lexicography, one would concur with the idea that lexicography is still a very new field in Tanzania. As mentioned above, Swahili dictionary compilers, especially in the 1st and 2nd phase, had been foreign missionaries or colonial administrators who depended on their linguistic intuition plus information they collected from their informants. Johnson's dictionaries in particular have valuable linguistic information which tell us that Johnson had a good linguistic background which has made his dictionaries authoritative and much depended upon for such a long time. Most dictionary compilers who published after 1939 used his dictionaries as their main sources.

The first Swahili dictionary by Tanzanians, the KKS, was compiled by a team of amateurs. The only expert was the team leader, Prof. R. Ohly, a Polish citizen who was then employed by the University of Dar es Salaam. With the little experience shared by the dictionary team, they very much depended on Johnson's dictionaries plus their personal intuition. The team set out to compile a standard language dictionary which was not well defined. The working team consisted of people not only of different ethnic groups but also people from different Swahili dialects. With such a mixed group therefore we cannot rule out the possibility that every one of them would have tried to include in the so-called standard dictionary some words from his own dialect. All these contributed to the rather poor quality of the product in comparison to Johnson's dictionaries. Too many synonyms and homonyms found their way into the *Standard Swahili Dictionary*.

Experience shows that like any new undertaking, several problems are

encountered by dictionary compilers, particularly when most of them are either amateurs or have little experience.

Dictionaries made by IKR after the KKS of 1981 had more advantages because the team included trained lexicographers. This factor has helped to speed up the process of dictionary making and a more critical approach was taken especially in the selection of entry words. Swahili lexicography has faced several problems and the main ones are discussed below.

3. Problems Encountered in Swahili Lexicography

3.1 Dictionary Objectives

The early Swahili dictionaries were compiled by people who wanted to use the product as one of their tools in performing whatever task was ahead of them. As such there was no earmarked target user — it was left open for all individuals who could get access to it. Such a target is too general and as such it is rather difficult to exhaust the language. Because they lacked identified users for their product the objectives were too wide to be embraced to the satisfaction of all users.

The KKS, published forty-two years after the first monolingual dictionary, also faced the problem of generality of objectives. It mentions a very wide range of users which would definitely create a problem in trying to satisfy them all. If we take the objectives mentioned in KKS we read '... it is targeted to students of all levels, teachers, writers, radio listeners, readers of books and newspapers...' (my translation — KKS 1981: xii). To satisfy such a wide range of users in a small-sized dictionary is too big an ambition. A primary school pupil definitely needs a more elementary dictionary whereas a college student requires a more advanced dictionary.

3.2 Choice of Language in Swahili Lexicography

Of the Swahili dictionaries mentioned above only KKS identified the kind of Swahili that was to be featured. IKR set out to compile a *Standard Swahili Dictionary*. The Inter-Territorial Swahili Language Committee under which Johnson published his dictionaries had chosen the dialect of Swahili spoken in Zanzibar at that period as the standard language. But because Swahili was used all over East Africa in the midst of other languages, and the fact that the language was rapidly growing, it lost more and more of its originality by incorporating words from other dialects and ethnic languages found in the mainland.

The problem of compiling a *Standard Swahili Dictionary* as pointed out by IKR arises from the fact that it was very difficult to identify the standard lan-

guage. The basic source of KKS was Johnson's dictionaries which had relied on informants and Johnson's linguistic knowledge. It is difficult to record a standard language without using documented materials which help to give an idea on the frequency use of the chosen words plus their meanings.

There is a time lapse from the time when Johnson's dictionaries were published to the time when the first Swahili monolingual dictionary by Tanzanians was compiled and published in 1981. As such one would expect KKS to be richer and more informative as a result of language growth. But this has not totally been the case because when KKS was declared a standard dictionary it restricted itself too much to the number of entry words purported to be standard, thus leaving dialectal and less standard words out.

Looking at what has so far been published in Swahili dictionaries we can conclude that very little has been done to increase the number of entries to cater for the growing language. This weakness can be attributed to the following problems. The first is the lack of a Swahili corpus — the main question remains — how did KKS collect data for a standard dictionary without a language corpus?

