

Lexikos 21

Lexikos 21

Redakteur

Editor

Elsabé Taljard

Resensieredakteur

Review Editor

T. Harteveld



African Association for Lexicography

AFRILEX-REEKS 21:2011

AFRILEX SERIES 21:2011



BURO VAN DIE WAT

STELLENBOSCH

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

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

Uitgewer Publisher

BURO VAN DIE WAT
Posbus 245
7599 STELLENBOSCH

Kopiereg © 2011 deur die uitgewer
Alle regte streng voorbehou
Eerste uitgawe 2011

Tipografie en uitleg
deur Tanja Harteveld en Hermien van der Westhuizen
Bandontwerp deur Piet Grobler
Geset in 10 op 12 pt Palatino

Gedruk en gebind deur Onyx Press
Stewartstraat 61 Goodwood

ISBN 978-0-9814434-3-0
ISSN 1684-4904

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

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

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

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

Lexikos is elektronies beskikbaar by Sabinet, AJOL, Ebsco en Proquest
Lexikos is available online from Sabinet, AJOL, Ebsco and Proquest

Indekse Indexes

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

Hierdie uitgawe van

Lexikos

in die

AFRILEX-REEKS

word opgedra aan

J.C.M.D. du Plessis

redakteur van *Lexikos* 7 tot 20 (1997 – 2010)

Tegniese Assistent: 1963 – 1965

Mederedakteur: 1965 – 1981

Senior Mederedakteur: 1981 – 1984

Assistenthoofredakteur: 1985 – 1990

Eindredakteur: 1991 – 2000

van die

WOORDEBOEK VAN DIE AFRIKAANSE TAAL

This issue of

Lexikos

in the

AFRILEX SERIES

is dedicated to

J.C.M.D. du Plessis

editor of *Lexikos* 7 to 20 (1997 – 2010)

Technical Assistant: 1963 – 1965

Co-editor: 1965 – 1981

Senior Co-editor: 1981 – 1984

Assistant Editor-in-chief: 1985 – 1990

Final Editor: 1991 – 2000

of the

WOORDEBOEK VAN DIE AFRIKAANSE TAAL

Adviesraad / Advisory Board

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

Redaksiekomitee / Editorial Committee

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

Inhoud / Contents

Voorwoord	xi
Foreword	xii
<i>Elsabé Taljard</i>	
'n Woord van AFRILEX	xiii
A Few Words from AFRILEX	xvi
<i>Gilles-Maurice de Schryver</i>	
Redaksionele doelstellings	xix
Editorial Objectives	xx
Redaktionelle Ziele	xxi
Politique éditoriale	xxii

Artikels / Articles

Between <i>designer drugs</i> and <i>afterburners</i> : A Lexicographic-Semantic Study of Equivalence	1
<i>Arleta Adamska-Salaciak</i>	
National Lexicography Units: Past, Present, Future	23
<i>Mariëtta Alberts</i>	
Needs-adapted Data Presentation in e-Information Tools	53
<i>Henning Bergenholtz and Theo J.D. Bothma</i>	
Voorbeelde en ooraadressering in tweetalige woordeboeke	78
<i>Herman L. Beyer</i>	
Equivalent Selection in Specialized e-Lexicography: A Case Study with Spanish Accounting Terms	95
<i>Pedro A. Fuertes-Olivera</i>	
From Corpus to Dictionary: A Hybrid Prescriptive, Descriptive and Proscriptive Undertaking	120
<i>Minah Nabirye and Gilles-Maurice de Schryver</i>	

The Compilation of Multilingual Concept Literacy Glossaries at the University of Cape Town: A Lexicographical Function Theoretical Approach <i>Dion Nkomo and Mbulungeni Madiba</i>	144
A Critical Analysis of the Lemmatisation of Nouns and Verbs in isiZulu <i>D.J. Prinsloo</i>	169
Management and Internal Standardization of Chemistry Terminology: A Northern Sotho Case Study <i>Elsabé Taljard and Mahlodi Jean Nchabeleng</i>	194
Pedagogical Lexicography: Towards a New and Strict Typology Corresponding to the Present State-of-the-Art <i>Sven Tarp</i>	217
Theoriebedingte Wörterbuchformprobleme und wörterbuchformbedingte Benutzerprobleme I. Ein Beitrag zur Wörterbuchkritik und zur Erweiterung der Theorie der Wörterbuchform <i>Herbert Ernst Wiegand und Rufus H. Gouws</i>	232
Leksikonotas / Lexiconotes	
Categorising Example Sentences in Dictionaries for Research Purposes <i>Lorna Hiles</i>	298
Problems of Usage Labelling in English Lexicography <i>Lydia Namatende Sakwa</i>	305
Leksikohuldeblyk / Lexicotribute	
Dr. J.C.M.D. du Plessis: Redakteur van <i>Lexikos</i> 1997 tot 2010	316
Dr J.C.M.D. du Plessis: Editor of <i>Lexikos</i> 1997 to 2010	318
Resensieartikels / Review Articles	
Quand la pratique lexicographique se modernise en RD Congo. Note sur <i>Nkòngamyakù Cilubà–Mfwàlànsa</i> , dictionnaire bilingue de NgoSemzara Kabuta <i>Crispin Maalu-Bungi</i>	320

- Regard critique sur *On Est Ensemble: 852 mots pour comprendre le français du Gabon* 337
Pierre Ondo-Mébiame et Guy-Modeste Ekwa Ebanéga

Resensies / Reviews

- Monika Bielińska: *Lexikographische Metatexte. Eine Untersuchung nichtintegrierter Außentexte in einsprachigen Wörterbüchern des Deutschen als Fremdsprache* 359
Maria Smit

- Heberto H. Fernández Urdaneta. *Dictionaries in Spanish and English from 1554 to 1740: Their Structure and Development* 366
J.C.M.D. du Plessis

- Pedro A. Fuertes-Olivera (Editor). *Specialised Dictionaries for Learners* 369
Birger Andersen

- Jan Posthumus. *Woordenboeken en hun lotgevallen: Studies over lexicografische praktijk (ca. 1850 tot heden)* 381
Jan Posthumus. *Lexicografie in Nederland: Peilingen in de negentiende en twintigste eeuw*
J.C.M.D. du Plessis

- Wiegand, Herbert Ernst und M^a Teresa Fuentes Morán. *Estructuras lexicográficas. Aspectos centrales de una teoría de la forma del diccionario* 384
Michael Mann

Publikasieaankondigings / Publication Announcements 391

- Voorskrifte aan Skrywers 394
Instructions to Authors 396
Richtlinien für Autoren 398
Instructions aux Auteurs 400

Voorwoord

Dit is grootliks te danke aan die meesterhand van die vorige redakteur, dr. J.C.M.D. du Plessis, dat Lexikos gedurende die afgelope 14 jaar homself as vaktydskrif van keuse vir beide plaaslike en internasionale leksikograwe gevestig het. Daarvan getuig die verskeidenheid van onderwerpe wat in Lexikos 21 aan die bod kom. Ook in hierdie uitgawe word die oorspronklike doelstellings van Lexikos vergestalt. Dit bevat naamlik artikels wat die metaleksikografie as tema het, en ander wat die toepassing van die nuutste teorieë op praktiese wyse illustreer. Die sinergie wat tussen die metaleksikograaf en die praktiserende leksikograaf bestaan, verseker dat die twee nie van mekaar vervreem raak en die metaleksikografie in die moeras van 'teorie ter wille van teorie' vasval nie. 'n Tweede gesprek wat in Lexikos 21 gevoer word, is die dialoog tussen Afrika en die internasionale leksikografiegemeenskap. Die term 'dialoog' word doelbewus hier gebruik, omdat hierdie nooit as 'n eenrigtinggesprek gesien mag word nie. Dit is net so belangrik vir die res van die wêreld om kennis te neem van die dinamiese en kreatiewe wyse waarop daar in Afrika met beide die praktiese en teoretiese leksikografie omgegaan word, as wat dit vir Afrika is om op die hoogte van die nuutste internasionale ontwikkelinge op dié gebied te bly.

As amptelike mondstuk van Afrilex rus daar 'n besondere verpligting op Lexikos om nie net op die tale van Suid-Afrika te konsentreer nie, maar ook om gestalte te gee aan Afrilex as 'n vereniging van en vir Afrika. Dit is daarom verblydend om ook in hierdie uitgawe van Lexikos artikels te kan plaas waarin tale van elders in Afrika die fokus vorm.

In die verlede is die elektroniese beskikbaarstelling van Lexikos deur Sabinet en AJOL behartig. Ebsco en ProQuest het vanjaar met die WAT kontrakte gesluit om Lexikos in elektroniese formaat te bemark en verkoop. Hierdie stap sal Lexikos se sigbaarheid in die internasionale gemeenskap verder verhoog.

Elke uitgawe van Lexikos is die resultaat van 'n spanpoging. Bydraers is sekerlik een van die belangrike spelers in die Lexikos-span, maar ewe belangrik is die keurders, wat onselfsugtig hul tyd en kundigheid beskikbaar stel om van elke Lexikos 'n gehalte tydskrif te maak. Die nougesetheid waarmee outeurs gewoonlik die kommentaar van keurders opvolg en implementeer, weerspieël die waarde wat aan die insette van keurders geheg word. Wat die tegniese en administratiewe aspekte van die publikasie van Lexikos betref, rig ek graag 'n besondere woord van dank aan me. Tanja Harteveld, wat besonder geduldig met 'n nuwe en onervare redakteur was. Saam met me. Hermien van der Westhuizen, was sy verantwoordelik vir die setwerk, en haar oog vir detail en nougesette werkswyse het veel bygedra tot die tegniese kwaliteit van die publikasie.

Elsabé Taljard
Redakteur: Lexikos 21

Foreword

It is largely due to the expert touch of the previous editor, Dr. J.C.M.D. du Plessis, that *Lexikos* has established itself during the last 14 years as the journal of choice for both local and international lexicographers. This is borne out by the range of subjects that is represented in *Lexikos* 21. In this volume, too, the original objectives of *Lexikos* are embodied. It contains articles with metalexigraphy as theme, and others that illustrate the practical application of the latest theories. The synergy that exists between the metalexigrapher and the practising lexicographer, ensures that these two do not become estranged, and prevents the metalexigraphy being embogged in 'theory for the sake of theory'. A second discourse in *Lexikos* 21 is the dialogue between Africa and the international lexicographic community. The term 'dialogue' is deliberately being used here, as it should never be regarded as a one-way conversation. It is as important for the rest of the world to take note of the dynamic and creative way in which both the practical and the theoretical lexicography in Africa are being dealt with, as it is for Africa to keep track of the latest international developments in this subject field.

As official mouthpiece of Afrilex *Lexikos* has a special duty not only to concentrate on the languages of South Africa, but also to realize Afrilex' goal to be an association of and for Africa. It is therefore heartening to be able to publish articles in this volume of *Lexikos* in which languages from elsewhere in Africa form the focus.

In the past *Lexikos* was made available in electronic format through Sabinet and AJOL. This year Ebsco and ProQuest signed contracts with the WAT to market and sell *Lexikos* in electronic format. This step will raise the visibility of *Lexikos* further in the international community.

Each volume of *Lexikos* is the result of a team effort. Contributors surely are some of the most important players in the *Lexikos* team, but the reviewers are as important. They provide their time and expertise unselfishly in order to make each *Lexikos* a journal of quality. The meticulous way in which authors usually follow up and implement the comments of reviewers reflects the value that is placed on these inputs. As regards the technical and administrative aspects of the publication of *Lexikos*, I extend a word of thanks to Ms Tanja Harteveld, who was exceptionally patient with a new and inexperienced editor. She was responsible for the typesetting, together with Ms Hermien van der Westhuizen, and her eye for detail and conscientious way of working has contributed much to the technical quality of the publication.

Elsabé Taljard
Editor: Lexikos 21

'n Woord van AFRILEX

Dr. J.C.M.D. du Plessis het ons in die Voorwoord van verlede jaar se uitgawe van *Lexikos* meegedeel dat daardie volume die laaste een sou wees waarvoor hy as redakteur van ons tydskrif sou optree. Hy het aangekondig dat die redakteurskap in die toekoms op 'n roterende basis deur proff. Elsabé Taljard, Danie Prinsloo en Rufus Gouws behartig sal word. Ons verwelkom hulle, en veral vir Elsabé, wat bereid was om eerste aan die beurt te wees. Elsabé neem oor op 'n tydstip waarin *Lexikos* maksimaal funksioneer, en soos in die geval van 'n 4 x 100-afloswedloop, is hierdie taak ewe afhanklik van bekwaamheid as spoed. Johan, wat *Lexikos* vir nie minder nie as veertien uit die vorige twintig jaar geredigeer het, het van krag tot krag gegaan met elke nuwe uitgawe. Met volume 20 het ons uitmuntende redakteur 'n puik ampstermyn op 'n skitterende wyse afgesluit. Met 800 bladsye het Johan se finale uitgawe gegroei tot meer as net 'n nommer van 'n tydskrif om sodoende sy plek te verseker op die rakke vir die mees gewaardeerde boeke in 'n biblioteek, waar dit nou sy aan sy staan met die gewigtigste woordeboeke — heel gepas vir ons vakrigting.

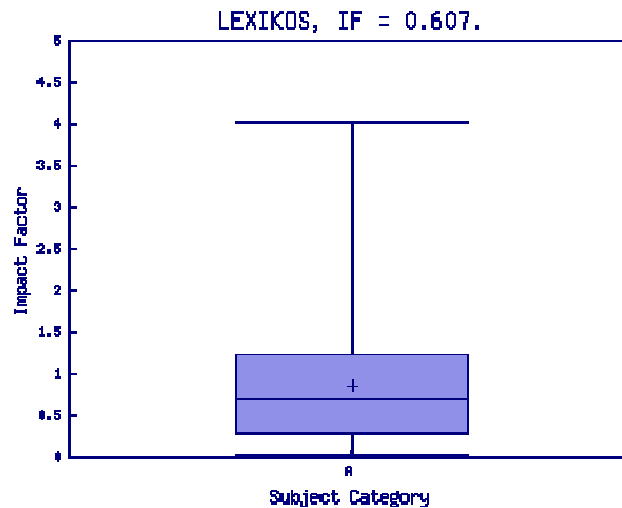
Dit wat die kwantitatiewe prestasies betref. Kwalitatief was die opgang selfs meer indrukwekkend. *Lexikos* het in 2008 aangesluit by die wêreld se top linguistiektydskrifte, toe dit 56ste geplaas is uit 68 tydskrifte in die kategorie vir Linguistiek. Indien hierdie kategorie in vier kwartiele verdeel word, het *Lexikos* gemaklik tot die lys toegetree in die laaste (vierde) kwartiel. In 2009 het *Lexikos* se impakfaktor meer as verdubbel, van 0.300 (in 2008) tot 0.667 (in 2009), en was dit toe 49ste uit 92 tydskrifte, wat die tydskrif in die derde kwartiel geplaas het. Met verlede jaar se uitgawe het *Lexikos* hierdie uitstekende plasing behou, met 'n impakfaktor van 0.607 wat toegeken is. Interessant genoeg, in 'n lys tydskrifte in die kategorie vir Linguistiek, wat nou baie langer is en waar *Lexikos* tans 66ste geplaas is uit 141 tydskrifte, het ons tydskrif weer verbeter, aangesien *Lexikos* die tweede kwartiel in 2010 betree het — net een kwartiel weg van die toppunt.

In die *Journal Citation Reports* van Thomson Reuters kan bogenoemde grafies gesien word met 'n sogenaamde 'Box Plot' of houerstipping, soos aangedui op die volgende bladsy. Die horisontale lyn wat die houer sny, is die mediaan-impakfaktor vir die kategorie Linguistiek (A).¹ Hierdie mediaan is 0.562 vir 2010; dit wil sê met 'n impakfaktor van 0.607 vir *Lexikos* in 2010 het *Lexikos* inderdaad die tweede kwartiel betree, m.a.w. die bestek tussen die mediaan en

1 *Lexikos* behoort tot net een kategorie, Linguistiek, wat as A aangedui word op die grafiek. Party tydskrifte behoort tot meer as een kategorie, en hulle impakfaktor word dan voorgestel op meervoudige houerstippings (A, B, ens.).

'n Woord van AFRILEX

die horisontale lyn wat die bokant van die houer vorm.² Om hierdie uitstekende prestasies te konsolideer is nou die uitdaging vir die volgende drie redakteurs in ons figuurlike afloswedloop. Dit is egter vir die snelle vordering tot hier waarvoor ons die dankbaarste behoort te wees.



Ek wil graag nog meer hulde bring aan Johan. Asof dit nie genoeg is dat hy die ekwivalent van ses maande van elke jaar gewy het aan die produksie van 'n *Lexikos*-uitgawe nie, met die gepaardgaande hantering van ongeveer 1 500 e-posse jaarliks,³ het Johan, die akademikus, ook ontwikkel in Johan, die grootmoedige. Hy het alles in sy vermoë gedoen om 'n ware kultuur van wetenskaplike verslaggewing oor leksikografie in Afrika te help vestig. Buiten die sorgvuldige leiding wat hy aan nuwelinge gegee het deur hulle te help met talle herskrywings aan hulle bydraes, het Johan ook in 2003 die "*Lexikos* encouragement prize for scholarly writing" bekendgestel. Die doel van hierdie prys was in sy eie woorde:

Om studente in die leksikografie en jong leksikograwe aan te moedig om betekenisvolle navorsing op hul vakgebied te lewer, en standarde van wetenskaplike skryfwerk op die gebied van die leksikografie te verhoog. Mededinging om die prys geskied dus deur middel van 'n wetenskaplike artikel.

— J.C.M.D. du Plessis (2003, uit 'n insetsel ingesluit by *Lexikos* 13)

-
- 2 Die derde kwartiel is die bestek tussen die mediaan en die horisontale lyn wat die onderkant van die houer vorm, terwyl die vierde kwartiel die bestek onder die houer beslaan. Die eerste kwartiel is die bestek bo die houer.
 - 3 Dit is 'n konserwatiewe ekstrapolering gebaseer op slegs vyf e-posse met die outeur(s) van elke manuskrip, en vier e-posse elk met die twee beoordelaars van elke manuskrip. Daar word aanvaar dat die aanvaardingskoers ongeveer die helfte is.

Johan het nie net gepraat nie, maar ook die prysgeld vir die weners beskikbaar gestel.

Ten slotte kan 'n mens met vrymoedigheid verklaar dat Johan nie alleen die aansien van *Lexikos* en daarom ook van AFRILEX verhoog het met sy onvoorwaardelike toewyding aan sy taak nie, maar ook van die leksikografie as 'n akademiese vakgebied as geheel. Ons betuig ons hartlikste dank aan dr. Johan du Plessis — 'n gewaardeerde kollega en 'n goeie vriend!

Gilles-Maurice de Schryver
President: AFRILEX

A Few Words from AFRILEX

In the Foreword to last year's volume of *Lexikos*, Dr. J.C.M.D. du Plessis informed us that that volume was going to be his last as Editor of our journal. He announced that the baton would henceforth be passed on to a rotating triumvirate: Proff. Elsabé Taljard, Danie Prinsloo, and Rufus Gouws. We welcome them, and especially Elsabé, who was found willing to be next in line. Elsabé takes over at a moment when *Lexikos* has reached top speed, and as in a 4 x 100 relay race, the task is just as much a skill event as it is a speed event. Johan, who had been editing *Lexikos* for no less than fourteen of the first twenty years, went from strength to strength with each new volume. With Volume 20, our Editor-par-Excellence put a magnificent capstone to an outstanding tenure. Weighing in at 800 pages, Johan's capstone simply outgrew the concept of a single journal issue, catapulting it to the most valued library shelves, where it now rubs shoulders with the heftiest of dictionaries — quite appropriate for our discipline.