The question posed above still holds because although the Swahili language is rich in printed material covering a wide range of topics, it still lacks modern technology which would allow the team to build up a corpus and a data bank that can be accessed and easily retrieved by dictionary compilers. Lack of modern technology still force dictionary compilers to depend on their intuition plus that of a few informants, thus leaving out a substantial amount of valuable information which is essential not only to Swahili language learners but also to the daily users.

Depending on informants and the intuition of compilers contributed to the poor choice of entry words plus definitions which do not identify the defined. A lexicographical research conducted by Mdee (1996 — in press) on the flora and fauna featuring in KKS revealed that a number of the items listed were not recognized by his informants. For example, he had a total of thirty-five queries on animal names out of which ten were identified by an expert in wild life and only thirteen were identified by an informant who had participated in compiling that same dictionary. Twenty two names could not be identified by both of them. This indicates how much more research is required before we have a standard Swahili dictionary.

Another Swahili monolingual dictionary published in Kenya by Bakhressa based its work on KKS with no new information except for the inclusion of illustrative sentences on all headwords.

3.3 Lemma and Grammar in Swahili Dictionaries

3.3.1 Grammar in Swahili Dictionaries

Compilers of Swahili dictionaries have always used foreign grammar to

describe features found in the Swahili language. When describing grammatical categories the English classification has been used without taking into consideration the differences existing between the two languages. The question of the exact number of Swahili grammatical categories is still a crucial issue to be solved by Swahili grammarians. Identification of Swahili grammatical categories, especially the functional words, is essential if all words existing in the language are to receive fair treatment in dictionaries.

3.3.2 Lemma Form in Swahili Dictionaries

The Swahili language is an agglutinative language which requires some skills in identifying canonical forms that are acceptable as dictionary lemmas. Not all research assistants paid to collect data from printed Swahili materials came out with the required information. A quick review of dictionary cards collected by inexperienced research assistants hired by the Institute of Kiswahili Research (1994) revealed that agglutinated forms were registered as dictionary lemmas. For example:

a-tu-wafikish-e (= subject prefix + object marker + verb + imperative ending = he pacify us)
a-na-kwawa (= subject prefix + present tense marker + verb = he is hesitating)

Obviously one would have expected the two quoted cards above to represent lemmas *wafikisha* and *kwawa* instead of the clauses *atuwafikishe* and *anakuwawa* respectively.

3.3.3 Potentiality of Derivational System

The derivational system of Swahili has great potentiality in producing new words. There are a lot of words created through noun and verb derivations. The verb derivational system makes use of suffixes to express such features as applicative (*-ia, -ea*), passive (*-wa*), reciprocal (*-ana*), and stative (*-ka*), plus a prefix for the reflexive form (*ji-*). Although the system is well known to users of the language, it is still not very consistent. Some derivatives carry extra shades of meaning from those realized in the canonical form. KKS (1981) indicated the derivational suffixes at the end of a verb entry with no further comments. Baba Malaika (1991) went a step further by including derived forms of verbs as dictionary lemmas. For a more systematic approach whereby no redundancy is allowed into dictionaries and whereby words with specific meanings are accorded fair treatment in Swahili dictionaries, more research, funds and time should be allocated to the compilers. This would guarantee a better selection of entry words for future Swahili dictionaries.

4. Conclusion

We envisage that due to fiscal problems and time taken in compiling a dictionary, any substantial work on lexicography in Tanzania will still be under institutions for a longer time. Both sides, institutions and compilers, should therefore work more closely together so that the projects put forward are allocated enough time and resources to be able to deliver the goods that will meet the required standard.