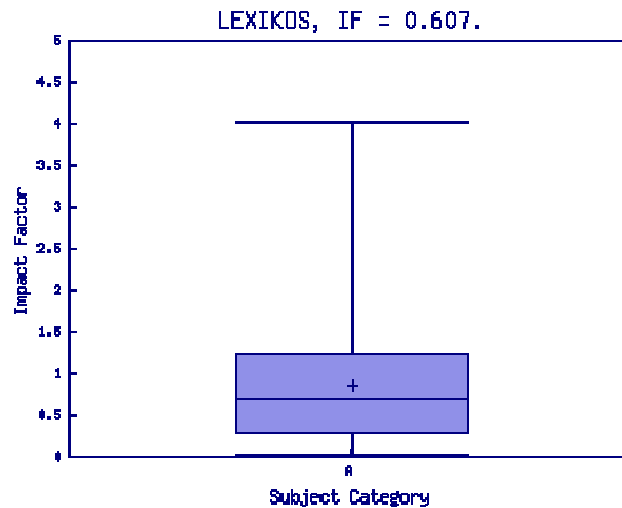
So much for the quantitative achievements. Qualitatively, the rise has been even steeper. *Lexikos* joined the world's top linguistics journals in 2008, when it was ranked 56th out of just 68 journals in the Linguistics category. Dividing the Linguistics category into four quartiles, *Lexikos* had safely joined the list in the last (fourth) quartile. In 2009, the Impact Factor of *Lexikos* more than doubled, from 0.300 (in 2008) to 0.667 (in 2009), now ranking 49th out of 92 journals, and bringing the journal into the third quartile. With last year's volume, finally, *Lexikos* held onto this excellent score, being awarded an Impact Factor of 0.607. Interestingly, in a now much larger list of journals in the Linguistics category, where *Lexikos* currently ranks 66th out of 141 journals, our journal gained yet again, as *Lexikos* entered the second quartile in 2010 — only one quartile away from the absolute top.

In the Journal Citation Reports of Thomson Reuters this can be appreciated graphically with a so-called 'Box Plot', as shown on the next page. The horizontal line that intersects the box is the median Impact Factor for the Linguistics category (A).¹ That median is 0.562 for 2010, so with an Impact Factor of 0.607 for *Lexikos* in 2010, *Lexikos* indeed just entered the second quartile, that is, the range between the median and the horizontal line that forms the top of the box.² Consolidating this excellent track record is now the challenge for the

1 *Lexikos* belongs to just one category, Linguistics, referred to as A in the graph. Some journals belong to multiple categories, and their Impact Factor is then visualized in multiple Box Plots (A, B, etc.).

2 The third quartile is the range between the median and the horizontal line that forms the bottom of the box, while the fourth quartile is the range under the box. The first quartile is the range above the box.

next three editors in our metaphorical relay race. But it's for bringing us this far this fast that we need to be most grateful.



I would like to pile even more praise onto Johan. As if it was not yet enough to spend the equivalent of six months of every year on the production of a volume of *Lexikos*, handling an annual 1 500 e-mails in the process,³ Johan-the-Academic also revealed himself as Johan-the-Generous. He went to extraordinary lengths to help create a true culture of scientific reporting on lexicography in Africa. In addition to the painstaking guidance he offered to new entrants, helping them through the numerous rewritings of their contributions, in 2003 Johan also launched the "*Lexikos* encouragement prize for scholarly writing". In his own words, the aim of this prize was:

To encourage students in lexicography and young lexicographers to conduct significant research in their field of study, and to raise the standard of scholarly writing in the field of lexicography. Competing for the prize therefore takes the form of a scientific article.

— J.C.M.D. du Plessis (2003, from a leaflet included with *Lexikos* 13)

Johan also put his money where his heart was, and further included prize money for the winners.

In conclusion, then, one can confidently state that through the unconditional devotion to his job, Johan not only raised the prestige of *Lexikos* and thus AFRILEX, but also of lexicography as an academic discipline as a whole. To Dr

3 This is a conservative extrapolation, based on just five e-mails with the author(s) of each manuscript, and four e-mails each with the two adjudicators of each manuscript. Acceptance rate is assumed to be about a half.

A Few Words from AFRILEX

Johan du Plessis — valued colleague and dear friend — our sincerest gratitude is extended!

Gilles-Maurice de Schryver
President: 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 interdisiplinêre aard van die leksikografie, wat ook terreine soos die taalkunde, algemene taalwetenskap, leksikologie, rekenaarwetenskap, bestuurskunde, e.d. betrek, onder die algemene aandag te bring;
- (5) om beter samewerking op alle terreine van die leksikografie moontlik te maak en te koördineer, en
- (6) om die doelstellings van die *African Association for Lexicography* (AFRILEX) te bevorder.

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

- (1) Bydraes tot die leksikografiese gesprek word in die vaktydskrif *Lexikos* in die AFRILEX-reeks gepubliseer.
- (2) Monografiese en ander studies op hierdie terrein verskyn as afsonderlike publikasies in die AFRILEX-reeks.
- (3) Slegs bydraes wat streng vakgerig is en wat oor die suiwer leksikografie of die raakvlak tussen die leksikografie en ander verwante terreine handel, sal vir opname in die AFRILEX-reeks kwalifiseer.
- (4) Die wetenskaplike standaard van die bydraes sal gewaarborg word deur hulle aan 'n komitee van 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 coordinate cooperation in all fields of lexicography; and
- (6) to promote the aims of the *African Association for Lexicography* (AFRILEX).

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

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

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

Redaktionelle Ziele

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

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

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

Gemäß den Zielsetzungen der AFRILEX-Serie werden:

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

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

Politique éditoriale

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

Les objectifs de la collection AFRILEX sont de :

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

Pour atteindre ces objectifs, la collection AFRILEX

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

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

Between *designer drugs* and *afterburners*: A Lexicographic-Semantic Study of Equivalence

Arleta Adamska-Sałaciak, *Department of Lexicology and Lexicography, School of English, Adam Mickiewicz University, Poznań, Poland (arleta@ifa.amu.edu.pl)*

Abstract: The lexicons of natural languages are not isomorphic. Reasons for the anisomorphism can be sought on three interrelated planes: language structure, extralinguistic reality, and conceptualisation. Simply put, the relevant differences may reside in the language, the world, the mind, or any combination of these. As a result, what goes under the name of *lexicographic equivalence* is a rather heterogeneous category. Growing awareness of this fact has resulted over the years in the creation of several tentative typologies of equivalence, one of which is presented below, together with a brief discussion of some strategies for dealing with imperfect equivalence.

The remaining part of the article comprises a detailed analysis of a single problem encountered while preparing a new edition of a bilingual dictionary for Polish learners of English. The task at hand involved choosing a viable counterpart for a (Polish) semantic neologism from among a few (English) equivalence candidates. In the discussion, reference is made both to the metalexigraphic categories introduced earlier and to such concepts developed by lexical (especially cognitive) semantics which may prove helpful in capturing the meaning differences between the source-language item and its competing target-language renditions.

This micro-scale dissection of a single specimen demonstrates that we are still some way from being able to classify, let alone deal with, all the instances of imperfect interlingual correspondence that come our way. Persisting in the efforts to advance our understanding of the complex issues covered by the blanket term *lexicographic equivalence* thus seems crucial for improving the treatment of meaning in bilingual dictionaries.

Keywords: EQUIVALENCE, LEXICOGRAPHY, BILINGUAL DICTIONARY, MEANING, LEXICON, CORPUS, ENCODING, CONCEPTUALISATION, COGNITIVE SEMANTICS, CONCEPTUAL METAPHOR, DENOTATION, CONNOTATION, FIXED PHRASE, FREE COMBINATION, TRANSLATION, CULTURE

Opsomming: Tussen *designer drugs (ontwerpersdwelms)* en *afterburners (naverbranders)*: 'n leksikografiese-semantiese studie van ekwivalensie. Die leksikons van natuurlike tale is nie isomorf nie. Redes vir die gebrek aan isomorfie kom op drie onderling verwante vlakke voor: taalstruktuur, die buitetalige realiteit, en konsepualisasie. Anders gestel kan die tersaaklike verskille in die taal, die wêreld, die verstand, of enige kombinasie hiervan aangetref word. Gevolglik is dit wat as *leksikografiese ekwivalensie* bekend staan 'n nogal heterogene kategorie. 'n Groeiende bewuswording van hierdie feit het oor die jare daartoe gelei dat verskeie tentatiewe tipologieë van ekwivalensie geskep is, waarvan een hieronder aangebied word, saam

met 'n kort bespreking van sommige strategieë om onvolledige ekwivalensie te hanteer.

Die oorblywende deel van die artikel bevat 'n uitvoerige ontleding van 'n enkele probleem wat teëgekem is toe 'n nuwe uitgawe van 'n tweetalige woordeboek vir Poolse aanleerders van Engels saamgestel is. Die betrokke taak het behels dat 'n lewensvatbare teenhanger vir 'n (Poolse) semantiese neologisme uit 'n paar ekwivalente (Engelse) alternatiewe gekies word. In die bespreking word verwys na sowel die metaleksikografiese kategorieë wat vroeër geopper is, as na konsepte in die leksikale (veral die kognitiewe) semantiek wat sou kon help om die betekenisverskil tussen die brontaalitem en sy kompeterende doeltaalvertalings vas te vang.

Hierdie ontleding van 'n enkele voorbeeld op mikroskaal demonstreer dat ons nog ver daarvan is om al die gevalle van onvolledige intertalige ooreenstemming wat ons teëkom te kan klassifiseer, wat nog te sê van te hanteer. Om vol te hou met ons pogings om die komplekse vraagstukke te verstaan wat deur die oorkoepelende term *leksikografiese ekwivalensie* gedek word, is dit dus uiters belangrik om die hantering van betekenis in tweetalige woordeboeke te verbeter.

Slutelwoorde: EKWIVALENSIE, LEKSIKOGRAFIE, TWEETALIGE WOORDEBOEK, BETEKENIS, LEKSIKON, KORPUS, ENKODERING, KONSEPTUALISERING, KOGNITIEWE SEMANTIEK, KONSEPTUELE METAFLOOR, DENOTASIE, KONNOTASIE, VASTE FRASE, VRYE KOMBINASIE, VERTALING, KULTUUR

1. General considerations

1.1 Why is there no perfect interlingual equivalence?

The naïve view of natural languages, which assumes that, for every language pair, any given lexical item found in one language can be matched semantically with its counterpart in the other, is fundamentally flawed. Philosophers and linguists, not to mention bilingual lexicographers, have repeatedly pointed out that there is, in fact, no such perfect interlingual correspondence. Reasons for this can be located on three interrelated planes: language structure itself, the extralinguistic world, and the human mind.

Cases of the first kind, which Zgusta (1971) sees as springing from a "difference in the inventory of non-designative items", are easily exemplified by situations in which a grammatical category present in one of the languages (e.g. English articles, Zulu ideophones) is absent from the other. Those of the second kind (Zgusta's "difference in denotata"; Gouws's (1996) "referential gap") prototypically include geographically restricted features of the natural environment, as well as culture-specific customs, institutions and dishes. Less prototypically, they may also involve creatures and phenomena whose ontological status depends on the (religious and/or folk) beliefs characteristic of a particular linguistic community.¹ Finally, cases of the third kind, i.e. those originating in conceptualisation (Zgusta's "difference in designata"; Gouws's (1996) "linguistic gap"), are a consequence of the fact that, rather than merely reflecting the world, language imposes structure upon it — in other words, the categorisation of extralinguistic reality is language-dependent. Countless examples have been

cited in the literature of individual concepts which are not lexicalised in particular languages: there is, for instance, no word for 'puppy' in Afrikaans (Gouws 2002: 200). Consider also the varying level of detail with which family relationships can be categorised linguistically: in some languages, special words exist for the maternal and paternal grandfather/grandmother/uncle or for various kinds of in-laws, despite the relationships themselves being, biologically speaking, the same for all humans.

This third, most complex category of mismatches, dependent on differences in conceptualisation, covers a multitude of non-obvious cases. Thus, language B may lack a lexical item of the same level of generality as language A — as in Keley-i, which has a number of words sharing the general meaning 'to alter or destroy the structure of something by pressure', but no word for the shared meaning itself:

- let-ik* 'to crush between fingernails, e.g. lice'
 - pedit* 'to crush with fingers or foot, e.g. insects'
 - kupikup* 'to crush with hands, e.g. paper'
 - gudu* 'to crush anything dried into bits or powder form with fingers, e.g. leaves, tobacco leaves'
 - bugay* 'to crush into smaller pieces or powder with hands, feet or other instrument, e.g. soil, lump of salt'
 - ubud* 'to crush any form of seed into powder with an instrument, e.g. coffee bean'
 - teptep* 'to crush something with a stone or heavy object, e.g. betelnut'
 - gemik* 'to crush shell, e.g. egg'
 - ledih* 'to crush pepper or fruit to produce juice, with an instrument'
 - pitpit* 'to crush a can totally, with feet or an instrument'
 - pitul* 'to crush a can or plastic container partially, with feet or an instrument'
- (after Hohulin 1986: 44f.)

In this Philippine language there is, thus, no hyperonym designating the whole class of actions, i.e. no word corresponding to the English *crush*, with the help of which the verbs have been glossed.

Less obviously still, the names of certain emotions, mental states, or social institutions, routinely given as interlingual equivalents, may be associated with concepts which are far from identical. Consider, for instance, the concept of marriage in different cultures — including aspects such as the admissible number of spouses, how easy it is to enter into or dissolve a marital union, etc. — or the concept of anger in most European languages (where traces of the ancient theory of the four humours can still be detected)² as opposed to 'the same' concept in Japanese,³ Chinese,⁴ Zulu,⁵ or the language spoken on the Micronesian atoll of Ifaluk⁶.

That the structure of a particular language, especially its lexicon, depends on conceptualisation has been stressed very strongly by contemporary seman-

ticists, particularly those of the cognitive persuasion; however, the idea itself is hardly new. Over three hundred years ago, Locke (1689/90) postulated that all knowledge consists of ideas which, in turn, begin in experience, that is, in sensations (directed outside, to the world) and reflections (directed inside, to the mind). Immediately after experience has played its part, the mind starts working of its own accord, reshaping the simple, experience-driven ideas into complex ones and making the latter conform to the needs of communication. It is only by means of language — which Locke conceives of primarily as a stock of words — that an individual can make his thoughts known to others. Words, which are signs of ideas, do not denote individual things, but classes of things. These classes, rather than being a feature of the world around us, are created by the mind. Human creativity is manifested in the names of substances (objects in nature), relationships (e.g. in mathematics) and, most of all, so-called mixed modes (e.g. moral notions like 'justice'). The fact that there exist different languages is due mainly to such complex ideas, which tend to be culture-specific. Moreover, complex ideas are often so vague and unstable that they do not allow for a unanimously accepted definition — at least not out of context.⁷

1.2 Varieties of lexicographic equivalence

Given that Locke's scenario, or something very much like it, pertains to every individual language, the identification of interlingual equivalents can hardly be a straightforward task. And that is, by no means, the whole problem. As noted by many (recently, e.g., Hallsteinsdóttir and Farø 2010), it makes little sense to talk about 'equivalence in general'; rather, what counts as equivalence will depend on a particular research perspective (e.g. lexicological, translational, lexicographic or didactic).

Even adopting an exclusively lexicological vantage point, as was done in the preceding section, it seems pretty obvious that one cannot expect a bilingual dictionary to offer perfect interlingual equivalents. What bilingual lexicographers have been doing for centuries is providing equivalents which, while not perfect, go some way towards meeting the dictionary user's expectations. Equivalence is thus a gradable phenomenon. Furthermore, different types thereof need to be distinguished if we want to understand what goes on in the pages of bilingual dictionaries. Working on such dictionaries for a number of years and studying the relevant metalexicographic literature (particularly the writings of Ladislav Zgusta) has led me to believe that what is called equivalence in lexicography comes in four different, though partly overlapping, varieties:

- (i) cognitive (semantic, systemic, prototypical, conceptual, decontextualised, notional)
- (ii) translational (insertable, contextual, substitutable, textual)
- (iii) explanatory (descriptive)
- (iv) functional (situational, discourse, dynamic, pragmatic, communicative).⁸

1.2.1 Cognitive equivalence

A cognitive equivalent is situated at the level of the language system. Being fairly general, it is appropriate as a translation of the source-language (SL) item in many, though not all, contexts. When not insertable in a particular context, it should ideally elicit in users an appropriate translational equivalent. This is possible thanks to the cognitive equivalent's explanatory potential, i.e. its ability to faithfully represent the meaning of the SL item. For all these reasons, the cognitive equivalent, wherever available, is the lexicographer's first choice. Not only that: cognitive equivalents tend to be reassuringly identical, no matter which bilingual dictionary for a given language pair one happens to consult. Thus, for instance, German *sterben* will normally be supplied with the English equivalent *die*, English *head* with French *tête*, and Spanish *pesado* with English *heavy* (in their primary senses, of course). Furthermore, the equivalents will often exhibit symmetry (*sterben* = *die* in a German-English dictionary; *die* = *sterben* in an English-German one, and so on).

1.2.2 Translational equivalence

A translational equivalent is situated at the level of the text, that is, it produces a good translation when substituted for the SL item in a particular context or contexts. It is what the bilingual lexicographer settles for when a cognitive equivalent is not available. Due to the (in principle unlimited) number of contexts in which a given SL lemma may occur, the number of its possible translational equivalents can sometimes be quite high. Dictionary entries reflect this, often featuring more than one equivalent per sense. Translational equivalents are especially useful in dictionaries, or parts thereof, aimed at foreign language production. Below are some examples.

- (1) Nguni *ubuntu* — E. *human-heartedness* (DSA, after Gouws 1996)
- (2) E. *awe* — Polish *respekt, trwoga, podziw* (Polish does not have a lexicalised concept corresponding to the English one; each of the three equivalents provided — respectively, 'respect', 'fear', and 'admiration' — covers part of the SL meaning.)
- (3) Russian *intelligent* — E. *intellectual* (The equivalent will work as a translation in many contexts, but, according to Farina (1996: 4), it fails to reflect the richness of the Russian concept.)

1.2.3 Explanatory equivalence

In contrast to the previous two types, an explanatory equivalent is a free combination, not an established lexical unit of the target language. It is a TL paraphrase of the SL item, situated somewhere between a TL equivalent and a

TL definition. Thanks to its explanatory power, it should help the dictionary user come up with a translational equivalent appropriate on a particular occasion (that is, perform a function similar to that of a non-insertable cognitive equivalent). Lexicographers resort to explanatory equivalents only in the absence of a 'proper' (i.e. cognitive or, at least, translational) equivalent. The solution tends to be employed when dealing with culture-specific items and is found more often in older dictionaries than in contemporary, corpus-based ones, which in similar circumstances turn to functional equivalence. Giving an explanatory equivalent is justifiable primarily in dictionaries, or parts thereof, aimed at foreign language reception. The following examples illustrate these points.

- (4) Venda *mbongo* — E. *food prepared from freshly harvested maize* (Madiba and Nkomo 2010: 319)⁹
- (5) Nguni *ubuntu* — E. *quality embodying all the traditional virtues and values of isintu*; the word *isintu* refers to African characteristics, style, values and traditions (DSA, after Gouws 1996)
- (6) Russian *intelligent* — E. *member of the intelligentsia* (This explanatory equivalent "implies a certain politeness of behavior and delicacy that are not necessary to qualify as an intellectual" (Farina 1996: 4). Thus, although not as readily insertable in context, it conveys the SL meaning better than the translational equivalent *intellectual*.)

The three equivalence types discussed so far are closely related. Zgusta (1987: 238) saw the ideal dictionary equivalent as combining insertability with high explanatory potential. A cognitive equivalent does precisely that (though its insertability is contextually restricted). A translational and an explanatory equivalent can be seen as complementary, each one performing half of the cognitive equivalent's job.

1.2.4 Functional equivalence

The last variety of equivalence is the most distinctive of the four. Functional equivalence holds between portions of text larger than individual lexical items. Crucially, the TL portion contains either a TL word of a different grammatical category than the SL lemma or a TL expression with no element whatsoever corresponding directly to the SL lemma. Unlike in translatology, where functional equivalence has always enjoyed a privileged status, in lexicography it was, until recently, relatively marginal. This is now changing rapidly. In an era when respectable dictionaries are more often than not corpus-based, lexicographers are free to exploit the rich stores of data (i.e. contextual occurrences) at their disposal for the purpose of tailoring such syntagmatic stretches containing the SL lemma which can be comfortably translated into the target language

without requiring a lexical equivalent for the SL lemma itself. As an example, consider the following entry from De Schryver's innovative Zulu-English school dictionary which demonstrates how Zulu ideophones can be rendered contextually into English:

- (7) **mpo ideophone 1** ► (of being erect) ♦ Wavuka wama mpo uGovu wathi: ... • Govu woke up, stood straight up, and said: ... **2** ► (of extreme action) ♦ Sekubanda, mpo, eNingizimu. • It is now cold, extremely cold, in the South. ♦ Kuthe ilanga selithe, mpo, ibandla lihlezi phansi nje emthunzini. • When it is extremely hot, the group of men simply sits on the ground in the shade. **3** ► (of being full) ♦ Ligcwaliseni isaka lithi mpo. • Fill up the bag to the brim.
(after De Schryver 2009: 40)

Note that no attempt was made here to offer decontextualised equivalents: the sense discriminators are followed directly by examples of usage and their English translations; equivalence obtains primarily between the sentences as wholes, and only to a limited extent between the Zulu ideophone *mpo* and whatever English phrase corresponds to it in the translation.