References

- Baba Malaika. 1991. *The Friendly Modern Swahili Modern English Dictionary*. MS-tryk Denmark.
- Bakhressa, Salim, K. 1992. *Kamusi ya Maana na Matumizi*. Nairobi: Oxford University Press.
- Johnson, Fredrick. 1935. *Kamusi ya Kiswahili yaani Kitabu cha maneno ya Kiswahili*. Nairobi, Dar es Salaam: Oxford University Press.
- Johnson, Fredrick. 1939a. *A Standard Swahili-English Dictionary*. Nairobi, Dar es Salaam: Oxford University Press.
- Johnson, Fredrick. 1939b. *A Standard English-Swahili Dictionary*. Nairobi, Dar es Salaam: Oxford University Press.
- Krapf, L.A. 1882. *Swahili-English Dictionary*. London: Trubner and Company Ludgate Hill.
- Legere, K. 1990. *Wörterbuch Deutsch-Swahili*. Leipzig: VEB Verlag Enzyklopädie.
- Lenselaer, A. 1983. *Dictionnaire Swahili-français*. Paris: Éditions KARTHALA.
- Madan, A.C. 1894. *English-Swahili Dictionary*. London: Oxford University Press.
- Madan, A.C. 1903. *Swahili-English Dictionary*. London: Oxford University Press.
- Mdee, James S. 'Nafasi ya Msamiati usiotumika katika Kamusi ya Lugha Sanifu. Chuwa and Kihore (Eds.). 1996. *Kiswahili Katika Karne ya 21*. Institute of Kiswahili Research. (in press)
- Miachina, E.N. 1987. *Kamusi ya Kiswahili-Kirusi*. Moscow: Idara Ya Uchapishaji ya Lugha ya Kirusi.
- Sacleux, Charles. 1891. *Dictionnaire Français-Swahili*. Paris: Institut d'Ethnologie.
- Sacleux, Charles. 1939. *Dictionnaire Swahili-Français*. Paris: Institut d'Ethnologie.
- Snoxall, R.A. 1958. *A Concise English-Swahili Dictionary*. Nairobi: Oxford University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1981. *Kamusi ya Kiswahili Sanifu*. Nairobi/DSM: Oxford University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1991a. *Kamusi ya Isimu na Lugha*. Dar es Salaam: Dar es Salaam University Press.
- Taasisi ya Uchunguzi wa Kiswahili. 1991b. *Kamusi ya Biolojia, Fizikia na Kemia*. Dar es Salaam: Dar es Salaam University Press.
- Taasisi ya Uchunguzi wa Kiswahili, CREDU and Department of Foreign Languages and Linguistics. 1992. *French-Swahili Dictionary*. Nairobi: CREDU.
- Taasisi ya Uchunguzi wa Kiswahili, CREDU and Department of Foreign Languages and Linguistics. 1996. *Swahili-French Dictionary*. Nairobi: CREDU.
- Taasisi ya Uchunguzi wa Kiswahili. *Kamusi ya Kiswahili Sanifu*. (Second Edition — in press).
- Taasisi ya Uchunguzi wa Kiswahili. *English-Swahili Dictionary*. (In press).

Lexicography in Kenya: A Historical Survey

M.D. Mageria, *Department of Linguistics, University of Nairobi, Kenya*

Abstract: This paper traces the historical development of lexicography in Kenya. Attempts are made to show that although Kenya's linguistic landscape boasts with about 50 indigenous African languages, very little has been done in the field of compiling their dictionaries, especially by Kenyans themselves. Particular attention has been paid to the works done so far on the main African languages spoken in Kenya. It is argued that in the past five years Kenyan scholars have started to take considerable notice of lexicography as an academic subject. I have briefly focused on the language situation in Kenya, Swahili lexicography, professional literature and seminars. It is proposed that more efforts need to be advanced, in updating the already existing dictionaries as well as venturing into specific genres of lexicography like dictionary criticism and dictionary use.

Keywords: KENYA, DICTIONARY, SWAHILI, LEXICOGRAPHY, DICTIONARY USE, LINGUISTS

Muhtasari: Katika kazi hii, tunajaribu kumulika maendeleo ya taaluma ya leksikografia nchini Kenya. Ingawa kuna takriban lugha zaidi ya hamsini nchini Kenya, hakujakuwepo na jitihada madhubuti za kutungia kila lugha kamusi yake hususan na Wakenya wenyewe. Baadaye makala yanaonyesha kuwa kumekuwepo na mwamko fulani kwani wataalam wengi wa Isimu sasa wameanza kutambua taaluma hii nchini Kenya. Kwa ufupi tumetaja historia ya leksikografia ya kiswahili, semina na hali ya utungaji kamusi kwa sasa hapa nchini.