1.3 Dealing with imperfect equivalence in a bilingual dictionary

All those cases where a lexicographer cannot come up with a cognitive equivalent of the SL item are jointly referred to here as instances of imperfect equivalence. There are two major groups of strategies for dealing with such cases. The first, obvious one involves employing equivalence types other than the cognitive — that is, other than the default. The bilingual lexicographer can thus provide:

- (i) two or more cumulative partial equivalents (a variety of translational equivalence, as in the *awe* example in 1.2.2)
- (ii) a paraphrase (i.e. an explanatory equivalent, as in the *intelligent* example in 1.2.2 and 1.2.3)¹⁰
- (iii) an extension of the syntagmatic scope of the lemma and its TL translation (which amounts to achieving functional equivalence, as in the *mpo* example in 1.2.4).

If none of the above appears feasible, one can try:

- (iv) giving a cultural counterpart¹¹ (e.g. the name of an educational institution, government body, etc., which plays a roughly similar role in the TL speech community)
- (v) giving a partial (translational) equivalent plus a gloss (in parentheses) or an explanatory note (following the entry proper)

- (vi) proposing an innovation (via sanctioning a lexical or semantic borrowing from the source into the target language).¹²

Having examined different varieties of lexicographic equivalence and isolated some useful strategies for coping with troublesome instances, we are now equipped for the discussion of a concrete case of imperfect equivalence which the concepts introduced above should help us handle.

2. Case study

2.1 The problem

One issue following from the absence of perfect interlingual equivalence is the subordinate question of how to choose between several imperfect equivalent candidates — the main problem of encoding into a foreign language. This will be the focus of the rest of the article.

The case related below is an actual dilemma faced by the lexicographic team preparing the second edition of LSW, a bilingual dictionary for Polish learners of English (of junior high and high school age). The main feature which distinguishes LSW from its competitors on the Polish market is that it is the first (and so far the only) truly encoding dictionary available. This means, among other things, that all the equivalents have been carefully chosen, the main requirement being that they must be insertable in the translations of the example sentences supplied (that is, they must work as translational equivalents).¹³

In the autumn of 2010, when the new edition of LSW was being compiled, a discussion raged in the Polish media about the dangers of so-called *dopalacze* — legally available substitutes of illegal drugs, sold in special shops, ostensibly as collector's items. Widespread use of the drugs had led to a series of teenage deaths, which caused a public outcry and eventually prompted the authorities to introduce a ban on selling them. Of course, the problem has not thereby disappeared, as one can still buy *dopalacze* via the internet or bring them in from abroad.

Given that LSW is directed at young people, it was deemed necessary to include the entry for *dopalacze* in its Polish-English part.¹⁴ This provoked a number of questions. No obvious equivalent sprang to mind (i.e. one that would qualify as cognitive in view of the classification in 1.2), but there were a few candidates for translational equivalents that the lexicographers could choose from. What follows is a reconstruction of the reasoning which informed the choices made in the end.

First, let us see what can be established about the meaning of the Polish word. Upon typing in *dopalacz* (the singular form of *dopalacze*) in the Polish version of Wikipedia, one is instructed to select one of the following senses:

- (1) a device used in jet planes which increases engine thrust
- (2) a device used in cars which reduces exhaust toxicity
- (3) (in plural) a popular name for certain intoxicants, often containing benzilpiperazine.

Note the grammatical difference, the last sense being formally distinguished by its use in the plural.¹⁵ The third definition differs from the remaining ones in yet another respect: it talks about the name, not the referent, thus creating some distance between the two.¹⁶ This implies that the name is somehow less natural or less literal, or simply less established, in that third sense. In the Wikipedia entry for *dopalacze* (plural), some more details are provided: it is a popular name for different types of products containing psychoactive substances which are not included in the list of legally prohibited drugs; the use of these products is meant to produce an intoxicant effect as close as possible to that of illegal substances.

It should be clear from the above that the drug sense is derived from the jet-plane sense of *dopalacz* by means of metaphorical extension: a drug, specifically a so-called upper, acts on the human psyche in a way analogical to the effects of a thrust augments in a jet plane. Accordingly, in cognitive linguistic terms, *dopalacze* can be seen as a figurative expression instantiating the conceptual metaphor THE HUMAN BODY IS A MACHINE, with its constituent submetaphors: THE BRAIN IS AN ENGINE; ENHANCED BRAIN PERFORMANCE IS INCREASED ENGINE THRUST; DRUGS ARE THRUST AUGMENTERS.¹⁷

2.2 Equivalent candidates: Evidence from texts, dictionaries, and corpora

2.2.1 Designer drugs

The *dopalacze* article in Polish Wikipedia mentions the term *designer drugs*, stating authoritatively that this is the name under which the substances are known in English. It thus makes sense to examine this equivalent candidate first.

When one googles (or yahoos) the phrase, one is told that a *designer drug* is 'a drug with properties and effects similar to a known hallucinogen or narcotic but having a slightly altered chemical structure, especially such a drug created in order to evade restrictions against illegal substances' (<http://www.answers.com/topic/designer-drug>; <http://education.yahoo.com/reference/dictionary/entry/designer%20drug>); or, 'a drug produced by a minor modification in the chemical structure of an existing drug, resulting in a new substance with similar pharmacologic effects, especially one created to achieve the same effect as a controlled or illegal drug' (<http://dictionary.reference.com/browse/designer+drug>).

So far, so good. If other dictionaries agreed with AHD (from which these definitions appear to have been taken), we would need to look no further. Unfortunately, English dictionaries are not unanimous. In MEDAL, a *designer drug* is defined as 'an illegal drug produced artificially from chemicals: *designer drugs like ecstasy*'; in LDOCE, as 'an illegal drug that has an exciting or relaxing effect, and is taken for pleasure'; in OALD, as 'a drug produced artificially, usually one that is illegal: *a tablet of the designer drug Ecstasy*'; in CALD, as 'any of various strong drugs that has been changed to give it a similar effect to an illegal drug such as cocaine'; in CED, as 'any of various narcotic or hallucinogenic substances manufactured illegally from a range of chemicals'; finally, in OED, as 'a drug synthesized to mimic a legally restricted or prohibited drug without itself being subject to such restriction'.

Two things are worrying here. For a start, *ecstasy*, which features in the OALD and MEDAL examples of use, would never be thought of as a *dopalacze* in Polish. This points to a difference in meaning between *designer drugs* and the Polish term: *dopalacze* are substances whose chemical composition is deliberately kept secret, there is an air of mystery about them — these things are part of the concept. Secondly, there is no agreement among the English definitions regarding the legal status of the drugs: of the seven dictionaries quoted, only AHD, CALD, and OED claim that designer drugs are legal.¹⁸ In a way, the confusion matches the Polish situation: the selling of *dopalacze* has been banned, but this does not seem to have led to a change in the collective consciousness. *Dopalacze* are still perceived as something less serious and less dangerous than straightforward *narkotyki* 'drugs' (the two names are often contrastively juxtaposed), both in terms of health risks and possible criminal consequences.

Designer drugs having thus failed to meet the requirements for a perfect equivalent, the search must continue.

2.2.2 Afterburners

Describing the media hysteria surrounding *dopalacze*, a blog post from the Economist uses three English terms: *designer drugs*, *legal highs*, and *afterburners*. The introductory paragraph goes as follows:

For a week now Poles have been in the throes of collective delirium over so-called **legal highs**. Known colloquially in Poland as "**afterburners**", these **designer drugs** began cropping up five years ago, offering party thrills and cognitive enhancement. Crafty vendors have skirted drugs laws by labeling their wares as "collectibles" or "plant food", "unsuitable for human consumption". Meanwhile, clever chemists on their payroll continually tweak the substances' chemical composition so whenever one compound is outlawed, a modified, legal cousin takes its place. (http://www.economist.com/blogs/easternapproaches/2010/10/legal_highs_poland)

Strictly speaking, contrary to what the passage says, it is not true that the sub-

stances in question are "known colloquially" as *dopalacze*: the name is, in fact, the only Polish name in use (although it is true that the word is occasionally preceded by *tzw.* 'so-called', testifying to its relative novelty). Throughout the rest of the text, the author uses *afterburners* interchangeably with *legal highs*.

Is there any independent evidence for *afterburners* being employed in this sense? Not if one consults the English Wikipedia or the OED. The relevant entry from Wikipedia gives a definition which corresponds to sense 1 of *dopalacz* in the Polish definition given in 2.1 above:

An **afterburner** (or **reheat**) is an additional component added to some jet engines, primarily those on military supersonic aircraft. Its purpose is to provide a temporary increase in thrust, both for supersonic flight and for takeoff (as the high wing loading typical of supersonic aircraft designs means that take-off speed is very high).

OED distinguishes two senses, corresponding to senses 1 and 2 of *dopalacz*:

1. An auxiliary burner fitted to the exhaust-pipe of a turbo-jet engine to increase its thrust.
2. An auxiliary burner in a flue, exhaust pipe, etc., designed to burn any remaining combustible waste gases.

No corpus occurrences have been found documenting the drug sense, either. All the instances of *afterburner(s)* that have been identified in corpora (BNC — 7; ukWaC — 214; enTenTen — 466) are either to do with plane or car engines, or they are metaphorical (typically in texts about running or other sport-related activities), often appearing as product names (title of a record album; title of a poetry volume; something to do with back lighting in Game Boy; a game you play in an arcade; software to make a program run faster). Nowhere are the effects of drugs referred to or even hinted at, although the metaphorical potential (of an energy boost or of enhancing performance) is both clearly present and often exploited, as in the following examples:¹⁹

ukWaC:

Afterburner is an add-on utility for Director which compresses animations for more efficient distribution via the WWW.

The impala normally travels in graceful arching bounds, but when this particular impala discovered that the race cat was on his track, he stretched out full length horizontally and shifted into **afterburner**.

Most energy is burned by metabolism; increasing the resting metabolic rate can have a dramatic effect. One of the best ways to do this is by exercising. Exercise burns energy. It also continues to burn extra energy by raising the RMR from 1 up to 24 hours after exercising. This is a bit like having an **afterburner**. To get

prolonged after burn you need to exercise for a longer period at a relatively moderate intensity. 30 minutes is perfect. However, sprinting to catch that bus in a morning does not generate metabolic after burn.

On entry the flavour is cool, fizzy, citrus — lemon and honey, then the taste of soft gooseberry and pear in syrup evolve as it glides across the palette with cereal and toasted muffins in the slipstream. There is no evidence of the power at this stage because the viscosity is almost like glycerine, however as it engages the taste receptors at the back of the tongue it really hits the booster button and an amazing heat floods deep into the chest. It's brilliantly fresh and fizzy with an extremely pleasant **afterburner** effect and leaves an aftertaste that is superb, not long, but unforgettable — believe me. In 1695 the spirit would have been drunk straight from the still and would certainly have been perilous indeed.

enTenTen:

Henderson was the Class 4 A Louisiana state high school champion in both the 100 meters (10.3) and 400 meters, then clocked the second-fastest 60-yard dash ever (6.7) at LSU. Last season, his **afterburners** enabled him to average 23.2 yards per reception, best in the National Football League.

Barber under-performed in two of her best events, the Long Jump and Javelin on Sunday, but set off at such a lick in the final event, the 800 metres that it looked as if she might yet steal the title. But with 200 to go, Klüft switched on the **afterburners**, and even ran a personal best, 2.08.89, to take the gold with 6 887 points, to Barber's 6 824.

Adams turned on the **afterburners** for the final 100 metres, creeping up to the first place spot.

After assuming the lead, Robinson puts on the **afterburners** to roll into the finish line at Loma Luz Clinic, completing the fifty-seven miles in two hours, thirty-three minutes.

If you invest too much on infrastructure then you draw more on the already already scarce people and resources in an economy that has the **afterburners** turned on.

One father referred to his son as "all **afterburner** and no rudder."

Let's just say that my Palm IIIc carries numerous bits'n'pieces around, but the most frequently used seem to be **AfterBurner** (overclocking software to make it run faster), Eudora (a very nice Email program), Wordsmith (for those unavoidable fits of creative writing), iSilo and Plucker (for reading compressed repositories of HTML), APCalc (because the standard PalmOS calculator is a load of rubbish), and Sim City (the killer app for the colour Palm Pilots).

In sum, it appears that the Economist blogger used *afterburners* idiosyncratically and creatively — in other words, he/she innovated.

2.2.3 Legal highs/Legal intoxicants

After typing *legal highs* in the search box of Wikipedia, one is redirected to *legal intoxicants*. Under the latter, a variety of things are mentioned: stimulants (e.g. caffeine), hallucinogens/psychoactive (e.g. nutmeg, but also synthetic cannabinoids, including the drug *Spice*), deliriant (e.g. certain antihistamines), depressants (e.g. alcohol), inhalants (e.g. laughing gas), and opioids (e.g. codeine). The noun phrase *legal intoxicants* thus appears to designate a much broader category of substances than *dopalacze*.

OED does not have an entry for *legal intoxicants*. Nor do any of the English corpora consulted (BNC, ukWaC, enTenTen, COCA, BASE, BAWE). This necessarily leads to the conclusion that *legal intoxicants* is not a fixed phrase, but a free combination.

By contrast, the phrase *legal high(s)*, though also absent from OED, can be found in corpora (e.g., BNC — 5; ukWaC — 24; enTenTen — 19; COCA — 1),²⁰ where it occasionally appears between quotation marks. The examples below (with their original spelling retained) clearly show that the term *legal highs* corresponds denotatively to *dopalacze*:

BNC:

Cannabis dealers, on the other hand, make far less money and can end up in prison. In order to save law-abiding FACE readers from a possible prison sentence, I decided to investigate some of the **legal highs** that are available.

After explaining the nature of my quest, they told me about their own favourite **legal high**. Potter's Herbal Cigarettes are a smoking preparation made from stramonium. My new companions break up the cigarettes, and make tea with the herb. They claimed that drinking the tea gave them an intense, trip-like experience.

ukWaC:

Draconian penalties for cannabis possession may encourage young people to experiment with "**legal highs**" in the mistaken belief that what is not banned must be relatively safe.

Last year's ban on 'magic mushrooms' by the Government left a gap in the market for '**legal highs**'.

But any biochemist will tell you that it is most unlikely that any group of natural products is to be found only in a single species, and that many other drug-containing plant materials are now freely on sale — either as "**legal highs**" or as plants and seeds.

And as for the **legal highs**, for something that supposedly are shit how come they have stoped a big pill muncher like me taking xtc after 4 years of regular weekend abuse?!

Its effects are more hallucinatory than other **legal highs** — though high doses of the raw plant are usually needed to achieve these effects.

The best **legal high** ive tried yet are the smilies now they are one to definately try if you havent already! id say they are closest to e ive tried.

ABC News reports that 9% of US 8th graders are regularly inhaling ordinary — and potentially deadly — household products as a cheap and **legal high**, according to 2001 federal drug statistics.

enTenTen:

Ace understands that there's no need for weed when your house is a veritable cornucopia of cheap, **legal highs**.

When you have a hideous irrational drug policy that arbitrarily puts the two most dangerous drugs in the "legal" category while most others are in the "illegal" column ... is it really surprising that bored youths will turn to such truly risky "**legal highs**" as choking or inhaling solvents?

This includes displaying, selling, or detailing the use of drug paraphernalia, as well as tips on **legal highs**, such as glue sniffing, the misuse of prescription drugs, or the abuse of other legal substances.

Legal Highs on the Rise Sep 29, 2006 (New Scientist) Recreational drugs that keep you on the right side of the law have never been more popular, but does that mean they're safe?

Chugging cough medicine for an instant high certainly isn't a new practice for teens, who have raided the medicine cabinet for a quick, cheap, and **legal high** for decades.

It is marketed as a "**legal high**" or a substitute for MDMA (Ecstasy) and is sold in solid and liquid forms.

I worked with methodone (heroin replacement) patients, and they more or less split into 2 groups, those who really wanted to get clean, and those who just wanted to go to a nice dark flat and spent time off thier tits without bothering anyone, including the scum dealing H. Of the later category, I'd rather they got a **legal high** and left everyone well alone than being in need of a fix and climbing through my window to pay for it.

It is a semi-synthetic **legal high** that is made from a specialty blend of powerful herbs, aphrodisiacs & chemical compounds that are strong but at the same time very safe to take, with no side effects.

2.3 Analysis: Lexicographic equivalence types and semantic categories

What are we to make of all this in terms of the equivalence types presented above (1.2) and the strategies for dealing with difficult cases (1.3)? What can semantic

theory contribute to the understanding of the relationships between the Polish word and the four English equivalent candidates? Let us look at the latter again, this time with a view to establishing the lexicographic potential of each one.

2.3.1 Designer drugs

The phraseme *designer drugs* can doubtless function as a translational equivalent of *dopalacze* in some contexts. It might be tempting to argue that it is, in fact, more than that, i.e. a fully-fledged cognitive equivalent, seeing as its definition mentions many of the same characteristics (*differentiae*). But the denotational meaning is not, in fact, quite the same as that of *dopalacze*. As we saw in 2.2.1, there are differences in the extensional range of the two terms (the designer drug ecstasy would not be called a *dopalacz* in Polish), as well as some controversy regarding the legal status of the denotata. Moreover, one intuitively feels that the interlingual correspondence is not perfect, which is probably why the Economist blogger did not use *designer drugs* as the only equivalent, but took advantage of other translational strategies as well.

What can hide behind the intuition that there is something not quite right with the Polish-English pairing? For a start, the compositional (literal) meaning of *designer drugs* is different from the motivation behind *dopalacze*: the English term concentrates on how the substances are created, while the Polish one highlights their effects. Further, no conceptual metaphor motivates the English term. Instead, a substance which acts like a drug is explicitly called a drug, and the name is further modified by a noun specifying the drug's method of creation. In addition to that, the Polish word (in the drug sense) has a touch of informality about it, possibly due to the crudeness of the underlying metaphor, resulting from the conceptual distance between its source and target domains: conceptualising the human body in terms of an engine-driven machine involves a significant trivialisation of the complex processes of our mental and emotional functioning.

The connotations are different, too: unlike the Polish word, which has obvious technical associations, the English phrase alludes to expensive, fashion-conscious lifestyle, by analogy with, e.g., *designer jeans* or *designer label*. This creates a clash between the SL and TL items, since the use of *dopalacze* is by no means restricted to fashionable circles, nor are the drugs so described particularly expensive. Quite the opposite, as a matter of fact: *dopalacze* are universally perceived as a cheap substitute for the real thing. All this suggests that the explanatory power of *designer drugs* as an equivalent of *dopalacze* is relatively low. By contrast, as has already been stated, its insertability potential is considerable: in many contexts it will fit just fine.

It might be worth mentioning here that various authors — linguists as well as translation theorists — have proposed various component factors necessary for talking about the equivalence of two lexical units in two languages. The common elements are always denotational and connotational meaning;

other factors may vary. Two proposals will be looked at below which seem representative of the bulk of the literature on the topic.

Talking about equivalence in phraseology, Günther (1990: 505) says we need to consider: 1) denotation; 2) type of lexical unit; word class and phraseme class; 3) collocability and valency; 4) connotation; 5) constitution of formatives, compositional structure, motivation or imagery, the euphonic means used. Taking this as the basis for comparison, the pair *dopalacze* — *designer drugs* can be seen to fulfil the first and the third criterion. Not the second, because *dopalacze* is a simplex and *designer drugs* a phraseme, and not the last two, because both the connotations and the imagery used are different.

Discussing equivalence in translation studies, Koller (1995: 197-198) recognises denotative, connotative, text-normative, pragmatic, and formal equivalence (involving, respectively, extralinguistic content, connotations, text and language norms, the receiver, and certain formal-aesthetic features of the SL text). Analysed in accordance with his criteria, the pair *dopalacze* — *designer drugs* would exhibit a high degree of denotative equivalence and partial equivalence in all the remaining dimensions, with the lowest degree of correspondence on the connotative plane.