Maneno muhimu: KENYA, KAMUSI, KISWAHILI, LEKSIKOGRAFIA ISIMU, MATUMIZI YA KAMUSI

Introduction

Lexicography has witnessed tremendous changes particularly in the past two decades. In Europe and America, for example, there have been deliberate endeavours to improve it academically and as a professional discipline. Unfortunately, apart from the past five years or so, the same efforts have not been registered in Africa. Hartmann (1990: 71, 72) catches this mood in Africa when he writes that there are still very few monolingual dictionaries for the official and national languages, very few African nationals are engaged in dictionary making and a negligible number of countries have a professional body specializing in lexicography.

This is the same picture in the Kenyan scene. Most dictionaries available

now were written by missionaries and colonial administrators initially as wordlists, glossaries and supplements to grammars. The national policy on language is to a great extent dependent on the one inherited from the colonial government. In this regard, English as the official language remains the dominant language at the expense of the innumerable indigenous languages in Kenya. This means that there have been very little efforts connected with the general improvement and researching of the "local" languages, both in terms of compilation of their vocabulary as well as designing their grammars.

On the other hand there is the Swahili language. This is the national language of communication. There have been discernible attempts to map it as a *lingua franca* in East and Central Africa. It has benefited more from the Tanzanian scene than from Kenya. Most of the works available in Swahili now are as a result of the input from the Institute of Kiswahili Research in Dar es Salaam University in Tanzania. This is why this paper occasionally will traverse and enter into the Tanzanian scene.

Language Situation in Kenya

Kenya is a multilingual country. It is not yet very clear how many languages are spoken in Kenya. Estimates put the figures at about 40-50 languages. These languages can broadly be categorized into 3 elaborate groups: Foreign languages, e.g. English, French, German, Lingala, Swahili, and indigenous languages, such as Kikuyu, Luo, Kalenjin and Luhya.

As mentioned earlier, English is the official language. It is becoming increasingly important for our communication needs. This is because it is the sole language of instruction in the schools, the language of communication in the courts and a language through which government business is conducted. It is also seen as the language of the elite. Ironically, most of the young people in Kenya today would like to associate themselves with the English language more than Swahili or their first languages (mother-tongues). This therefore means that the indigenous languages are being sidelined further and further.

In the same breath, a new scene is developing in Kenya. Due to the insurmountable dependence of our economy on foreign aid, languages such as German, French, Japanese and Italian are now taking the centre stage. Most people are enrolling in foreign cultural centres like Goethe Institute, French Cultural Centre, Alliance Française and Japanese School. This is as a result of seeking employment in multinational firms, tourist companies or migrating to those countries, which necessary means ushering in better living standards.

Although Swahili should historically be grouped with the other indigenous languages, synchronically it deserves a special mention. It is the fastest growing language that has continued to serve as a viable political vehicle for nation building.

It is spoken extensively in Kenya as well as the whole of the East and

Central Africa region. In 1969, the ruling party in Kenya declared Swahili the official language of the country. But there were no concrete policy actions that supported this decision. Recently the government has realised the importance of Swahili and it has been given a higher status in the new educational system.

It is a compulsory subject examinable up to the secondary level. It is also used in parliament and political public meetings. According to Heine and Möhlig (1982: 61) Swahili is the most widely used spoken language in Kenya with a 65.3% figure compared to English 16.1% and vernacular second languages 13.7%.

With the advent of multiparty politics in Kenya ethnicity has continually threatened the unity that was hitherto seen in the country. This means that the government is not very keen on supporting the indigenous languages that exist. Numerically, the number of speakers of each of these languages corresponds to the demographic size of the ethnic community where they are spoken as first languages. In this regard Kikuyu, Luo, Kalenjin and Luhya are the largest in this category. Very little deliberate efforts are being made to augment their status in the country. As it is now, these languages are seen as the most divisive catalysts in the country. This is the reason we find a negligible number of recent publications that cover their vocabularies or grammars. These languages are therefore threatened with extinction since the younger generation is no longer interested in their acquisition nor are the policy makers enthusiastic about dealing with them.

Swahili Lexicography

Swahili has few monolingual dictionaries but many bilingual ones. This is because the avant-garde of Swahili lexicography were foreigners who came to East Africa as missionaries or colonizers. Basically they needed bilingual dictionaries which would help them to communicate with Africans. Kabalimu (1993: 1) writes:

Swahili lexicography can be traced back to the year 1811 when travellers and adventurers were compiling English-Swahili word lists. These works were fragmented and lacked linguistic sophistication and standard orthography.

As Swahili was used more and more in business administration and Christian evangelism, there was need to compile its lexicon and in 1882 Krapf published his collection of extensive Swahili vocabulary and grammar for missionary use. Later we find works by Madan (1894, 1903), Sacleux (1890, 1939 and 1959), Velten (1910), Johnson (1939). [See Kabalimu (1993) for full bibliographical details of these references].

Recent lexicographic pursuits have recorded works done by Africans

themselves. Of great significance is the monolingual *Standard Swahili Dictionary* (1981) published by Oxford University Press and compiled in Tanzania by Kiswahili scholars at the lexicography section of the Institute of Kiswahili Research, Dar es Salaam University.

This monumental step in Swahili lexicography was motivated by pragmatic reasons. After independence, Tanzania placed a lot of importance on self-reliance. Swahili was declared the official language. This meant that there was need for updating the existing Kiswahili literature. Khamis (1987: 194) claims that this dictionary, initially meant to be a bilingual piece, was compiled in order to "help supplement its (Tanzania's) egalitarian adult programme". It is instructive to note that only one Kenyan was involved in its compilation compared to 13 Tanzanian scholars. Ironically, this dictionary has sold more in Kenya than in Tanzania (in 1989, 13 editions were published in Kenya and only one in Tanzania). The Institute has also published several specialized dictionaries such as a dictionary of science, a dictionary of linguistics and language (1990), and a Swahili-English slang pocket dictionary (1978).

In Kenya, a new education system implemented in the late eighties gave special status to Swahili. There was need to change the syllabus in order to meet the stated goals of education in the country. Dictionaries have thus been published in order to meet the specific needs of the new system. Bakressa's *Dictionary of Meaning and Usage* (1992) best explains this situation. Despite its shortcomings (see Mageria 1995) this dictionary is a turning point in Swahili lexicography as far as the Kenyan scene is concerned.

Dictionaries

Dictionaries are regarded as language banks. They play an important role in preserving and registering the vocabulary of a language. Dictionaries also help to enhance the standardization of a language.

As already said, the category of foreign language has been given great importance in Kenya. It is not surprising to find that most dictionaries available in the country are those written for learners of English, French or German as foreign languages. They are the most expensive but the most used dictionaries. We are not going to deal with these types of dictionaries in this article.

Most of the dictionaries available to the third category of languages in our dichotomy were written by foreigners who visited Kenya. They came and stayed amongst the indigenous people, learnt their language and culture and hence used their vast lexicographical experience to compile dictionaries in these languages.

For example, Dr A. Barrett, a native of Ireland, worked for 16 years among the Turkanas of northern Kenya before compiling the *English-Turkana Dictionary* (1988). He says that he hoped "the dictionary will help preserve some of the wisdom of the Turkana themselves, benefit those (foreigners) who work in

Turkana ...". He therefore gives an introduction on the language and uses examples extensively and grammatical information to help those learning the language. English is used mainly as the source language, meaning that such dictionaries are geared towards those who already know English and are trying to learn Turkana. For example:

flight, n akerit, aporet; erot; to put to flight, akisur; flight of birds, a siriit a ngikuni.

We also find the *Kikuyu-English Dictionary* compiled by T.G. Benson in 1964. Mr Benson was given permission by the School of Oriental and African Studies in London to undertake the work of editing a standard dictionary of the Kikuyu language in 1954. This project lasted for over six years.

As far as we are concerned this is the most comprehensive dictionary ever, dealing with an African indigenous language in Kenya. It gives notes for the user of the dictionary. It uses Kikuyu language for illustration purposes, it gives the pronunciation of the Kikuyu alphabet and an in-depth guide to pronunciation indicators prepared by E. Sharp, a lecturer in phonetics at SOAS. The introductory matter takes 20 of the over 400 pages of the dictionary.