2.3.2 Afterburners

Using *afterburners* as a dictionary equivalent of *dopalacze* would be a case of lexicographic innovation, more precisely, of sanctioning a semantic borrowing. The term is closest to the Polish original in terms of motivation (the same metaphor), as well as formally (a transparent derivative of the verb *to burn*, exact equivalent of Polish *palić*, from which *dopalacz* is derived). It would thus appear to be an ideal equivalent, save for one crucial fact: we lack sufficient evidence for it. The Economist blogger was only able to use it because he/she had first defined it. Otherwise, the readers would have been at a loss as to how to interpret the word. Thus, despite its high lexicographic potential, *afterburners* must be judged the weakest of our candidates, as it is unlikely to be interpreted correctly by an English-speaking person when encountered as a stand-alone, decontextualised equivalent of *dopalacze*.

Had the direction of our equivalence search been the reverse of what it was (i.e. from English to Polish, rather than from Polish to English), the grounds for sanctioning the innovation would naturally have been much stronger. But, given that English has borrowed virtually nothing from Polish so far (save for a handful of culinary terms), introducing *afterburners* as an equivalent in a Polish-English dictionary would hardly have been a judicious move.

2.3.3 Legal highs/Legal intoxicants

Our next candidate, *legal highs*, appears to be an excellent translational equivalent in most contexts. What may be debatable is whether it is already an estab-

lished lexical unit of English. Not only is it not recorded by OED, but in some of the corpus examples quoted (e.g. *cheap and legal highs*) it is still clearly a free combination, while in others its tentative status is signalled by scare quotes. However, the bulk of the corpus evidence suggests that, if not yet a fully-fledged phraseme, it is already well on the way to becoming one, thus fulfilling Zgusta's conditions for a translational equivalent. Equally importantly, *legal highs* agrees with the source-language item in register.

On the down side, it cannot be considered a perfect equivalent any more (though it could at the time the LSW entry was being written), because of the legality aspect, which is explicitly lexicalised in the English phrase. There is another difference in meaning as well: *highs* denotes uppers only, not uppers as well as downers (not to mention other conceivable effects drugs may have). Arguably, this is not a very significant difference, since prototypical intoxicants (and the substances called *dopalacze*) reportedly act as stimulants, at least for the first few hours after being consumed.²¹

As for the motivation behind *legal highs*, it is partly metaphorical and partly metonymic. At the most general level, the relevant conceptual metaphor is GOOD IS UP. This is manifest in the diachronically related adjectival senses:²² "emotionally exalted; elated, merry, hilarious"; then, by specialisation, "excited with drink, intoxicated"; and, analogically, "under the influence of, stimulated by, a drug or drugs". Those adjectival senses, describing the sensations a person experiences (as in *to be high*), gave rise to nominal senses denoting the sensations themselves. Thus, the last adjectival sense mentioned above produced the noun *high* defined as a "euphoric state induced by the taking of a drug or drugs". From that, the novel sense (not yet recorded by OED) must have developed metonymically: the name for the state came to stand for the drug that induces it (EFFECT FOR CAUSE metonymy). All in all, the motivation is only similar to that of *dopalacze* insofar as in both cases we are dealing with figurative expressions.

Our last candidate, *legal intoxicants*, is a free combination, not an established lexical unit of English, and thus at best qualifies as an explanatory equivalent. Since we need an equivalent for the productive (encoding) part of the dictionary, explaining the meaning of the SL item is low on our list of priorities. Moreover, the meaning of the TL phrase is more general than that of the SL word: the substances referred to as *dopalacze* are merely one type of intoxicant (recall the Wikipedia entry discussed in 2.2.3), which makes *legal intoxicants* a superordinate term. On top of all that, there is a clear register mismatch, *legal intoxicants* being more formal.

2.3.4 Implications

Summing up, when translating a Polish text featuring *dopalacze*, or when writing an original English text discussing the phenomenon, any of the four equivalent candidates might do the job. The choice would depend on which property needed stressing: the legal status of the substances, the way they are created, or the effect they are meant to produce when consumed. However,

when deciding upon the best equivalent to be placed in a bilingual Polish-English dictionary, other factors acquire prominence. Considering those factors takes us back to where we started from, i.e. to the various types of reasons behind the lack of perfect interlingual equivalence. Strictly speaking, what we are dealing with here does not fall under any of the three scenarios identified at the beginning of the paper. The imperfect equivalence is not language-structure-based, because *dopalacze* is a noun, and English has nouns. It is not world-based, because the referent (the substances in question) does exist in both cultural settings. It is not mind-based, because both speech communities possess the relevant concepts. Upon inspection, however, it turns out to be closest to that last case, albeit in a rather non-obvious way.

First, in both cultural milieus the concepts are somewhat fluid, with the substances' legal status being a matter of controversy — recall the lack of agreement between various English dictionaries concerning the legality of *designer drugs*, and the fact that in Poland the substances were delegalised in 2010, although the situation may well change again. Secondly, the concepts as such are broadly similar, but not identical: as is evident from their linguistic manifestations, they highlight (make explicit) different aspects of the designated phenomenon. From the preceding discussion, we can see how the actual (figurative) meanings of the Polish and English items are derived, in a fairly transparent manner, from their literal meanings; what differs are the literal meanings themselves. And it is precisely those differences, still available to the native speaker's consciousness — thanks, in no small measure, to the compositionality of *designer drugs*, *legal highs*, and *legal intoxicants* — which are the reason why one or the other English equivalent fits best in a particular context.

It will have become clear by now which of our candidates were judged the most promising as lexicographic equivalents to be placed in LSW. In the end, we opted for a combination of *legal highs* and *designer drugs*, preceding the equivalents by the sense discriminator *narkotyki* (in order to make it clear that that was the intended meaning rather than either of the two technical senses, whose English equivalents our target users can live without).

3. Conclusion

It goes without saying that few lexicographers can afford the luxury of devoting such a great deal of time to the search for a single equivalent. Moreover, when reconstructing the way we proceeded in this particular instance, I have no doubt refined the argument, filling in some details which were only intuitively grasped at the time — as one does when telling a story. This in itself is quite significant. It shows that the metalexigraphic and linguistic tools discussed in the preceding pages can definitely help us understand the choices made *after* we had made them. It is much less certain, however, to what extent they were also helpful in making those choices in the first place — which, after all, was the main rationale behind developing the said tools.

For the moment then, despite the considerably enriched methodological apparatus at the (linguistically sophisticated) lexicographer's disposal, we are still some way from being able to neatly classify, let alone deal with, everything that comes our way in the day-to-day business of bilingual dictionary making. The lexicon of a language — especially when examined from the point of view of the lexicon of another language, with which it is to be matched pairwise, item by item — proves too complex and elusive to fit comfortably into the pigeonholes metalexicographers have painstakingly constructed for it. Still, I see no principled reason why we should not get there one day.

Notes

1. See, for instance, the discussion of Zulu *tokoloshe* and similar cases in Swanepoel (2005).
2. See Geeraerts and Grondelaers (1995) for a detailed discussion.
3. According to Kövecses (2000: 162), although Japanese shares with English the underlying conceptual metaphor ANGER IS A HOT FLUID IN A CONTAINER, "in addition to the body as a whole, the stomach/bowels area (called *hara* in Japanese) is seen as the principal container for the hot fluid that corresponds to anger".
4. According to Yu (1995: 59), Chinese tends to utilise more body parts in its metaphors of anger and other emotional states. This allegedly has its roots in the theories of *yin-yang* and of the five elements of Chinese medicine. The Chinese counterpart of anger (*nu*) is based on the culturally significant notion of *qi* — energy that is conceptualized as a gas or fluid flowing through the body. When an excess thereof is produced, anger is felt (Kövecses 2000: 163).
5. According to Taylor and Mbense (1998: 192), "Zulus characterise anger, not only in terms of the impairment of rational faculties, but also in terms of nausea, illness, perspiration, and hyperactivity. (...) a number of Zulu expressions highlight the ferociousness of anger, and the devastation it may cause; while these aspects are certainly not absent from the English conceptualisation, they appear to be perspectivised more frequently, and more forcefully, in Zulu than in English".
6. According to Lutz (1987: 292), the inhabitants of Ifaluk "define, explain, and understand emotions primarily by reference to the events and situations in which they occur. This aspect of their ethnotheory of emotion contrasts with our own emphasis on the internal and private, rather than the social, nature of emotion". For instance, "the pragmatic information encoded in the term *song* (justifiable anger) includes the notion that judgements about whether a particular act constitutes a rule violation are more aptly made by people of higher social status, including the chiefs and older individuals" (Lutz 1987: 297).
7. For an accessible summary of Locke's views on the origin of ideas, see Harris and Taylor (1997: Chapter 10) or Hüllen (2004: 164ff.).
8. The terms in parentheses are the most common alternatives found in the relevant metalexicographic sources; a list of the latter can be found in Adamska-Salaciak (2010).
9. Actually, Madiba and Nkomo (2010: 319) treat this as an instance of explanation in the target language, not an explanatory equivalent. They do not supply any examples which might suggest what they take an explanatory equivalent to be or how it is supposed to differ from their "explanation in the target language".

10. This seems to correspond to Gouws and Prinsloo's (2010) surrogate equivalent.
11. This is sometimes called a 'cultural equivalent', but note that the correspondence here is mainly referential: it is not, or at least not primarily, the lexical items that can be treated as each other's approximate cultural equivalents, but the real-world institutions which those items denote.
12. This last strategy, no doubt the most controversial of those mentioned, is discussed in detail in Adamska-Salaciak (2006: 133-145).
13. Naturally, if they happen to be cognitive equivalents at the same time, so much the better.
14. The first edition of LSW was published in 2004, when *dopalacze* was not yet a household word. To the best of my knowledge, the underlying problem did not exist, either.
15. Singular use is possible, but much less frequent.
16. Following Hanks (1987), the term *displacement* could be used to refer to this definitional strategy.
17. For an introduction to Conceptual Metaphor Theory, see Lakoff and Johnson (1980, 1999).
18. It should, perhaps, be added that the noun phrase *designer drugs* is also attested in corpora in another sense, that of (as yet non-existent) drugs produced by new technologies and expected to combat diseases against which today's drugs are powerless. In this sense, it is often modified by the adjectives *utopian* or *futuristic* and occurs in the vicinity of *genetic engineering*, *gene therapies*, and similar.
19. All the English examples quoted in this paper have been obtained using the Sketch Engine corpus query system.
20. These are exclusively instances in the relevant sense. Cases such as, e.g., *legal high priests*, *hold the legal high ground*, or *legal high fliers*, have been discarded.
21. The author cannot vouch for the accuracy of this information (gathered from different discussion fora).
22. All the definitions in this paragraph come from OED.

References

Dictionaries

- AHD = *The American Heritage Dictionary of the English Language*. Boston: Houghton Mifflin Company.
- CALD = *Cambridge Advanced Learner's Dictionary* online (<http://dictionary.cambridge.org/>)
- CED = *Collins English Dictionary* online (<http://www.collinslanguage.com/>)
- DSA = Branford, J. and W. Branford (Eds.). 1991. *Dictionary of South African English*. Fourth edition. Cape Town: Oxford University Press.
- LDOCE = *Longman Dictionary of Contemporary English* (<http://www.ldoceonline.com/>)
- LSW = Fisiak, J., A. Adamska-Salaciak, M. Idzikowski, E. Jagła, M. Jankowski and R. Lew. 2011. *LONGMAN Słownik Współczesny Angielsko-Polski, Polsko-Angielski*. Second edition. Harlow: Pearson Education Limited.
- MEDAL = *Macmillan English Dictionary for Advanced Learners* online (<http://www.macmillandictionary.com/>)

OALD = *Oxford Advanced Learner's Dictionary* online (<http://www.oxfordadvancedlearnersdictionary.com/>)

OED = *Oxford English Dictionary* online (www.oed.com/)

Corpora

BASE (British Academic Spoken English Corpus; 1,252,256 tokens)

BAWE (British Academic Written English Corpus; 8,336,262)

BNC (British National Corpus; 112,181,015 tokens)

COCA (Corpus of Contemporary American English; ca. 425,000,000 tokens)

enTenTen (English Web Corpus; 3,268,798,627 tokens)

ukWaC (British English Web Corpus; 1,565,274,190 tokens)

Other references

Adamska-Salaciak, A. 2006. *Meaning and the Bilingual Dictionary: The Case of English and Polish*. Frankfurt am Main: Peter Lang.

Adamska-Salaciak, A. 2010. Examining Equivalence. *International Journal of Lexicography* 23(4): 387-409.

De Schryver, G.-M. 2009. The Lexicographic Treatment of Ideophones in Zulu. *Lexikos* 19: 34-54.

Farina, D. 1996. The Bilingual Lexicographer's Best Friends. *Lexicographica* 12: 1-15.

Geeraerts, D. and S. Grondelaers. 1995. Looking Back at Anger: Cultural Traditions and Metaphorical Patterns. Taylor, J.R. and R.E. MacLaury (Eds.). 1995. *Language and the Cognitive Construction of the World*: 153-180. Berlin: Mouton de Gruyter.

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

Gouws, R.H. 2002. Equivalent Relations, Context and Cotext in Bilingual Dictionaries. *Hermes* 28: 195-209.

Gouws, R.H. and D.J. Prinsloo. 2010. Surrogaatekwivalensie in tweetalige woordeboeke met spesifieke verwysing na zero-ekwivalensie in Afrikataalwoordeboeke. *Tydskrif vir Geesteswetenskappe* 50(4): 502-519.

Günther, K. 1990. Äquivalenzbeziehungen in der Phraseologie. *Zeitschrift für Slawistik* 35: 504-509.

Hallsteinsdóttir, E. and K. Farø. 2010. Interlinguale Phraseologie: Theorie, Praxis und Perspektiven. *Yearbook of Phraseology* 1: 125-158.

Hanks, P. 1987. Definitions and Explanations. Sinclair, John M. (Ed.). 1987. Looking Up: An Account of the COBUILD Project in Lexical Computing and Development of the Collins COBUILD English Dictionary: 116-136. London/Glasgow: Collins ELT.

Harris, R. and T.J. Taylor. 1997 [1989]. *Landmarks in Linguistic Thought I: The Western Tradition from Socrates to Saussure*. Second Edition. London: Routledge.

Hohulin, E.L. 1986. The Absence of Lexical Equivalence and Cases of Its Assymetry. *Lexicographica* 2: 43-52.

Hüllen, W. 2004. *A History of Roget's Thesaurus: Origins, Development, and Design*. Oxford: Oxford University Press.

- Koller, W.** 1995. The Concept of Equivalence and the Object of Translation Studies. *Target* 2: 191-222.
- Kövecses, Z.** 2000. The Concept of Anger: Universal or Culture Specific? *Psychopathology* 33: 159-170.
- Lakoff, G. and M. Johnson.** 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lakoff, G. and M. Johnson.** 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Locke, J.** 1689/90 [1975]. *An Essay Concerning Human Understanding*. (Edited by P.H. Nidditch.) Oxford: Clarendon Press.
- Lutz, G.** 1987. Goals, Events, and Understanding in Ifaluk Emotion Theory. Holland, D.C. and N. Quinn (Eds.). 1987. *Cultural Models in Language and Thought*: 290-312. Chicago: University of Chicago Press.
- Madiba, M. and D. Nkomo.** 2010. The *Tshivenda-English Thalusamaipfi/Dictionary* as a Product of South African Lexicographic Processes. *Lexikos* 20: 307-325
- "Questions of Substance", Oct 8th 2010, 10:38 by J.P. (http://www.economist.com/blogs/easternapproaches/2010/10/legal_highs_poland) (date of access: 15.10.2010.)
- Swanepoel, P.** 2005. On Defining "Imaginary" Beings and Attributes: How Do Lexicographers Cope with Culturally Determined Differences in Beliefs about Cosmology, Ontology and Epistemology? *Lexikos* 15: 179-195.
- Taylor, J.R. and T.M. Mbense.** 1998. Red Dogs and Rotten Mealies: How Zulus Talk about Anger. Athanasiadou, A. and E. Tabakowska (Eds.). 1998. *Speaking of Emotions: Conceptualisation and Expression*: 191-226. Berlin: Mouton de Gruyter.
- Yu, N.** 1995. Metaphorical Expression of Anger and Happiness in English and Chinese. *Metaphor and Symbolic Activity* 10: 223-245.
- Zgusta, L.** 1971. *Manual of Lexicography*. The Hague: Mouton.
- Zgusta, L.** 1987. Translational Equivalence in a Bilingual Dictionary: *Bahukošyam. Dictionaries* 9: 1-47. Reprinted in Zgusta, L. 2006. *Lexicography Then and Now: Selected Essays*: 236-261. Tübingen: Niemeyer.

National Lexicography Units: Past, Present, Future

Mariëtta Alberts, *Research Unit for Languages and Literature in SA Context,
Northwest University, Potchefstroom Campus (malberts@lantic.net)*

Abstract: This article deals with the national dictionary offices of the previous bilingual dispensation, the eleven official national dictionary offices in the present multilingual dispensation, and the future prospects of these offices. It discusses the past dispensation in terms of the need and reasons for the establishment of national dictionary offices, i.e. national lexicography units (NLUs). Attention is given to the prescripts of the National Lexicography Units Bill (1996) for the establishment of NLUs, as well as the transfer of these units from the Department of Arts, Culture, Science and Technology to the Pan South African Language Board. The restructuring of dictionary units that existed prior to the multilingual dispensation is considered, together with the establishment of new dictionary units for the official African languages. The present situation is dealt with by describing the status quo at the NLUs in terms of housing, administration, funding, management, training, computerisation, cooperation, production and the like. The article concludes with some questions and reservations about the future of the NLUs, followed by a number of apposite recommendations.

Keywords: BILINGUAL DISPENSATION, CULTURAL DIVERSITY, GENERAL DICTIONARY, LEXICOGRAPHY, MONOLINGUAL DICTIONARY, MULTILINGUALISM, NATIONAL DICTIONARY, NATIONAL LEXICOGRAPHY UNITS, OFFICIAL LANGUAGE, TERMINOLOGY, TERMINOLOGY

Opsomming: Nasionale leksikografiese eenhede: Verlede, hede, toekoms. Hierdie artikel handel oor die nasionale woordeboekkantore tydens die voormalige tweetalige bedeling, die huidige meertalige bedeling waarin kantore vir die elf amptelike nasionale woordeboeke funksioneer, en die toekoms van hierdie kantore. Aspekte van die vorige bedeling word bespreek ten opsigte van die behoefte aan en redes vir die stigting van nasionale woordeboekkantore oftewel nasionale leksikografiese eenhede (NLEe). Die soeklig val op voorskrifte van die wetsontwerp oor nasionale leksikografiese eenhede (1996) vir die stigting van sodanige eenhede, sowel as op hul oordrag van die Departement van Kuns, Kultuur, Wetenskap en Tegnologie na die Pan-Suid-Afrikaanse Taalraad. Verder fokus die artikel op die herstrukturering van daardie woordeboekeenhede wat voor die meertalige bedeling bestaan het en die stigting van nuwe woordeboekeenhede vir die amptelike Afrikatale. Die huidige bedeling word ondersoek en die praktiese situasie by die NLEe word bespreek ten opsigte van behuising, administrasie, befondsing, bestuur, opleiding, rekenarisering, samewerking, produksie en dergelike. Die artikel eindig met vrae en bedenkinge oor die toekoms van die nasionale leksikografiese eenhede, gevolg deur enkele aanbevelings.

Sleutelwoorde: ALGEMENE WOORDEBOEK, AMPTELIKE TAAL, EENTALIGE WOORDEBOEK, KULTURELE DIVERSITEIT, LEKSIKOGRAFIE, MEERTALIGHEID, NASIONALE LEK-

SIKOGRAFIESE EENHEDE, NASIONALE WOORDEBOEK, TERMINOGRAFIE, TERMINOLOGIE, TWEETALIGE BEDELING

1. Introduction

The Memorandum on the Objects of the National Lexicography Units Bill, 1996 (DACST 1996: 16) states that "dictionaries are a mirror of a country's social, cultural, economical, scientific and technological development". With the onset of the multilingual dispensation in South Africa, indications were that the language communities of all the official languages needed national dictionaries — not only to bridge the apparent communication gap between communities, but also to document the rich variety of South African languages.

The compilation of national monolingual general dictionaries would also preserve South Africa's linguistic diversity in all its forms, regardless of social, economic, political, demographic or linguistic status. The envisaged dictionaries could assist with improving citizens' understanding of the diversity of culture and language through research, advocacy, public education and community support.

Language services such as translation and interpreting, as well as institutions such as schools, tertiary institutions, libraries, the media and publishers could benefit from the national dictionary products and language services provided by dictionary units.

The proposal for the establishment of national dictionary offices was completely in line with the spirit of the Constitution (DAC 2002) in that it made equitable provision for national monolingual general dictionaries for all the official languages of South Africa with a view to the eventual promotion of multilingualism.

At the time (1995) of the drafting of the National Lexicography Units Bill, the major dictionary projects in South Africa all had different time schedules, venues and aims and it was difficult to align the projects and lexicographical activities. National lexicography (dictionary) units had to be established for all the official languages in order to create a better alignment of the State-funded part of these national dictionary projects. These units would comply both within themselves and in the broader national perspective with the goals set out in the national language policy regarding multilingualism (DACST 1996: 16).

There was (and still is) an apparent scepticism towards and total disregard for multilingualism at some institutional levels in South Africa. These sceptics might find positive pointers to the attainment of multilingualism in the endeavours of official national dictionary units. National Lexicography Units (NLUs) were namely established and managed "to make equitable provision for national general monolingual dictionaries for each of the official languages of South Africa and for matters incidental thereto" (DACST 1996: 2; Alberts 1996; DAC 2002: 7).

Political goodwill and government resources evidently had to be mobilised towards the operationalisation of constitutional multilingualism by estab-

lishing national lexicographical units. The question that remains, however, is: "How successful are the NLUs in their contribution towards the development of the official languages within the multilingual dispensation and do they have a future?"

This article addresses the issues related to the official dictionary compilation endeavours in the period before and during the establishment of NLUs, the motivation for the establishment of these units, the establishment process itself, and the current situation where a national dictionary office was established for each of the official languages to manage the lexicographical data of these languages. It concludes with some conjecture about what the future holds for these units. The article therefore aims to give a critical historical account of the NLUs and to consider possible future scenarios for the management of institutionalised lexicographic processes.

2. Past

2.1 Background

It is the democratic right of every citizen to be empowered in any language of his/her choice in order to make a meaningful contribution to political and socio-economic activities (DACST 1996: 16; DAC 2002). The new political dispensation after 1994 thus urged the Department of Arts, Culture, Science and Technology (DACST) to make a constructive contribution to reconciliation, nation building, reconstruction and development in South Africa. The Department had no alternative but to establish NLUs to govern and manage the documentation of the vocabulary of the official languages of South Africa by giving financial support to various national dictionary projects in South Africa. By planning, managing and financing national dictionary projects, various lexicographical activities could be coordinated. DACST wanted to accomplish this in close cooperation with the Pan South African Language Board (PanSALB).

Language is a collective interest of the whole community and involves the communication needs of every person in South Africa. By establishing and funding national monolingual dictionaries, the Department could develop an umbrella lexicography policy for all official languages. At the time it was of critical importance to formulate a policy to meet the general need for a properly funded and managed lexicographical practice for the national monolingual general dictionary projects of South Africa (Alberts 1996; 1997).

A national monolingual general dictionary deals with the lexical inventory of a specific language over a broad spectrum of the vocabulary in order to describe that vocabulary in detail. This type of dictionary usually contains the general words (e.g. vernacular, common words, literary words, etymology, archaic words, dialectal variants, slang, vulgar words, core terminology of various subject areas, etc.) of the vocabulary of a specific language needed by speakers of the language needed to communicate. Such a dictionary does not

only document the language, but also preserves it. The dictionary compilation process usually takes a number of years; the office has a permanent staff and the project is funded externally by either the State or a large publisher.

In the previous bilingual dispensation the State, through DACST, financed only two national dictionary units, namely the national Afrikaans Dictionary at the Bureau of the "*Woordeboek van die Afrikaanse Taal*" (WAT) in Stellenbosch and the national English Dictionary at the *Dictionary Unit for South African English* (DSAE) in Grahamstown.

Although the Department did not then finance any of the general dictionaries for the African languages, some public funds were made available to support these dictionary projects indirectly through university funding at various universities. Ongoing projects in an African language at that stage were the Xhosa dictionary project (English/Xhosa/Afrikaans) at the University of Fort Hare, the Zulu dictionary project at the University of Zululand, Tswana dictionary at the University of the North and a Northern Sotho dictionary project at the University of Pretoria. There were also a few private initiatives and projects of publishers regarding the compilation of general dictionaries for the indigenous South African languages. None of these projects were, however, state funded. In the post-1994 dispensation this situation of unequal government support was considered unacceptable (Alberts 1996; 1997).

Since the major dictionary projects at the time had different time schedules, venues and aims, it was difficult to align them. The Afrikaans, English and Xhosa dictionaries were ongoing projects with fixed offices, assets and staff, while those for Zulu, Northern Sotho and Tswana were university projects. A South Sotho project was also in a planning phase. All these were ongoing projects planned to last for several years. Although mostly university projects, it was clear that they also needed State funding. Two national dictionary projects planned for Swati and Tsonga respectively would probably be government projects and therefore require State funding. A need was also expressed for a Ndebele national dictionary, but no project plan had yet been submitted at that stage. It was also known that the Tshivenda Language Board was working on an English-Venda-Afrikaans glossary, although no response was received on the project during the needs assessment study that was conducted by DACST. Since this was a multilingual project it could not be considered as a national project (Alberts 1996).

In view of the constitutional provisions for multilingualism, the development of all the official languages, and the lack of alignment of State-funded dictionary projects, it was necessary to make equitable provision for the operating of national dictionary offices for all these languages. Such a dispensation required legislation along the lines of the Act governing the *Woordeboek van die Afrikaanse Taal* (WAT) (Act 50 of 1973, with its Amendment Acts, 9 of 1986 and 22 of 1991). This could be achieved by either amending the WAT Act or by drafting completely new legislation (DACST 1996: 16).

In 1995 DACST, funder of the WAT and DSAE dictionary projects, decided to draft legislation to establish national lexicography units for all the offi-

cial languages. The ensuing Bill made provision for dictionary units for all the official languages. The linguistic heritage of the country could be preserved and further developed in these units seeing that general lexicography comprehensively describes the vocabulary of a given language by documenting it. The Bill was supposed to bring about change and transformation.

By supporting monolingual dictionary projects, the policy of multilingualism was also maintained since it involved the documentation of all the official indigenous languages. By planning, managing and financing national dictionary projects, DACST intended to coordinate lexicographic activities in South Africa and eliminate the possible duplication of work. Thus, DACST aimed to finance the dictionary units (as was already the case with Afrikaans and English) and to provide recurrent State financing to the NLUs. The envisaged establishment of the units was advertised widely in the media. Media coverage would ensure the correct positioning of the dictionary units to be established and attract appropriate collaborators to work on these projects.

Various bodies such as the Language Task Group (Langtag), the Langtag Subcommittee for the Development of the (South) African Languages, the Pan South African Language Board (PanSALB), the African Association for Lexicography (Afrilex), the Board of the "Woordeboek van die Afrikaanse Taal" (WAT), the Portfolio Committee on Arts, Culture and Language, Science and Technology, the Select Committee on Education, Sport and Recreation, Arts, Culture, Science and Technology, the Task Group appointed to revise the WAT Act, the departmental legal adviser, departmental consultants and 98 stakeholders in the field of lexicography were consulted, and various needs assessment studies were conducted (Alberts 1996).

It was obvious from the reaction of stakeholders that responded to the proposed revision of the existing WAT Act that they all welcomed the establishment of the envisaged National Lexicography Units. The respondents were all satisfied with the idea of an umbrella Act to govern general monolingual dictionary projects for the official languages of South Africa. The Task Group appointed to revise the WAT Act discussed the suggestions made by respondents and these suggestions were incorporated in the Bill (Alberts 1996).

The Department (DACST) subsequently recommended that the existing WAT Act (Act 50 of 1973, with its Amendment Acts, 9 of 1986 and 22 of 1991) be repealed and replaced by a new Act (the National Lexicography Units Bill) to govern national lexicography (dictionary) units for the official languages (Alberts 1996).

2.2 The National Lexicography Units Bill

In 1995 the National Lexicography Units Bill was drafted as umbrella legislation to incorporate the existing WAT and DSAE dictionary units, and to make provision for the establishment of national dictionary units for each of the official South African languages (Alberts 1996; DACST 1996). According to the Bill,

each national lexicography (dictionary) unit would be managed by an Editor-in-Chief who would be accountable to a Board of Directors as management committee. The Board of Directors (BoD) was to be appointed by the Minister of Arts, Culture, Science and Technology (ACST) in consultation with stakeholders and it would in turn be accountable to the Minister of ACST. The individual BoD of each NLU would appoint the Editor-in-Chief for its NLU as well as the staff. The BoD also had the power to dismiss any person appointed. The NLUs would be managed as section 21 companies and although financed by the State, they would be able to generate their own funds. The NLUs would furnish DACST with annual and financial reports (Alberts 1996; 2003).

Every NLU would have a fixed staff component and the State undertook to finance the dictionary projects as long as possible or necessary (Alberts 1996; 1997).

The Bill stipulated that the NLUs should be situated at tertiary institutions in the relevant geolinguistic area where the bulk of first-language speakers reside. The tertiary institutions could provide the necessary infrastructure and administrative facilities while the academic staff teaching the respective languages could be included in the editorial committees. The involvement of linguists would bring the dictionary products in line with the overall development and standardisation of the respective languages. First-language speakers in the geolinguistic areas concerned would obviously be stakeholders in the dictionary compilation process and would be available for fieldwork and the gathering of linguistic data (Alberts 1996).

Although the use of the official languages is not limited to specific geolinguistic areas, it was decided not to relocate existing decentralised dictionary units such as the WAT and DSAE since this would disrupt and even destroy infrastructure and support systems that were developed over time (Alberts 1996). In an era where computers rule the world, a coordinated approach towards the compilation of dictionaries does not require centralisation of dictionary offices, resources or facilities. In fact, centralisation could well detach dictionary making from the speech communities concerned and undermine the independence, resources and skills of minor language groups and under-resourced regions (Alberts 1996; 1997).

In order to develop the necessary capacity to achieve lexicographical goals such as the compilation, editing and publishing of dictionaries, the staff appointed at the WAT and DSAE, as well as the newly appointed staff of the new dictionary units required training pertaining to lexicographical principles and practice. Several institutions, bodies or associations skilled in lexicographical principles and practice, e.g. universities presenting courses in lexicography, the African Association for Lexicography (Afrilex) and existing dictionary units (e.g. WAT and DSAE) indicated that they could provide the required training to staff of the envisaged dictionary units (Alberts 1996).

The National Lexicography Units Bill served before the Cabinet's Committee on Social and Administrative Affairs on 12 June 1996 and was summarily approved. The Bill then served before Cabinet on 19 June 1996, and input

from the Chief State Law Adviser was received on 18 September 1996. Officials from DACST's National Terminology Services (NTS) met with the Portfolio Committee: Arts, Culture and Language, Science and Technology on 14 February 1997 to discuss suggestions made by the Portfolio Committee. Once these suggestions had been incorporated, the Bill was ratified (Alberts 1996; DACST 1997).

2.3 PanSALB and the tabling of the National Lexicography Units Bill

The former Minister of Arts, Culture, Science and Technology, Dr Ben Ngubane, requested the revision of legislation pertaining to the Afrikaans dictionary project at the Bureau of the "Woordeboek van die Afrikaanse Taal" with a view to covering the governance of the envisaged national lexicography units (NLUs) for the official South African languages. The aim with the new legislation was also to ensure greater equity in allocating resources to the different NLUs (DACST 1996; DACST 1997).

After wide consultation with various stakeholders in the field of lexicography (see § 2.1), a Bill — the National Lexicography Units Bill — was drafted by the Task Group appointed for this task (see § 2.2). It was obvious from the reaction of respondents to the questionnaires and working documents related to the Bill that they were looking forward to the establishment of the envisaged NLUs. As mentioned earlier, the respondents were all satisfied with the idea of an umbrella Act to govern national general dictionary projects for the official South African languages (DACST 1997).

During discussions in the Minister's office regarding possible amendments to the Pan South African Language Board Act (Act 59 of 1995), the future process regarding the establishment of the NLUs was again raised. The discussions focused mainly on the future role of PanSALB, but various issues were considered — among others the relationship between PanSALB and the envisaged lexicography units (DACST 1997).

It was recommended that the Minister consider the relationship between PanSALB and the envisaged NLUs, and discuss this matter with the Executive and Board of PanSALB. It was further recommended that the issue be discussed with the Chairperson of the Portfolio Committee: Arts, Culture and Language, Science and Technology. Political binding had to be received before the end of April 1997, since the policy review of the whole language issue was at stake (DACST 1997).

In 1998 DACST revised and amended the Pan South African Language Board Act, Act 59 of 1995 and one of the major proposals to these amendments was that the Board "shall in the manner prescribed by notice in the Gazette and the provincial Gazette establish NLUs for each of the official languages of South Africa" (PanSALB 1998: 14). The PanSALB Board did not approve the proposal from DACST and suggested that the National Lexicography Units Bill

be resubmitted to be passed as an Act of Parliament as a matter of urgency (PanSALB 1998: 14).

The amended PanSALB Act (PanSALB Amendment Act 1999), however, placed on PanSALB the responsibility to establish "national lexicographic units that will operate as companies limited by guarantee under section 21 of the Companies Act No 61 of 1973" (PanSALB 1999: 28). The articles of the National Lexicography Unit Bill were incorporated as resolutions into the revised PanSALB Act, and the NLUs were to be established as structures of the Pan South African Language Board.

In terms of the PanSALB Amendment Act, 1999, Section 8 (2) (c), the Board had to establish the following (PanSALB 1999: 19; 2000: 26):

"Units to operate as companies limited by guarantee under section 21 of the Companies Act, 1973 (Act 61 of 1973) and shall allocate funds to the units for the fulfillment of their functions. Provided that the memorandum and articles of association of such units shall include the following principles:

- (i) The unit is accountable to the Board for the moneys allocated to it.
- (ii) The unit shall abide by the policies of the Board.
- (iii) The unit shall adhere to the principles of promoting language development."

The Bureau of the *Woordeboek van die Afrikaanse Taal* (WAT) operated in terms of the "Woordeboek van die Afrikaanse Taal" Act, 1973 (Act 50 of 1973). The Bureau of the WAT had been registered as a section 21 company to await the promulgation of the repealed WAT Act, after which it ceased to function as a statutory body and had the same status as the other NLUs in the country (PanSALB 2001: 22).

The Dictionary of South African English (DSAE) had been operating as a section 21 company since 1991 and functioned under the auspices of PanSALB since April 2001 (PanSALB 2001: 22).

The subsidies that DACST used to pay to the Bureau of the WAT and the DSAE were thenceforward to be handled by PanSALB. An equitable distribution of subsidies to all NLUs countrywide would be pursued over the next ten years (PanSALB 2001: 22).

2.4 The establishment of the National Lexicography Units (NLUs)

PanSALB realised the need for coordinating all the isolated efforts in the country so as to formulate a strategy for lexicography and terminology development in respect of all the official languages in the country, particularly the previously marginalised languages. On 31 October 1997 a consultative meeting of stakeholders was held in Johannesburg. Two follow-up meetings took place in 1998: the first on 19 and 20 March for those languages with existing dictionary

units, and the second on 14 and 15 May for those languages without dictionary units. Norms and regulations for the establishment and recognition of lexicography units were discussed with delegates on these occasions and the reports on the proceedings of these meetings were published by PanSALB (PanSALB 1998: 11).

A meeting for existing dictionary units was held from 25 to 27 August 1999 in Pretoria to develop a strategic planning framework for each unit and to do lexicographic planning (PanSALB 1999: 20).

These processes assisted the Board in the effective handling of its new responsibility of establishing lexicographic units and actions were developed to facilitate the establishment of such units. The norms and regulations for establishing lexicographic units were translated into all the official languages and gazetted by government notice to facilitate this process (PanSALB 1999: 28).

The main purpose of the units would be to "conserve, preserve, research and document the official languages concerned, by compiling a monolingual explanatory dictionary and such other dictionaries as may be required to satisfy the needs of the target users of that language" (PanSALB 2000: 26).

Consultative meetings were held in 1999 with various stakeholders and role-players from the different official languages with the aim of informing them of the nature of section 21 companies, the main purpose of the dictionary units, and to elect founder members (PanSALB 2000: 26).

By March 2001 a national lexicography unit (NLU) for each of the official languages in South Africa had been established and registered as a section 21 company.

2.5 Training workshops for NLUs

As mentioned above, between August and November 1999 PanSALB held four workshops for those languages that already had lexicography units, as well as for those languages for which units had to be established. The workshops were attended by staff already working in dictionary offices and other people involved in lexicography projects on an individual and part-time basis, such as university lecturers who were compiling dictionaries (PanSALB 2000: 26).

The training sessions dealt with strategic and lexicography planning. Vision and mission statements were formulated and guidance was given for the setting of lexicographic goals and the formulation of action plans. Training also focused on the computerisation of dictionary compilation processes at the dictionary units (PanSALB 2000: 26).

Expert lexicographers from within South Africa and abroad dealt with topics such as the following (PanSALB 2000: 26):

- Different types of dictionaries and their nature
- Data collection policies and the creation of a database

- Inclusion policies
- Lexicography processes and data types
- Dictionary design and layout
- Computers and computer literacy in lexicography
- Computer programs (software) and corpus building

3. Present

3.1 NLUs: Accommodation and venue

The National Lexicography Units Bill stipulated that national lexicography units should be established at a tertiary institution in the geolinguistic area where the most first-language speakers of the relevant language reside (National Lexicography Units Bill 1996). At the different consultative meetings held during 1997 and 1998 it was decided that "the units should be established at or close to a tertiary institution where the language concerned is taught and where there is a concentration of mother-tongue speakers" (PanSALB 2000: 26-27).

In 1999 PanSALB held meetings with the following universities and institutions to secure accommodation for NLUs:

- IsiXhosa: the University of Fort Hare (the Greater Xhosa Dictionary project > isiXhosa National Lexicography Unit), Alice — NLU registered as a section 21 company in 2001.
- IsiZulu: the University of Zululand (Isikhungo Sesichazamazwi SesiZulu), Umlazi Campus — project established in 1981. NLU registered as a section 21 company in 2000.
- Sesotho sa Leboa: the University of the North (Sepedi Dictionary Unit > Sesotho sa Leboa Dictionary Unit), Pietersburg/Polokwane — NLU registered as a section 21 company in 2000.
- Xitsonga: Tivumbeni College of Education (Yuniti Ya Lekisikhogirafi ya Xitsonga), Tzaneen — NLU registered as a section 21 company in 2001.
- Setswana: University of Northwest (Sefala sa Setswana), Mmabatho — NLU registered as a section 21 company in 2001.
- Siswati: Pretoria Technikon Nelspruit Campus (Silulu Sesiswati Lexicography Unit), Nelspruit (PanSALB 2000:27) — NLU registered as a section 21 company in 2001.

Accommodation was at the time (1999) (PanSALB 2000: 28; 2001: 25) also secured for

- Afrikaans: the University of Stellenbosch (*Die Woordeboek van die Afrikaanse Taal* (WAT)), Stellenbosch — c. 1930, NLU registered as a section 21 company in 2000.
- English: Rhodes University (Dictionary of South African English), Grahamstown — c. 1969, in existence as a section 21 company since 1991.
- IsiNdebele: the University of Pretoria (IsiNdebele Dictionary project), Pretoria — NLU registered as a section 21 company in 2000.

Accommodation was later secured (PanSALB 2001:25) for

- Sesotho: University of the Free State (Sesiu sa Sesotho NLU), Bloemfontein — NLU registered as a section 21 company in 2001.
- Tshivenda: University of Venda (Tshivenda NLU), Thohoyandou — NLU registered as a section 21 company in 2001.

3.2 NLUs: Management structures

The NLUs at the outset consisted of two management structures: members and directors. Members of the companies were elected during consultative meetings, e.g. annual general meetings. Directors were nominated by institutions, structures and bodies in terms of article 41 of the articles of association (PanSALB 2000: 27).

The Boards of Directors (BoDs) that controlled and managed the NLUs consisted of eleven directors each. They appointed the editors-in-chief and the requisite staff members who compiled the dictionaries and managed the offices (PanSALB 2000: 27). The Boards of Directors were also responsible for the development, adaptation and implementation of strategic planning frameworks for all NLUs (PanSALB 2000: 30).

PanSALB developed guidelines for the reporting by an NLU to its Board of Directors, the reporting of NLU Boards of Directors to PanSALB, and the reporting of PanSALB to DACST (PanSALB 2000: 30). Guidelines were provided regarding progress and financial reports.