There is a detailed treatment of lemmata. For example:

Mumenyi, a — n, well informed, knowledgable, skilful person, an expert, if nyamenya, njorua kimenyi; — (in sarcastic sense) one with a special knowledge, a know-all.

-muno ndaturaga one who thinks he knows everything lives but for a short while.

This kind of explanation benefits both the learner of English as well as Kikuyu. There is the use of usage notes, grammatical information and illustrative examples all under one lexeme.

Multilingual dictionaries seem to be the predominant feature of Kenyan lexicography. In 1972 T.P. Gorman, an accomplished educationalist in East Africa, helped to prepare what he called *A Glossary in English, Kiswahili, Kikuyu and Dholuo*. He acknowledges that this was actually a very ambitious attempt. He wrote: "The matter of finding equivalents for English terms in the three vernacular languages is complicated by the fact that each (language) has a number of dialects; each language is also undergoing a rapid expansion in certain lexicon fields ..."

The issue of dialects is a tricky one and many lexicographers who associate themselves with Kenyan languages find themselves facing profound problems. Most of these languages do not have a standard structure or orthography. Towett's *English-Swahili-Kalenjin Nouns Pocket Dictionary* (1979) elucidates this fact. "Kalenjin" is not one particular language, it is a conglomeration of several languages namely Nandi, Tugen, Marakwet, Kipsigis and Elgeyo.

Although their level of intelligibility is high, they are mostly distinct languages. What Towett records in this dictionary is his own Kipsigis dialect which may not be very clear to a Nandi or a Tugen for that matter. The same can be said of the Luhya language which has Kabaras, Marogoli, Tiriki, Bukusu, Tachoni among other subdialects.

The Summer Institute of Linguistics has had its share of contributions. *A Tharaka, Swahili and English Dictionary* (1993b), *A Pokomo, Swahili and English Dictionary* (1993a), and a preliminary wordlist for *Duruma-Swahili-English* edited by Raphael Mkala in 1988 are some of the works compiled at their various stations in the country. These works are supposed to help in Bible translation and general literacy projects.

Recent Trends in Kenyan Lexicography

In the light of the foregoing, we have seen that for a long time Kenyans have not actively participated in their lexicographical orientations. This is true in respect of both Swahili and other vernacular languages.

Although we cannot claim that there has been much change in dictionary making, we have witnessed considerably more interest being shown in dictionary research.

When the Institute of Kiswahili Research (IKR) in Dar es Salaam in conjunction with Oxford University Press decided to review the *Standard Swahili Dictionary* 1981 edition, they decided to incorporate the views of the Kenyans. Two renowned Swahili scholars in Kenya were initially recruited to compile words peculiar to the country's Swahili speakers. These were Professors Ileri Mbaabu and Jay Kitsao. Consequently, there have been two lexicographical seminars held in Tanzania to discuss the intended new face of this dictionary. At each of these meetings, Kenya scholars have actively participated. Some of the papers presented by them have been the subject of two resultant lexicographical books edited by Dr J. Mdee and Dr J.G. Kiango at the Institute.

In 1995, for example, at a seminar held in Arusha, Tanzania, seven Kenyan scholars and publishers were involved. These included Prof. M. Abdulaziz, Prof. Okoth-Okombo, Prof. Bakari, Mr J. Habwe, Mr Mbatiah and Mrs E. Gatao, who works at the Oxford University Press. They looked at what was involved in the "publishing, marketing and distribution" of such a dictionary.

With regard to the training of lexicographers, very insignificant attempts have been made so far. Lexicography is now taught as a unit to the fourth year students of linguistics at the University of Nairobi. The course CLL 404, is meant to assist students to acquire basic skills in lexicography. It begins with a historical survey of lexicography then goes into the theory and practice of lexicography with an attempt to relating it to the African context. Due to this introduction we find some postgraduate students writing their M.A. theses in this area, for example Mageria (1995) and Iribe (forthcoming).

We have not been able to find the exact data on the state of actual involvement in dictionary making by Kenya publishing houses.

As seen in this overview, there remains a lot to be done especially in the training of lexicographers in Kenya. Unless this is done, the gap portrayed between the interest shown in dictionary criticism (theory) and in dictionary making (practice) will continue to widen in the country.