3.3 NLUs: Memorandum and articles of association

Every NLU operating as a section 21 company was to be governed by two constitutive documents:

- A Memorandum of Association (MoA) — the founding document of a company to which the founder members are signatories
- Articles of Association (AoA) — the document that determines the manner in which a company is supposed to function (PanSALB 2000: 27)

These documents were drafted by PanSALB in a form adaptable to the unit-specific characteristics of each unit and language group. The norms and regulations for the establishment of the NLUs developed at the consultative meetings were integrated with the constitutive documents of the companies and replaced by the MoA and AoA in accordance with the Companies Act, 1973 (Act 61 of 1973). These documents were continually being reworked and redesigned during the process of the establishment of the NLUs in an effort to take unit-specific factors into account (PanSALB 2000: 27).

In 1999 PanSALB appointed a practising attorney to draft powers of attorney and to lodge the prescribed forms with the Office of the Registrar of Companies in terms of the Companies Act once the companies (NLUs) were ready for registration (PanSALB 2000: 27).

3.4 NLUs: Funding and financial management

Since the establishment of the national lexicography units, PanSALB has been continually involved in negotiating funding for the NLUs with DACST and the National Treasury (PanSALB 2001: 23). Liaison and negotiations regarding funding for NLUs were handled through DACST (PanSALB 2000: 30).

PanSALB devised a ten-year plan for the attainment of funding equity between the NLUs. This plan was submitted to the Minister of Arts, Culture, Science and Technology for approval (PanSALB 2001: 23).

Funding is released on a monthly basis to the units upon receipt of a business plan and a budget. These are based on strategic planning documents compiled during workshops with role-players. Representatives from the units received guidance from PanSALB with regard to strategic planning focusing on short-, medium- and long-term goals, environmental analysis, future strategic positions, plans of action, etc. (PanSALB 2001: 23).

3.5 NLUs: Staff

Three NLUs, i.e. the Bureau of the WAT, the Xhosa Dictionary Project and the Zulu Dictionary Unit, initially operated as either statutory bodies funded by government (Bureau of the WAT) or as projects at tertiary institutions (the Xhosa and Zulu projects). The transitional provisions in the Repeal WAT Act made provision for the transfer of the staff of the Bureau of the WAT as a statutory body to the Bureau as a section 21 company (PanSALB 2001: 23). In the case of the Xhosa and Zulu projects the situation was more complicated since their personnel were employed and paid by the tertiary institutions that funded these projects (i.e. the universities of Fort Hare and Zululand respectively). These staff members had to enter into new service contracts with the section 21 companies, and PanSALB assisted the units and their staff in negotiating service benefits (PanSALB 2001: 23).

The newly established NLUs appointed staff members and offered employees fixed-term contracts. In these cases PanSALB provided the Boards of Directors with guidance on the organisational structure of an NLU, job specifications, and the terms and conditions of appointment (PanSALB 2001: 23-24).

3.6 NLUs: Administrative support, planning and functioning

Staff members in PanSALB's Lexicography and Terminology Development focus area managed the NLUs in terms of measuring progress with dictionary production and assisting with capacity building to strengthen the operation of these units. The focus area supplied administrative, training, financial and technical support to the NLUs. Assistance was given regarding the development of three-year strategic plans, as well as annual budgets and action plans. PanSALB transferred funds to the NLUs via the universities where the units are situated. The Editors-in-Chief were required to submit quarterly reports to PanSALB, including financial reports to ensure that NLUs' compliance with the relevant regulations could be monitored (PanSALB 2003/2004: 6; 2004/2005: 32; 2005/2006: 26; 2007/2008: 35). NLUs were furthermore assisted with the management of staff according to human resources prescripts (PanSALB 2003/2004: 6; 2004/2005: 32).

The Lexicography and Terminology Development focus area also established an information support service (PanSALB 2000: 30), guidance to NLUs with regard to procurement policies for information resources, and essential reference works that should be consulted during the lexicographic process (PanSALB 2001: 24; 2001/2002: 12). The NLUs obviously require permanent access to essential dictionaries, encyclopaedias, reference works and other information resources.

Copyright issues had to be solved — not only for the NLUs to obtain copyright clearance from the publishing companies to access material (i.e. to scan books and electronic material during the process of corpus building that leads to dictionary compilation) (PanSALB 2001/2002: 12), but also to safeguard the products of the NLUs (Alberts and Jooste 1998: 122-139). It was therefore suggested that the NLU Boards of Directors had to ensure that the service contract of every employee should include a clause to the effect that copyright in any work produced by the employee during his/her employment at the Company (NLU) would vest in the Company (NLU) (Alberts 2003).

The focus area staff and members of the Board visited the NLUs regularly to monitor progress and provide assistance where needed (PanSALB 2003/2004: 6-7).

NLUs were also assisted with the marketing of the units and the obtaining of community involvement (PanSALB 2004/2004: 7).

3.7 NLUs: Co-operative lexicography

The idea of co-operative lexicography already manifested itself in 1991 when it

was suggested that an institute for Southern African lexicography be established under the control of the then envisaged Pan-South African Language Board. In an attempt to concretise the idea of co-operative lexicography, the Board of Control of the WAT requested in 1992 that a feasibility study on the creation of an Institute for Southern African Lexicography be undertaken. This study was completed in July 1993 by an external research team led by the author.

The report, *Feasibility Study: Institute for Southern African Lexicography*, contained information that was indispensable for the further planning of lexicographical practice in South Africa. The report was based on questionnaires sent to 186 State-supported lexicographical projects and language offices, associations, private lexicographers, publishers, educational or cultural committees of political groupings, lecturers teaching lexicography, lexicographers at universities, various opinion formers, Ministers and MECs responsible for education and/or cultural affairs, and language boards (Alberts 1993).

Although the research indicated that an Institute for Southern African Lexicography was viable, it also showed that the important role-players did not feel the need for such a body. The success of the institute would rely on their collaboration. Respondents were afraid of a governing body with enforceable power. There was, however, a great need for collaboration, co-operation, co-ordination and communication. The respondents did not want an authoritative and bureaucratic body. They required a flexible structure with a lot of freedom. It was decided to establish an association for lexicography and thus the African Association for Lexicography (Afrilex) was founded in 1995. *Lexikos* of the *Afrilex Series* of the Bureau of the WAT served as its accredited journal and mouthpiece (Alberts 1993). Since its establishment, Afrilex supported co-operative lexicography and terminology through its international conferences, symposia and the publication of excellent articles in the *Lexikos* journal.

The lexicographic products rendered by lexicographers at the NLUs are used by a wide spectrum of target users, inter alia by learners, students, educators, academics, the economic sector, scientists, language practitioners, etc. It is therefore essential for lexicographers and other parties interested in the field of lexicography to co-operate as extensively as possible. Such co-operation may take place at various levels, including within and between national dictionary units, with members of the speech community, with language departments at tertiary institutions, with language bodies (i.e. National Language Bodies and Provincial Language Committees), with government and semi-governmental institutions, with associations for lexicography and terminology (i.e. Afrilex, Prolingua, ISO TC/37, SABS TC 37), with the private sector, with language practitioners such as translators, interpreters and terminologists, and with the media.

PanSALB had (and still has) a policy of actively promoting the notion of "co-operative lexicography" (PanSALB 2001: 24). The Board established systems of co-operation among all existing NLUs and other lexicographers nationally,

and obtained particulars of relevant stakeholders (PanSALB 2000: 30; 2001: 24-25). The experience in practical lexicographical work gained by the WAT and DSAE was shared freely with the other nine African language NLUs (PanSALB 2001/2002: 9).

Lexicographers at the NLUs had to seek input and feedback from target and end users of lexicographic products on a regular basis. By sharing information and expertise, the lexicographical process and products could be improved and promoted worldwide. The NLUs also initiated cross-border co-operation with countries sharing the same languages, especially in the case of marginalised languages (PanSALB 2004/2005: 38).

The NLUs were encouraged to submit their dictionary products to the relevant National Language Bodies (NLBs) for verification of the language matter (e.g. spelling and orthography conventions; word-forming principles) (PanSALB 2003/2004: 7). Closer collaboration between the NLUs and NLBs would greatly enhance the development of the various languages.

The NLUs were requested to become involved in the Human Language Technologies (HLT) project that was the brainchild of DACST and PanSALB (PanSALB 2001/2002: 11-12; 25-27). The HLT initiative could well lead to the establishment of an HLT industry in South Africa that will support translation and interpreting services, while the NLUs could provide crucial linguistic data to the project.

As part of the *Afrilex Series*, the Bureau of the WAT published various issues of *Lexikos*, a journal for lexicographers containing articles by authors from South Africa, other African countries and Europe — all who are experts in the field of lexicography.

3.8 NLUs: Training

PanSALB facilitated lexicography planning by assisting NLUs with information and training on issues such as data collection norms for dictionaries, norms for the inclusion and establishment of micro- and macrostructures for each dictionary project, and developing an editorial style guide for each dictionary to be compiled by a lexicography unit (PanSALB 2000: 30).

In 1999 PanSALB facilitated the training of the editors-in-chief, boards of directors and staff of existing NLUs in the principles and practice of lexicography; metalexigraphy; the use of computer hardware and software; the planning of projects and the management of NLUs; negotiation skills and marketing (PanSALB 2000: 30).

Follow-up training sessions were held for existing NLUs in October 2000 and in February 2001 for newly established NLUs. Topics covered at these workshops included the following (PanSALB 2001: 22):

- Directors' briefs (the King Code of Corporate Practices and Conduct)
- Drafting of business plans and budgets

- Reporting procedures
- Organisational structures and job descriptions
- Terms and conditions of employment
- Advertising and recruiting

Between 27 and 29 November 2000, staff members of existing NLUs were also given the opportunity to attend a training course entitled: *An introduction to theoretical lexicography and its practical applications*. The course was repeated at the end of March 2001 for suitable candidates identified by the boards of directors of the newly established NLUs. Prof. Rufus Gouws of the University of Stellenbosch was contracted by PanSALB to facilitate these training courses (PanSALB 2001: 23).

Prof. Gouws was also contracted by PanSALB in 2000 to facilitate, review and revise the then current and proposed projects at existing units, to assist newly established units in deciding what type of dictionary to compile first and to assist with the formulation of a dictionary conceptualisation plan and instruction book. This resulted in an action plan for the secondary lexicographic process within the NLUs (PanSALB 2001: 24). Prof. Gouws trained the staff at the units (PanSALB 2001/2002: 11) and training themes included the following:

- Dictionary typology
- The structure of dictionaries
 - Frame structure
 - Textbook structure
 - Macrostructure
 - Microstructure
 - Access structure
 - Mediosstructure (i.e. cross-referencing)
- Linguistic data and the presentation thereof in dictionaries

On 25 and 26 February 2002 a two-day training workshop on *The lexicographic process of the NLUs: Progress and planning* was held at the CSIR Conference Centre in Pretoria (PanSALB 2001/2002: 11). The training offered covered the following issues:

- The lexicographic process in the NLUs: recent issues
- Lexicographic management: the board of directors and the staff of the NLUs
- Contracts and legal matters
- Continued PanSALB support
- Implementing a dictionary plan

- Compiling and using corpora
- Using *Onoma* software
- Co-operation between the DSAE and the other NLUs

Prof. Gouws went on to develop a three-phased training plan on lexicographic planning on behalf of PanSALB (PanSALB 2001/2002: 11):

- Phase one: General theory of lexicography
- Phase two: Problems related to language groups
- Phase three: Language-specific features and problems

During the 2003/2004 financial year, four workshops were held for the staff of the NLUs aimed at improving the quality of lexicography work at the units. The following issues were addressed:

- Administrative and financial skills
- Improving the quality of lexicography work at the NLUs
- Training of Editors-in-Chief on managing NLU offices and staff (Alberts 2003)
- Improving computer skills of NLU staff (PanSALB 2003/2004: 6)

3.9 NLUs: Computerisation

In 2000, needs assessment studies with regard to computer hardware and software were undertaken by PanSALB at nine of the NLUs (PanSALB 2001: 24; 2001/2002: 11). The Bureau of the WAT started the computerisation of its lexicographic data by using the *Lexi* program developed especially for the WAT by the Human Sciences Research Council, while the DSAE was using X-Write software.

PanSALB established a relationship with a Swedish company on computational linguistics called Lexilogik AB. Their flagship program, *Onoma*, replaced the *Lexi* program at the Bureau of the WAT. PanSALB and Lexilogik AB concluded a contract at the end of March 2001 with regard to the computerisation of the other NLUs. After a detailed project plan had been drafted during a visit to South Africa by Lexilogik AB at the end of March 2001, the *Onoma* program was installed at the Sesotho sa Leboa and isiNdebele NLUs (PanSALB 2001: 24).

PanSALB contracted Prof. Daan Prinsloo of the University of Pretoria to assist the NLUs with the process of computerising collected data into a corpus (PanSALB 2001/2002: 11). This process entailed primary practical corpus building and dictionary compilation by using software. The training was done at the NLUs and continued during the 2003/2004 financial year. Prof. Prinsloo

assisted with the following issues (PanSALB 2003/3004: 6-7; 2004/2005: 32):

- Computer literacy of staff
- Improving computer skills of NLU staff
- Consultation of reference material
- Knowledge of information sources
- Data collection norms and policies
- Development and building of databases
- Development and building of corpora
- Expansion of corpora
- Use of the Internet and accessing of electronic data
- Strengthening the NLUs' information technology (IT) systems
- Putting systems in place to produce small dictionaries and other lexicographic products

The TshwaneLex dictionary production software that was developed in South Africa was demonstrated to PanSALB and in the 2004/2005 financial year the Board approved the obtaining of TshwaneLex licences for the eleven NLUs during the next financial year (PanSALB 2004/2005: 32; 2005/2006: 26). The staff of the NLUs received training in the use of TshwaneLex once the program had been installed at the units.

3.10 NLUs: Dictionary products

The ultimate goal of the NLUs is to produce dictionaries. The facilities, funding, guidance and training provided since the establishment of the units were efforts to accelerate the process of producing suitable and needed dictionaries for the local speech communities (PanSALB 2001/2002: 12).

Afrikaans NLU

- Volume XII of the *Woordeboek van die Afrikaanse Taal*, containing the letters P and Q was published in 2005 (PanSALB 2004/2005: 32).
- The electronic version of the *Woordeboek van die Afrikaanse Taal (A tot O) (Elektroniese WAT)* was made available in 2003 through the websites of six South African universities: Stellenbosch, Free State, Pretoria, Johannesburg, Northwest (Potchefstroom Campus) and South Africa (UNISA) (PanSALB 2004/2005: 33). The *Elektroniese WAT (A tot Q)* followed in 2006 (Wolvaardt 2008).

- An Afrikaans synonym dictionary was published in 2006 (PanSALB 2004/2005: 33; 2005/2006: 28; Wolvaardt 2008).
- An etymology dictionary of Afrikaans was published in 2005 (PanSALB 2005/2006: 28), with a Supplement that followed in 2007 (Wolvaardt 2008).
- New project: Volume XIII of the *Woordeboek van die Afrikaanse Taal* was published in 2009 (Wolvaardt 2008).

English NLU

- A Dictionary of South African English on Historical Principles (*DSAE-Hist*) was published in 1996 (Wolvaardt 2008).
- The DSAE edited the South African Concise Oxford Dictionary that was published by Oxford University Press (South Africa) in 2002 (PanSALB 2004/2005: 33).
- Research was done for a revision of the Illustrated School Dictionary for Southern Africa (PanSALB 2004/2005: 33).
- Additional material was identified for a revision of the Dictionary of South African English on Historical Principles (PanSALB 2004/2005: 33).
- The DSAE developed a comprehensive 'Reading and Intake Strategy' in order to ensure that research undertaken covers a broad range of South African English (PanSALB 2004/2005: 33).
- Other publications included the following:
 - South African Concise Oxford Dictionary (SACOD) (6th Impression, Nov. 2006) (Wolvaardt 2008)
 - Editorial consultants SA English: Encarta World Dictionary 2004 (Wolvaardt 2008)
 - Oxford Dictionary of English 2005 (Wolvaardt 2008)
 - Database of South African English, in progress (Wolvaardt 2008)
 - SACOD 2nd edition, in progress (Wolvaardt 2008)
 - DSAEHist, 2nd edition, in progress (Wolvaardt 2008)

IsiXhosa NLU

- (Former Xhosa Dictionary Project, 1968). Volume 2 (letters K to P) of the Trilingual Dictionary (Xhosa, English and Afrikaans) was published during the 2004/2005 financial year (PanSALB 2004/2005: 34).
- Volume 1 of the Trilingual Dictionary (letters A to J) was finalised (PanSALB 2004/2005: 34; 2005/2006: 28).
- The Greater Dictionary of Xhosa was published as follows: Vol. 2 in 2003, Vol. 1 in 2006 and Vol. 3 in 1989 (Wolvaardt 2008).

- A monolingual dictionary was published in 2008.
- A new project involves the compilation of a Maths, Science and Technology dictionary for Grades 4 to 6 (Wolvaardt 2008).

IsiZulu NLU

- The unit finalised the editing of a comprehensive monolingual Zulu dictionary (PanSALB 2004/2005: 35). This dictionary, *Isichazamazwi Sesizulu*, was launched in 2007 (PanSALB 2007–2008: 35) (earlier shorter versions, 1982 and 1985) (Wolvaardt 2008).
- The first draft of an isiZulu-English bilingual dictionary was published in 2010 (final product due in March 2012) (PanSALB 2009/2010: 43).

Siswati NLU

- The unit finalised the compilation of a draft English/Siswati dictionary. The draft was published and circulated to all stakeholders for comments and recommendations (PanSALB 2004/2005: 37), after which the *Silulu sesiSwati/Swati Bilingual Dictionary* (Siswati/English) was launched in 2007 (PanSALB 2007–2008: 35).
- The draft bilingual dictionary was launched in SA and Swaziland, and editing for its final publication is in progress (Wolvaardt 2008).
- The work on a Siswati monolingual dictionary was placed on hold in order to complete the bilingual draft (PanSALB 2004/2005: 37). A new project involving a monolingual dictionary is in progress (Wolvaardt 2008).

IsiNdebele NLU

- The NLU continued with the compilation of the English/IsiNdebele explanatory dictionary (PanSALB 2004/2005: 37; 2005/2006: 29) that was published in 2010 (PanSALB 2009/2010: 43). A monolingual dictionary was published in 2009 (Wolvaardt 2008).
- A draft bilingual scholars' dictionary was launched in March 2006 (Wolvaardt 2008).
- A new project that is under way involves a trilingual English, Afrikaans and isiNdebele dictionary (Wolvaardt 2008).

Sesotho sa Leboa NLU

- The NLU transferred the lemmas for the comprehensive monolingual dictionary from the records held in MS Word to TshwaneLex. A draft

dictionary was presented to stakeholders at the NLU's annual general meeting in 2004 (PanSALB 2004/2005: 36). The dictionary *Pukuntsutlhalosi ya Sesotho sa Leboa — Volume I* (Monolingual Dictionary) was launched in 2008 (PanSALB 2007–2008: 35; Wolvaardt 2008).

- Work was done on the compilation of a bilingual Sesotho sa Leboa/English dictionary (PanSALB 2004/2005: 36). A draft copy was published in 2005 (PanSALB 2005/2006: 30) and the dictionary was launched in 2008 (PanSALB 2007–2008: 35).

Sesotho NLU

- The unit completed the letters A to E of the draft bilingual Sesotho/English dictionary during the financial year 2004/2005 (PanSALB 2004/2005: 35). The English–Sesotho dictionary was published in the 2008 financial year (Wolvaardt 2008).
- The unit aimed to commence work on a Sesotho monolingual dictionary (PanSALB 2004/2005: 35). A corpus of 4,160,362 running tokens was built, and six additional Sesotho books were scanned and OCR errors cleaned up for inclusion in the corpus (Wolvaardt 2008).
- A mini scholars' bilingual dictionary produced by the unit was published in 2005 (PanSALB 2005/2006: 31).
- A primary school dictionary with 3 371 lemmata/headwords with Sesotho equivalents was compiled (Wolvaardt 2008).

Setswana NLU

- The unit prepared a mini dictionary during the 2004/2005 financial year and the lexicographical data was sent to various language experts for editing and quality control (PanSALB 2004/2005: 36).
- The unit aimed to start work on a bilingual Setswana/English dictionary (PanSALB 2004/2005: 36).
- A monolingual dictionary was compiled and sent through a consultative process for verification (PanSALB 2005/2006: 30). The monolingual manuscript was completed in June/July 2006, and a process of review and editing is currently in progress (Wolvaardt 2008).

Xitsonga NLU

- The NLU continued work on a bilingual explanatory Xitsonga/English dictionary. The English/Xitsonga section had been completed and work was done on the reverse side, namely the Xitsonga/English section (PanSALB 2004/2005: 36). The dictionary was published in 2005 (PanSALB 2005/2006: 31).

- Work on a comprehensive monolingual Xitsonga dictionary project was started in 2004 (PanSALB 2004/2005: 36; 2005/2006: 31).
- Both bilingual and monolingual draft dictionaries were presented for consideration at the unit's annual general meeting in 2005 (PanSALB 2004/2005: 36). The monolingual dictionary was published in 2009 (Wolvaardt 2008).
- Revised Terminology No. 3 was published in March 2009 (Wolvaardt 2008).
- A Xitsonga traditional dictionary was launched in the 2009/10 financial year (Wolvaardt 2008).

Tshivenda NLU

- The NLU published a draft of the first volume of a bilingual Tshivenda/English dictionary during the 2004/2005 financial year, after which copies of the draft were disseminated to stakeholders for comments and suggestions (PanSALB 2004/2005: 38). The project was completed in 2005 (PanSALB 2005/2006: 32) and the bilingual dictionary published in 2006 (Wolvaardt 2008).
- The unit subsequently started work on a monolingual Tshivenda dictionary (PanSALB 2004/2005: 38) that is still in progress (Wolvaardt 2008).

3.11 NLUs: Management Forums

After consultations in 1996 with the former editor-in-chief of the Bureau of the WAT, Dr Dirk van Schalkwyk, the author (then of the Department of Arts, Culture, Science and Technology) suggested that a management committee be created to oversee the lexicographical processes. The suggestion — involving the establishment of a co-coordinating lexicography board consisting of the eleven editors-in-chief of the eleven NLUs, with the Director-General of DACST (or his delegate) as chairperson and a member of DACST (i.e. a staff member of the National Terminology Service) as secretary of this board — was presented to the Portfolio Committee (Alberts 1996). This board was, however, not established as the National Lexicography Units Bill was incorporated into the PanSALB Amendment Act of 1999.

There is, however, a need for a coordinating body to encourage coordination and to share problems and solutions. The present incumbents of editors-in-chief created such a committee, namely the LexiEditors' Forum, and they all served *ex officio* in the Forum. The PanSALB Board approved the establishment of the LexiEditors' Forum and encouraged the NLUs to cooperate with each other in matters related to lexicography development (PanSALB 2005/2006: 26). The CEO of PanSALB also attends the meetings of this Forum.

The LexiEditors' Forum is a discussion and consultation group that en-

ables editors-in-chief to share their experiences and information about lexicography and the operation and progress of their particular NLU, as well as seek solutions to the challenges they all face in their work. The aim of this Forum is therefore to create better working relationships among the NLUs and to discuss issues of mutual interest (PanSALB 2004/2005: 38; 2005/2006: 26).

The NLUs responsible for the four most marginalised languages, namely Tshivenda, Siswati, isiNdebele and Xitsonga also formed a committee in order to collaborate on matters affecting their units and languages (PanSALB 2004/2005: 38).

3.12 NLUs and Terminology development

Several factors influence the development of terminology in the (South) African languages. Most of these factors are of a political and sociolinguistic nature. In the past, terminology development in South Africa was hampered by a number of ideological, historical, and educational factors, *the most fundamental of which were the language policies adopted in the Republic of South Africa (RSA)* (Mtintsilana and Morris 1988: 109). With English and Afrikaans being the official languages in the RSA up to 1994, the State actively supported terminology development in Afrikaans to try to ensure parity with English (Alberts, Botha and Kapp 2010).

Since 1995 structural and policy factors have had an effect on the practice of terminology in South Africa. Some of the language bureaux were either dissolved or reduced to perform a minimum of functions. The State furthermore accepted its responsibility to also develop and promote the African languages (Alberts, Botha and Kapp 2010).

Government policy in language promotion before 1994 stipulated that the promotion of the African languages was the task of the former national states and the Self-Governing Regions, and not that of the South African Government (Alberts, Botha and Kapp 2010).

In 1998, PanSALB's Lexicography, Terminology and Place Names sub-committee and focus area realised the need for coordinating all isolated lexicography and terminology efforts in the country so as to develop a strategy for approaching lexicography and terminology development for all the official languages in the country (PanSALB 1998: 11). PanSALB subsequently consulted various stakeholders involved in these areas.

The purpose of the Lexicography and Terminology Development focus area was inter alia to provide advice on lexicography and terminology development (PanSALB 2000: 26). PanSALB (2001: 24) later decided not to deal with terminology per se under this focus area, since this was the function of the Terminology Section of the National Language Service in DACST (currently known as the Terminology Coordination Section of the National Language Service (NLS), Department of Arts and Culture).

The terminologist working in a multilingual society such as South Africa thus deals with conflicting situations. The African terminologies should be developed to enhance the multilingual heritage of the country, but there are too many languages to make this practical and economically viable. There are also very few trained terminologists with even less terminology posts available. Various tertiary training institutions offer official courses in terminology and lexicography, and aspects relating to terminology principles are dealt with in courses such as translation or lexicography. The Terminology Coordination Section (TCS) provides in-house training and will have to offer bursaries on terminology training to students interested in this profession if the country's communication problems are to be solved through multilingual terminology in a variety of subjects or domains (Alberts, Botha and Kapp 2010).

These factors stress the importance of the Terminology Coordination Section as an organisation working on a national basis to document, standardise and modernise the various technical vocabularies used in South Africa (i.e. by means of compiling multilingual polythematic technical dictionaries and providing terminology training to collaborators in the field of terminology). The Language Research and Development Centres (LRDCs) assisted for many years with the documenting of terminology but they were closed down. Language units at governmental departments could assist with the documentation of the terminology of the various departments but they will have to work in collaboration with TCS and PanSALB. The translators working at the National Language Service harvest terminology by means of the translation-orientated terminography process where SL and TL documents are aligned and terminology extracted from both documents. The terms harvested in this way go directly to the national term bank (Alberts, Botha and Kapp 2010).

Some of the NLUs assisted with the compilation of terminology lists, e.g. a draft list of bird names in Siswati (PanSALB 2004/2005: 37) and the collaborative project of the IsiXhosa NLU and the Medical Research Council regarding the translation of a set of English medical terms into Xhosa (2005/2006: 28).

In 2004 PanSALB established a new focus area: Standardisation and Terminology Development to advise PanSALB on and implement projects related to the standardisation and terminology development of the official languages, the Khoe and San languages and South African Sign Language. The key strategic objectives of this directorate involved issues relating to primary standardisation processes, primary terminology development processes and secondary standardisation and terminology development processes. This directorate worked in close collaboration with the Terminology Coordination Section of the National Language Service in the Department of Arts and Culture and other organisations involved in terminology projects in order for the National Language Bodies to verify and authenticate terms (PanSALB 2004/2005: 25). This focus area was shut down with the retirement of the Director heading the section in 2010.

4. Future

4.1 DAC

The Department of Arts and Culture through the National Language Service is not in a position to accommodate the NLUs. The Department should do language planning and oversee legislative processes (e.g. a national language policy framework; legislation to protect language practitioners and enhance their working conditions), develop policies and matters related thereto. It should furthermore provide facilities and funding for research into language-related matters (i.e. the HLT endeavours) to improve the language situation in South Africa and to ensure that the constitutional demands for multilingualism are met. Practical language-related work such as dictionary compilation (general as well as for special purposes (terminology)) and translation work should not be done at the NLS. DAC should therefore not take care of the NLUs, translation or terminology services.

4.2 PanSALB

The 2003/2004 objective of the Lexicography and Terminology Development focus area was "to create conditions conducive to the development of all the official languages in South Africa" (PanSALB 2003/2004: 6). The focus area (consisting of a Deputy Director and an Assistant Director) and its subcommittee focused on the following strategic objectives (PanSALB 2003/2004: 6):

- Ensuring the appointment of Editors-in-Chief for all NLUs
- Capacity building and training of staff at NLUs
- Putting Information Technology (IT) systems in place for the NLUs
- Providing administrative support to NLUs in terms of provisioning systems, human resource management and financial management
- Supporting NLUs in terms of dictionary production processes
- Facilitating cooperation between NLUs and other bodies

It is clear from information supplied in §3 that various tasks were successfully undertaken regarding the administrative, financial and technical management of the NLUs. It seems, however, that the current focus area manager at PanSALB experienced problems regarding the administrative management of the NLUs, i.e. "owing to staff shortages, both at head office and at the NLU offices, and the inequitable salaries paid to staff members" (PanSALB 2007-2008: 36) — it is not clear whether the salaries mentioned were that of head office staff or NLU staff. The lack of information on the activities of the focus area

and of the NLUs after 2007 (compared to previous years) was also an indication of troubled waters (PanSALB 2007/2008: 35-36; 2009/2010: 43-44).

According to the annual report of 2007/2008 the strategic goal of the **focus area** was "to compile comprehensive monolingual and/or bilingual dictionaries" (PanSALB 2007–2008: 35). Surely PanSALB head office staff was NOT appointed to *compile dictionaries* but rather to *provide support services to the NLUs* whose main task is dictionary compilation.

The only measurable objectives of the focus area since 2009 have been "to assist in accelerating the production of dictionaries and other products that could generate income for NLUs in the near future. (To) ensure appointment of new Board of Director (BoD) members for all NLUs" (PanSALB 2009/2010: 43).

The only major event mentioned in the 2009–2010 Annual Report of PanSALB involves the actions taken to appoint new members on the Board of Directors. These include the nomination of new members for Board of Directors for all eleven NLUs; the publication of the preliminary name list for public comments; gazetting of the new BOD members, and the inaugural meetings of the new BoD members (PanSALB 2009/2010: 43-46). No mention was made of any other activities or achievements.

It has been difficult to obtain information on the dictionaries compiled by the various NLUs. It therefore seems as if PanSALB is no more in a position to create a working environment conducive to dictionary production.

4.3 Management forum

At present the editors-in-chief are *ex officio* members of the LexiEditors' Forum. It might however be advantageous to create another management structure with an enhanced membership structure, involving the following persons:

- Chairpersons of the Board of Directors
- Editors-in-Chief (as managers of the NLUs)
- Final Editors (as chief lexicographers)
- Directors responsible for Finances on the BoDs

In 1991 the Bureau of the *Woordeboek van die Afrikaanse Taal* requested a feasibility study to determine the need for an institute for lexicography (see § 3.7; Alberts 1993). At the time it transpired that the respondents did not want another bureaucratic institution. The idea of such a South African Institute for Lexicography (and Terminography) might, however, be justifiable in the current situation, seeing that there is no existing institution that is eager to take responsibility for the NLUs.

A South African Institute for Lexicography (and Terminography) could be a virtual office managing all the NLUs as well as terminology services (see § 3.12; § 4.4). It has been difficult to determine what type of dictionaries were

published, are ongoing or in a planning phase — a national register on lexicographical and terminographical projects would assist target users.

4.4 Dictionary services — general and special purposes

In the past, PanSALB considered lexicography and terminology "to be important building blocks in the development of languages and the promotion of multilingualism" (PanSALB 2000: 26). The Board "established lexicography units as section 21 companies for each official language in South Africa. These units [would] be responsible for the development of dictionaries and terminology for their respective languages" (PanSALB 2000: 26).

There is, however, also a dire need for terminology to be documented in rural, urban and academic fields, where fieldwork is needed to document community-based terminology (indigenous knowledge systems, academic subjects).

Terminologists could be appointed at the NLUs to do fieldwork and to harvest terminology from the speech communities in the rural, urban and academic environment so as to ensure that the existing terminology in these areas is documented and preserved. Each lexicography unit is presently seated at tertiary institutions within the relevant speech community.

Terminologists who are presently working at the Terminology Coordination Section, National Language Service, Department of Arts and Culture (see § 4.1) could become employees at the NLUs together with the budget allocation from the TCS attached to the specific terminologist to finance the position. These terminologists employed at the NLUs could do fieldwork and harvest terminology in urban, rural and academic areas. Trained terminologists could assist with language-specific terminology development. In cases where terminologists cannot relocate due to logistic problems, they could work online (through the TshwaneTerm/TshwaneLex software) and the envisaged HLT virtual network. The combination of lexicography and terminography work within the same dictionary unit is a model followed by several dictionary offices in Africa (Alberts 2010: 616-617).

4.5 Department of Higher Education and Training

According to Bamgbose (1991: 109-111) there is an escalating consciousness in Africa and worldwide of the positive impact of multilingualism, especially in respect of the role of African languages in advancing multilingualism in education. This is reflected in the language policies that acknowledge the need for teaching and acquiring these languages at all levels of education. Terminologists should coin new terms as necessary according to the new function required of the specific language (terms which might not be in existence, especially in technical areas), while lexicographers would be required to write dictionaries.

In October 2010 the Department of Higher Education and Training held a Roundtable discussion on the position and developmental status of the African languages. According to Language Policy in Higher Education, the use of South African languages in instruction and learning in higher education would require the development of dictionaries and other teaching and learning materials. The meeting acknowledged the existence of various general dictionaries compiled by the NLUs of PanSALB, the various multilingual terminology lists compiled by the terminologists of the TCS of NLS, DAC, and other teaching and learning materials that have since been produced. The question raised by the delegates, however, was whether these dictionaries and materials were enough to meet the requirements of the target group and whether more should be done to meet the requirements of constitutional multilingualism and target users. It was agreed that the State should start with systematic and widespread translation into African languages, supported by lexicography work and terminology development (Alberts, Botha and Kapp 2010; Nosilela 2010).

Delegates also agreed that more funding should be made available to dictionary compilation efforts, albeit general or technical dictionaries. A suggestion was made that the NLUs should be relocated to the Department of Higher Education and Training since PanSALB interpreted its mandate as the monitoring of languages (through the National Language Bodies and Provincial Language Committees) and NOT as the development of languages (Alberts, Botha and Kapp 2010).

5. Conclusion

This article started with an overview of the national dictionary offices in South Africa during the bilingual dispensation. It stressed the need for national dictionary offices in all the official languages of the country to adhere to the constitutional multilingual prescripts of post-apartheid South Africa. Next followed a discussion of the drafting of the National Lexicography Units Bill which had led to the establishment of eleven official national lexicography units to document, standardise and preserve the linguistic heritage. These issues were discussed in terms of the past, present and future. Taking into account the value of national dictionary projects, it is obvious that there would be serious implications should these dictionary offices cease to exist.

It is of vital importance that attention should be given to the further training of lexicographers, terminologists and students to enhance knowledge and skills in the principles and practice of general and technical dictionary compilation. Government should give financial assistance for lexicography and terminology awareness campaigns and for the training of practitioners in these fields. Collaboration with national and international experts regarding the standardising of principles and procedures of lexicography, terminography and computational linguistics will enhance the work of practitioners working

in these fields. Participation in the envisaged Human Language Technology (HLT) virtual network will expedite the HLT endeavours.

By registering various completed, ongoing and planned lexicographical and terminological projects, a clear picture of lexicography and terminology practice will emerge and the duplication of work will be prevented.

Serious consideration should be given to the seats of the NLUs. A logical overseeing institution is the Department of Higher Education and Training since the NLUs are already situated at tertiary institutions. The NLUs need — on a continuous basis — proper funding, administrative, technical and IT support and training to ensure excellent lexicography work.

The NLUs need to be given the resources that are necessary to revert to the original idea of producing properly researched, definitive, monolingual dictionaries for each of the official languages instead of compiling all kinds of dictionary products to generate funds (dictionaries that commercial publishers are able to compile much better, and that they are already publishing successfully). The NLUs are essentially the only structures suitable for undertaking major monolingual dictionary products, but they need to be subsidised substantially. The NLUs should be cultural and linguistic databases from which other material for special purposes could be developed. The NLUs should be seen as a cultural investment; a view that is essential if multilingualism is ever going to be feasible.

Consideration should also be given to moving all lexicographical and terminographical activities — even translation activities — to the NLUs in order to provide a proper language service to the citizens of South Africa.

There might at this stage be room for an independent (virtual) South African Institute for Lexicography and Terminography to manage the decentralised NLUs situated in their respective geolinguistic areas at tertiary institutions. It is also the author's view that the time is ripe for the establishment of such an institute, as this might promote citizen awareness of the value of lexicography and terminology for communication.

6. References

- Alberts, M. 1993. *Feasibility Study: Institute for Southern African Lexicography*. Stellenbosch: Bureau of the WAT.
- Alberts, M. 1996. *Background Information on the Drafting of the National Lexicography Units Bill*. Unpublished document. Pretoria: DACST.
- Alberts, M. 1997. *Reallocation of Funds to be used for Lexicography Training and Lexicography Awareness Campaign*. Pretoria: DACST.
- Alberts, M. and M. Jooste. 1998. Lexicography, Terminography and Copyright. *Lexikos* 8: 122-139.
- Alberts, M. 2003. *Appointment of Editors-in-Chief*. Unpublished paper read at the PanSALB workshop held at the Sinodale Building, Visagie Street on 3 March 2003 in Pretoria.
- Alberts, M., W.F. Botha and P.H. Kapp. 2010. *Historical Experiences in Developing Afrikaans as a Language*. Paper presented at the Roundtable on African Languages, organised by the Depart-

ment of Higher Education and Training and held at Unisa, Sunnyside Campus, Pretoria on 22 October 2010.

- Bamgbose, A.** 1991. *Language and the Nation: The Language Question in Sub-Saharan Africa*. Edinburgh: Edinburgh University Press.
- DAC.** 2002. *National Language Policy Framework, Final Draft*. Pretoria: Department of Arts and Culture (DAC).
- DACST.** 1996. *National Lexicography Units Bill*. (As introduced) [B 103–96]. *Wetsontwerp op Nasionale Leksikografie-eenhede*. (Soos ingedien) [W 103–96]. Pretoria: Minister of Arts, Culture, Science and Technology, Department of Arts, Culture, Science and Technology (DACST).
- DACST.** 1997. *PanSALB and the Tabling of the National Lexicography Units Bill*. Unpublished document (5/3/2/7/2/3). Pretoria: Department of Arts, Culture, Science and Technology (DACST).
- Mtintsilana, P.N. and R. Morris.** 1988. Terminography in African languages in South Africa. *South African Journal of African Languages* 8(4): 109-113.
- Nosilela, B.** 2010. *Current State of African Languages in Institutions of Higher Learning*. Paper presented at the Roundtable on African Languages, organised by the Department of Higher Education and Training and held at Unisa, Sunnyside Campus, Pretoria on 22 October 2010.
- PanSALB.** 1998. *Annual Report*.
- PanSALB.** 1999. *Annual Report*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2000. *Annual Report*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2001. *Annual Report*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2002. *Annual Report 2001/2002*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2004. *Annual Report 2003/2004*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2005. *Annual Report 2004/2005*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2006. *Annual Report 2005/2006*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2008. *Annual Report 2007/2008*. Pretoria: Pan South African Language Board (PanSALB).
- PanSALB.** 2010. *Annual Report 2009/2010*. Pretoria: Pan South African Language Board (PanSALB).
- Wolvaardt, J.** 2008. *Summary of NLU Products and Ongoing Projects* (Updated July 2008). Unpublished document. LexiEditors' Forum.

Needs-adapted Data Presentation in e-Information Tools

Henning Bergenholtz, *Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark (hb@asb.dk) and extraordinary professor, Department of Information Science, University of Pretoria, Pretoria, South Africa*

and

Theo J.D. Bothma, *Department of Information Science, University of Pretoria, Pretoria, South Africa (tbothma@up.ac.za) and visiting professor, Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark*

Abstract: In the current debate about the status of lexicography there are at least three quite different opinions:

- (i) Lexicography does not have or need any kind of own theory but can use all relevant linguistic theories.
- (ii) Lexicography needs a special theory for the lexicographical praxis, but this discipline is still a part of linguistics.
- (iii) Lexicography is a genuine part of information science and can use theories and learn from practice in the information society, but it also needs special theories for lexicography.

It is the third opinion we will maintain in this paper by discussing the information needs in the information society and partly using the function theory of lexicography.