Conclusion

The art of dictionary making in Kenya is still at its nascent stages. Previously seen as a preserve for foreign visitors, now more Kenyans are developing a keen interest in lexicography. Unfortunately this is more evident when it comes to English and Swahili languages. Most of the indigenous languages have been neglected due to political and social reasons. There is a need to train more people in this craft so that they may do research and be involved in projects intent on writing dictionaries for Kenyan languages.

Of paramount importance is the establishment of closer co-operation with other institutions that deal with lexicography, particularly in Africa. A lot could be learnt from the lexicography division at IKR and the Bureau of the WAT at Stellenbosch. But this is only possible if we establish our own language research station in this country. Otherwise lexicography will remain an elusive art in Kenya.

References

- Bakhressa, S. 1992. *Kamusi wa Maana na Matumizi*. Nairobi: Oxford University Press.
- Barett, A. 1988. *English-Turkana Dictionary*. Nairobi: Macmillan.
- Benson, T.G. 1964. *Kikuyu-English Dictionary*. Oxford: Clarendon Press.
- Gorman, T.P. 1972. *A Glossary in English, Kiswahili, Kikuyu and Dholuo*. London: Cassell.
- Hartmann, R.R.K. (Ed.). 1990. *Lexicography in Africa*. Exeter Linguistics Studies 15. Exeter: Exeter University Press.
- Heine, B. and W.J.G. Möhlig. 1982. *Recent German Research on Africa: Language and Culture*. Tübingen: Institute for Scientific Co-operation.
- Iribe, M.P. Forthcoming. *Translation and Lexicographical Methods in Kamusi Ya Kiswahili (1939)*. Unpublished M.A. thesis. University of Nairobi.
- Institute of Kiswahili Research. 1981. *Kamusi ya Kiswahili Sanifu*. Dar es Salaam: University of Dar es Salaam.
- Kabalimu. 1993. *Treatment of Verb Derivatives in Swahili Dictionaries*. Unpublished M.A. thesis. University of Dar es Salaam.
- Khamis, A.M. 1987. Trends in Swahili Lexicography. *Kiswahili Vol I*. Dar es Salaam: Institute of Kiswahili Research.

- Mageria, M.D.** 1995. *The User Perspective in Swahili Dictionary of Meaning and Usage (1992)*. Unpublished M.A. thesis. University of Nairobi.
- Mdee, J.S.** 1990. *Theory and Methods of Lexicography in the Standard Swahili Dictionary*. Unpublished Ph.D. thesis. Leipzig: University of Karl-Marx.
- Summer Institute of Linguistics.** 1993a. *A Pokomo, Swahili and English Dictionary*. Nairobi: Bible Translation and Literacy Project.
- Summer Institute of Linguistics.** 1993b. *A Tharaka, Swahili and English Dictionary*. Nairobi: Bible Translation and Literacy Project.
- Towett, T.** 1979. *English-Swahili-Kalenjin Nouns Pocket Dictionary*. Nairobi: East Africa Literature Bureau.
- Wa Mberia, K.** 1993. Language and National Development. *Cultural Week Symposium Report: 2-3*. Nairobi.

Publikasie-aankondigings / Publication Announcements

William Branford. *The South African Pocket Oxford Dictionary of Current English*, 2nd South African edition 1994, 2nd impression 1995, xxiii + 1138 pp. ISBN 0-19-570760-5. Cape Town: Oxford University Press. Price R54,95. (Review in this issue)

J.G.H. Combrink en R. McD. Dodds. *Retrograde Woordeboek van Afrikaans*, 3de druk 1995, 291 pp. ISBN 0 9583222 8 7. Stellenbosch: Buro van die WAT. Prys R36,00.

Johan Combrink en Johan Spies. *SARA. Sakboek van regte Afrikaans*, 2de uitgawe, 1ste druk 1994, 214 pp. ISBN 0 624 03252 3. Kaapstad: Tafelberg. Prys R49,95. (Resensie in hierdie uitgawe)

Jonathan Crowther (Editor), Kathryn Kavanagh (Assistant Editor) and Michael Ashby (Phonetics Editor). *Oxford Advanced Learner's Dictionary of Current English*, 5th edition (New International Students' Edition) 1995, x + 1428 pp. ISBN 0-19-431423-5. Oxford: Oxford University Press. Price R54,95. (Review in this issue)

Dictionary of Grammar, 1994, reprint 1995, 256 pp. ISBN 1 85471 702 2. London: Claremont Books. Price 99p.

Margaret Doyle. *the A—Z of Non-Sexist Language*, 1st edition 1995, 112 pp. ISBN 0 7043 4430 0. London: The Women's Press Ltd. Price £6,99. (Review in this issue)

I.E. du Toit. *Tweetalige Polisiewoordeboek. Bilingual Police Dictionary*, 1ste uitgawe / 1st edition 1994, 469 pp. ISBN 0 627 01903 X. Pretoria: J.L. van Schaik. Prys / Price R57,50. (Resensie in hierdie uitgawe / Review in this issue)

M.L.A. Kgasla and J. Tsonope. *Thanodi Ya Setswana*, 1st edition 1995, xxii + 330 pp. ISBN 99912 66 39 9. Gabarone: Longman Botswana. Price P38,50.

J. Kruijzen en J. Kokkelmans. *Woordenboek van de Limburgse Dialekten I. Agrarische Terminologie. Aflevering 6: Boerderij, bedrijfsgebouwen*, 1995, XII + 162 pp. ISBN 90 232 3009 4. Assen: Van Gorcum. Prijs f44,00.

Longman Dictionary of Contemporary English, 3rd edition 1995, xxii + 1668 + B18 pp. ISBN 0-582-23750-5. Harlow: Longman Group Ltd. Price R49,95. (Review in this issue)

Steven McGill. *Who's Who in Lexicography. An International Directory of EURALEX Members*, 1st edition 1996, iii + 157 pp. ISBN 1 900778 00 9. Exeter: Dictionary Research Centre. Price £5,00.

Rajend Mesthrie. *A Lexicon of South African Indian English*, 1st edition 1992, xxix + 148 pp. ISBN 0 94883 10 6. Leeds: Peepal Tree Press. Price £7,95. (Review in this issue)

N.E. Osselton. *Chosen Words. Past and Present Problems for Dictionary Makers*, 1st edition 1995, x + 188 pp. ISBN 0 85989 419 3. Exeter: University of Exeter Press. Price £27,50. (Review in this issue)

A.F. Prinsloo en F.F. Odendal. *Afrikaans op sy beste. Hulp met moderne taalkwessies*, 1ste uitgawe 1995, 231 pp. ISBN 0 627 01990 0. Pretoria: J.L. van Schaik Uitgewers. Prys R54,95.

Penny Silva (Managing Editor), Wendy Dore, Dorothea Mantzel, Colin Muller and Madeleine Wright (Editors). *A Dictionary of South African English on Historical Principles*, 1996, xxx + 825 pp. ISBN 0-19-863153-7. New York: Oxford University Press Inc. Price R550,00.

Terminology Committee for Social Work. Vaktaalkomitee vir Maatskaplike Werk. *New Dictionary of Social Work. Nuwe Woordeboek vir Maatskaplike Werk*, 1995, xxiii + 140 pp. ISBN 0-621-16908-0. Pretoria: Government Printer. Die Staatsdrukker. Price / Prys R45,60. (Review in this issue / Resensie in hierdie uitgawe)

Della Thompson (Editor). *The Concise Oxford Dictionary of Current English*, 9th edition 1995, xxi + 1673 pp. ISBN 0-19-861319-9. Oxford: Clarendon Press. Price R125,00. (Review in this issue)

Annelize van Rooyen. *Die Afrikaanse Naamboek*, 1ste uitgawe, 2de druk 1995, 121 pp. ISBN 1 874901 18 X. Kaapstad: Queilleries-Uitgewers. Prys R44,95. (Resensie in hierdie uitgawe)

A.A. Weijnen. *Etymologisch dialectwoordenboek*, 1996, XXVII + 269 pp. ISBN 90 232 2967 3. Assen: Van Gorcum. Prijs f65,00.