Keywords: LEXICOGRAPHY, FUNCTION THEORY, INFORMATION TOOL, DICTIONARY, FIXED EXPRESSION, INFORMATION OVERLOAD, INFORMATION STRESS, USER NEEDS, COMPLEX USER SITUATION, CONTEXT, COGNITIVE SITUATION, COMMUNICATIVE SITUATION, OPERATIVE SITUATION, INTERPRETATIVE SITUATION, MONOFUNCTIONAL INFORMATION TOOL, POLYFUNCTIONAL INFORMATION TOOL, INTEGRATED INFORMATION TOOL, DATABASE, DATA PRESENTATION

Opsomming: Data-aanbieding aangepas volgens behoefte in e-inligting-hulpmiddels. In die huidige debat oor die status van leksikografie is daar minstens drie wyd uiteenlopende standpunte:

- (i) Leksikografie het nie en benodig ook nie enige tipe eie teorie nie, maar kan van alle relevante taalkundige teorieë gebruik maak.
- (ii) Leksikografie benodig 'n spesiale teorie vir die leksikografiese praktyk, maar hierdie dissipline is nog steeds deel van taalkunde.
- (iii) Leksikografie is 'n ware deel van inligtingkunde en kan die teorieë gebruik en uit die gebruike in die inligtingsamelewing leer, maar dit benodig spesiale teorieë vir leksikografie.

Ons handhaaf die laaste standpunt in hierdie artikel waarin inligtingbehoefes in die inligting-gemeenskap bespreek word met gedeeltelike gebruik van die funksieteorie van leksikografie.

Sleutelwoorde: LEKSIKOGRAFIE, FUNKSIETEOORIE, INLIGTINGHULPMIDDEL, WOORDEBOEK, VASTE UITDRUKKING, INLIGTINGOORLADING, INLIGTINGSPANNING, GEBRUIKERBEHOEFES, KOMPLEKSE GEBRUIKERSITUASIE, KONTEKS, KOGNITIEWE SITUASIE, KOMMUNIKATIEWE SITUASIE, OPERATIEWE SITUASIE, INTERPRETATIEWE SITUASIE, MONOFUNKSIONELE INLIGTINGHULPMIDDEL, POLIFUNKSIONELE INLIGTINGHULPMIDDEL, GEÏNTEGREERDE INLIGTINGHULPMIDDEL, DATABASIS, DATA-AANBIEDING

1. Introduction

In the function theory of lexicography (Bergenholtz 2010; Bergenholtz and Bergenholtz 2007; Bergenholtz and Bergenholtz 2011; Bergenholtz and Gouws 2007; Bergenholtz and Tarp 2002; Bergenholtz and Tarp 2003; Bergenholtz and Tarp 2005; Gouws and Leroyer 2009; Leroyer 2009; Nielsen 2009; Tarp 2007; Tarp 2008a; Tarp 2008b; Tarp 2008c; Tarp 2009a; Tarp 2009b; Tarp 2011; Tono 2010; Verlinde, Leroyer and Binon 2009) the term 'information tool' is used rather broadly to include any information object from encyclopaedias to telephone directories, travel guides and instruction manuals in addition to dictionaries, lexica, glossaries, etc. This extension is based on two issues, namely that all these information objects are consulted in a non-linear fashion when there is a need for specific information and all information objects are organised in a similar way, that is, they are highly structured based on a predefined set of criteria (alphabetically, grouped according to topics, etc.). The terms 'information object' and 'information tool' will be used in this article to emphasise the fact that the principles described here do not apply only to traditional dictionaries, but to this broader range of information objects that include dictionaries as well. Furthermore, the 'e'-designation is added because even though the theoretical principles apply equally to print and digital information objects, the practical implementation is only feasible in the digital or e-environment.

2. Users and information stress

The information society is characterised by an information explosion. Many researchers try to estimate the rates at which information doubles (see, for example, Lyman and Varian 2003). Numbers in such research differ widely and it is probably impossible to be accurate in the estimates. What is clear, however, is that the amount of information available increases consistently and there is no possibility that the rate will decrease. This simply means that there is more information available than any user can conceivably need in any given situation and that, when a user searches for information, he/she is

overwhelmed by the results. This information overload (in terms of what is presented to the user) causes information stress in the user. The terms 'information overload' and 'information stress' are therefore related causally — 'a' causes 'b' and both are flip sides of the same coin. When a system presents too much information to a user (information overload), the user experiences this as information stress.

The following searches, done in October 2010, illustrate the point. Searching for example for the phrase 'information management' on Google results in close to 14 million results. If this search is repeated in a research environment on Google Scholar, nearly 700 000 research items are retrieved. Repeating the same search on EbscoHost (*EBSCO Publishing* 2010) provides more than 104 000 results with more than 51 000 peer-reviewed articles, of which more than 8 500 have appeared since January 2009. ScienceDirect (*ScienceDirect* 2010) (SciVerse) provides references to more than 22 000 articles and 1 750 books, which include more than 4 000 articles and 300 books that have been published since January 2009.

This implies that both the general user and the researcher are totally overwhelmed by the amount of available information. (It obviously also means that a search for information should be narrowed down by using more search terms combined with Boolean operators. However, this requires a fairly high level of information literacy that the average user does not have.) The information overload provided by search engines and journal platforms/databases results in information stress for the user. The user becomes totally overwhelmed by the amount of information and does not know how to cope with what is available. This may also result in a situation referred to as 'information death' — the user is so overwhelmed by what is presented that he/she simply turns away from the problem and does not try to satisfy the information need.

A recent survey by LexisNexis (October 2010) reports that "information overload is a remarkably widespread and growing problem among professionals around the world, and one that exacts a heavy toll in terms of productivity, performance and employee morale". It "reveals the pervasive nature of the problem in white collar workforces around the world, and the very real impact that information overload is having on worker productivity, their state of mind and quality of their work product", and adds that "[a]pproximately one in two white collar professionals surveyed report feeling demoralized when they can't manage all the information that comes their way at work" (*LexisNexis 2010 International*, 2010).

Users would therefore prefer to get exactly the information they need to satisfy their information need in a specific situation. In the context of dictionaries, Haas already formulated this in 1962: "The perfect dictionary is one in which you can find the thing you are looking for preferably in the very first place you look" (Haas 1962: 48). In this quotation, 'dictionary' can equally well be replaced by 'information source'. Any user consulting any information

source would prefer that it provides him/her with exactly the information he/she needs in the given situation — not more and not less. Bergenholtz and Gouws (2010: 3) formulated this as follows: "For the user the type of information source is not important. Important is that he/she retrieves the exact required information as quickly as possible."

The information that the user retrieves should therefore be relevant to his/her information need in the context of the information need, whether this is in "[s]ocial, systemic domain and work task contexts" (Ingwersen 2007), both individual and collective (see Ingwersen 2001; Ingwersen 2007; Ingwersen and Järvelin 2004; Ingwersen and Järvelin 2005). In addition, the information should be correct and complete (i.e. detailed enough) to solve the user's information need, and it should also be understandable to the user (see Bergenholtz and Bergenholtz 2007; Bergenholtz and Gouws 2007).

Search engines such as Google try to increase the relevance of search results (*inter alia*) based on the country of origin of the user and the country-specific version of Google. For example, searching for 'information technology' on Google renders around 65 million hits. However, the first pages of results of country-specific versions of Google list different results. This is evident when comparing the results for this search to a sample of other Google versions, such as the American (.com), South Africa (.co.za), Dutch (.nl), Danish (.dk), Spanish (.es), French (.fr) and German (.de) versions. In each case there are a few generic first listings (e.g. to Wikipedia), but thereafter the results differ quite widely, especially in terms of the origin and language of the retrieved sites. Google therefore assumes that a user from South Africa would find results from websites in South Africa on average more relevant than websites from, say, Denmark or Spain (except for a number of very commonly used sites).

Websites sometimes list retrieved items according to how relevant the retrieved items are to the search query. For example, in a digital library such as the *ACM Digital Library* (2010) and in commercial systems such as Amazon.com (2010) retrieved items can be displayed using different criteria such as relevance, even though the rationale for this relevance ranking is not apparent to the user. In all preceding examples the principle remains the same: from among the full set of records in the database the system tries to display first those items that it perceives as most relevant to the user and then all items are displayed in a descending order of relevance. The reason for this is obvious: the user should not be overloaded with less relevant or irrelevant information, precisely to avoid the stress that information overload causes.

In all the above cases relevance is simply system relevance, i.e. a match between the search query input by the user and the data items in the database. Further dimensions of relevance, such as the possible relevance of the retrieved item to the user's situation (situational relevance), state of knowledge (cognitive relevance), etc. (Cosijn and Ingwersen 2000; Cosijn 2006; Järvelin and Ingwersen 2010) are not taken into account.

System relevance does not take into account the unique characteristics of the user in any given situation simply because the system has no knowledge about the user. The system does not know what the user's level of knowledge of the topic is — expert, semi-expert or lay person. The system does not know what the user's general and technical language proficiency in the language of the document is — mother tongue speaker or specific level of proficiency in L2/L3. The system does not know what level of detail the user needs — an in-depth discussion or a brief overview. The system does not know for what type of task the user needs the information — a work task, leisure, study, etc. In each of the preceding cases the variables can be plotted on a continuum and be combined in multiple ways. For example, based on only three variables — subject knowledge, general language proficiency and technical language proficiency — Bergenholtz and Bergenholtz (2007) and Bergenholtz and Gouws (2007) identified 30 different types of user. They did not consider the other variables of a user's profile and context that were mentioned above.

The characteristics and context of the user are therefore complex factors that are not taken into account when the system executes a query and presents the results of a search. The information that is offered to a specific user is not optimally relevant to him/her in a specific context and may result in information overload by the system and information stress on the user.

In the preceding section examples were provided from e-resources in general. A similar situation obviously prevails in respect of current e-dictionaries and other e-information tools.

3. Different information needs over time

As suggested above, the characteristics of information differ. The characteristics of the user may also differ in various contexts and user needs differ over time in a specific context. Each of these aspects will be discussed briefly.

- Information could be characterised based on detail, e.g. from very little detail to comprehensive detail. (This correlates to a certain extent with the length of the information element — from a single sentence or paragraph to a comprehensive essay/article to a book.) The complexity of the information can differ: an article on the human genome project in a popular news magazine differs widely in complexity from a technical research article on the same topic, but may be less complex than an article prescribed for students at undergraduate level. A specific unit of information could therefore be aimed at a lay person, a semi-expert or an expert. Many permutations of this phenomenon are possible. A short definition can be a very easy definition aimed at the lay person while a highly complex definition could be considered easy by the expert. Similarly, a long essay can be written 'in simple English' for

the lay person or could contain highly complex scientific terminology and argumentation aimed at the expert.

- Three typical user categories, based on their domain knowledge, were mentioned in the preceding paragraph — a lay person, a semi-expert (or an 'interested'/knowledgeable lay person) and an expert. However, no expert is an expert in all fields. An expert in nuclear physics is most probably not an expert in biomedical research as well, but may be a semi-expert in another field either related to his/her research or in a field that he/she practises as a hobby. One person, therefore, has different characteristics depending on the situation and the context, and this has an impact on the characteristics of the information he/she may need to solve an information-related problem.
- A user's information needs may differ within a specific context. In completing for example a work task (a single context), an individual user (for example an expert) may have a number of different information needs that may differ vastly in terms of the characteristics of the information that he/she requires.

This could be illustrated by means of the following example. In the process of doing research the researcher will typically need only detailed information aimed at the expert. However, in this process, he/she may want to read somewhat more widely than the narrowly defined research problem within a field where he/she is not an expert, especially if the research undertaken is of an inter- and multidisciplinary nature. At this stage the expert may need fairly comprehensive information aimed only at the semi-expert. He/she may, if the situation requires it, move to more complex and detailed information aimed only at the expert. In these readings the expert may come across technical terms for which he/she needs only brief explanations or definitions, again aimed at the semi-expert (or even a lay person). The expert may also read articles (either aimed at the expert or the semi-expert) in a language other than his mother tongue. In this case he/she may need help with text reception, a simple communicative information need, to help him/her to understand the text in L2. The complexity of this reception problem may also differ. In an uninflected language it may be very simple to find a translation equivalent from L2 to L1; however, in a highly inflected language it may be more problematic to identify the root (infinitive/lemma) of a verb to enable the user to find the required translation equivalent.

The above example is represented diagrammatically in Figure 1 below, and is further explained in the bulleted points following the diagram. Attributes of an element are given below the element, indented and placed in square brackets.

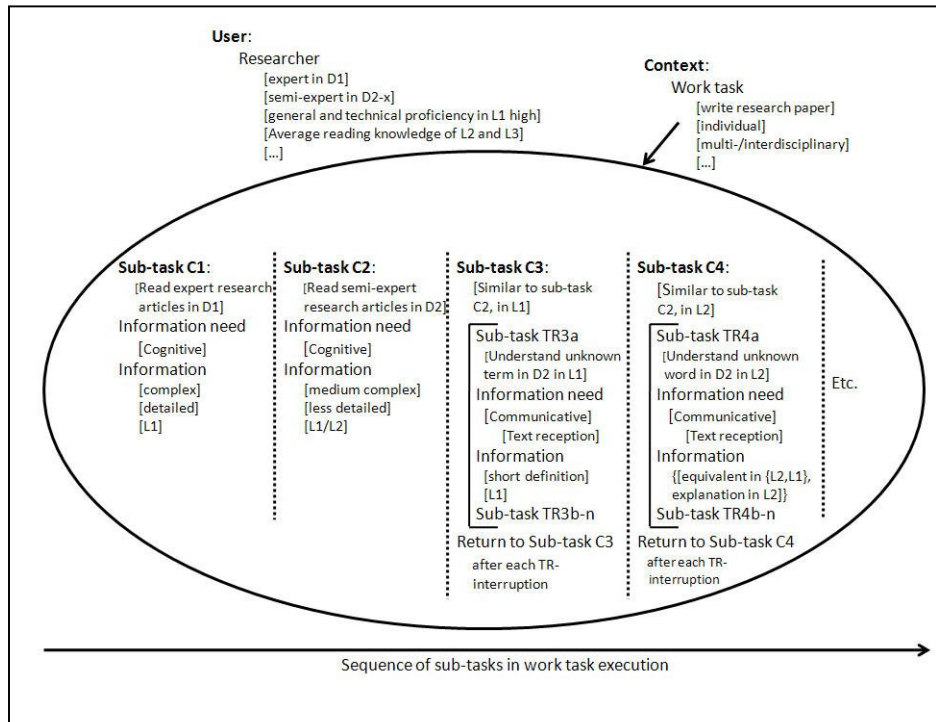


Figure 1: Attributes of user, context, information needs and information during work task execution

Explanatory notes to Figure 1:

- Both the user (a researcher) and the context (a work task) have certain attributes that are given in square brackets. Only some attributes of either are explicated, indicated by [...].
- The arrow at the bottom indicates that the sub-tasks are performed sequentially. However, they are separate and individual, as emphasised by the dotted lines.
- Sub-tasks C1 and C2 (C = cognitive) are straightforward. The researcher reads articles to increase his/her knowledge of his/her primary research domain and, in the inter-/multidisciplinary research, also in research domain 2, D2, as well.
- Sub-tasks C3 and C4 are similar to C2, albeit with interruptions.
 - While reading an article in sub-task C3, the researcher is interrupted in this cognitive process because of a text reception problem, namely not understanding a technical term in D2, indicated as Sub-task TR3a (TR = text reception; the "a" indicates that this is the first of such interruptions; TR3b-n indicates that such interruptions may reoccur any number of times). He/she therefore first has to solve this information need

before returning to sub-task C3, typically by referring to a brief definition of the term. Since the document he/she is reading is in L1, the information required to solve the need is also in L1. After solving the information need satisfactorily, the researcher carries on with sub-task C3 itself.

- A similar problem occurs in executing sub-task C4. While reading an article in sub-task C4, the researcher is interrupted in this cognitive process because of a text reception problem, namely not understanding a word in L2, indicated as Sub-task TR4a and possible further occurrences, indicated by TR4b-n. He/she therefore first has to solve this information need before returning to sub-task C4. Since the document he/she is reading is in L2, the information required to solve the need requires an equivalent in either L2 (a synonym) or L1 (simply the 'meaning' from a bilingual dictionary) or an explanation in L2 (an 'understandable' explanation in a monolingual dictionary). After solving the information need satisfactorily, the researcher carries on with sub-task C4 itself.
- The diagram does not make provision for any extra-work task information needs or activities not related to the work task itself, for example external interruptions, etc.

Each of the variables discussed above can be plotted on a continuum, i.e. in terms of complexity there are multiple possibilities, not only three as mentioned above. The same applies to the other variables.

If all these variables and the continuum on which they can be plotted are combined in a single diagram, the result is an n-dimensional complex matrix of elements. This is further complicated if the standard requirements for information (i.e. being accurate, up-to-date, reliable and relevant) and the requirement that e-information should be available on the user's platform of choice, from desktop to mobile, are also taken into account. This complexity is represented diagrammatically in Figure 2, where only two characteristics of information are considered without the further variables or the fact that all of these elements can be plotted on a continuum.

It is therefore "clear that it is extremely complex to satisfy a specific information need of a specific user in a specific situation, taking into account all the possible permutations of the characteristics of users and information" (Bothma 2011). Bergenholtz and Gouws (2007: 584) come to the conclusion that a reference work is of high quality if it contains what the user requires — and *only* what the user requires — to satisfy his/her information needs. Such products are not feasible in a paper-based environment and are currently only to a very limited extent available in the e-environment. By means of modern information technologies it is, however, becoming feasible to address these issues.

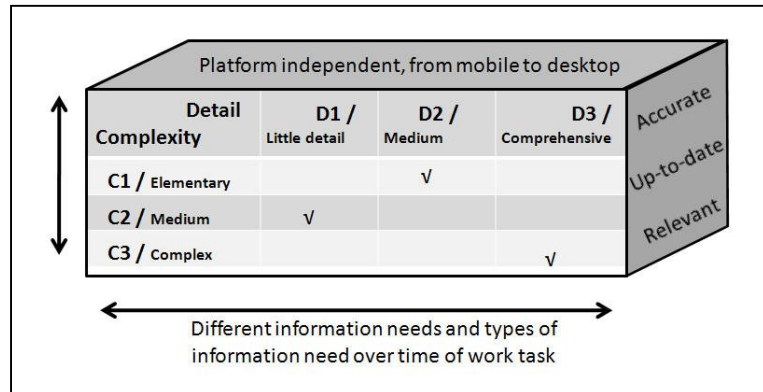


Figure 2: Matrix of possible characteristics of information (Bothma 2011)

4. Dictionaries as information tools

If someone does not know something, he/she is in what we call an 'extra-information tool' situation. This person could decide not to find out what he/she does not know. Alternatively, he/she could ask someone for help, look in a manual, a user guide, a handbook, a telephone directory or a dictionary or undertake a Google search or any other kind of information search in printed or electronic information tools. If he/she uses a dictionary (or a lexicon or an encyclopaedia or whatever name the lexicographical tool might have), the potential dictionary user is in an 'extra-lexicographical' situation. Looking in one or more dictionaries is a 'dictionary use' situation. In some cases the use of a dictionary would be the best choice but in many others certainly not. We therefore prefer to speak about 'extra-information tool' situations. We can distinguish between four basic situations:

- Cognitive situations
- Communicative situations
- Interpretative situations
- Operative situations

In **cognitive situations** the potential user has a need for knowledge of some kind. The purpose is not to use it in the concrete situation, e.g. during text reception or in a specific situation to act or react in that situation, but simply to get the knowledge. The user wants to know or to have the sought-after knowledge and stores it in the brain for later use. The most commonly used informa-

tion tools in cognitive situations are dictionaries, handbooks, databases and different search engines used for searching the Internet.

In **communicative situations** the potential user has a problem with or doubts about the process of written or oral communication and needs some help or recommendation for the text problem. The most commonly used information tools in communicative situations are dictionaries and different search engines used for searching the Internet.

In **operative situations** the potential user has a need for assistance in connection with physical or mental operations. There is a need for guidelines for what and how the user has to act to carry out a certain instruction. The most commonly used information tools in operative situations are user guides, manuals, road maps and search engines.

In **interpretative situations** the potential user has a need for assistance to interpret a non-language sign of some kind, e.g. a traffic sign. The right interpretation is normally needed to be able to act in the correct way. The most commonly used information tools in interpretative situations are information material from authorities and different organisations.

To understand the above classification of user situations it is important to make a clear distinction between an **operative situation** where someone simply intends to do something (act in a certain way) and a **cognitive situation** where someone selects knowledge that can be used later in another situation where a decision must be made to act in the most relevant way. Likewise, it is not the same to be in a **communicative** situation and a **cognitive** situation. In the former, the user has a text problem and, for example, does not understand a part of a text, cannot translate it or needs help to be able to go on with the production of the text. In the latter, the user selects knowledge that can be used later in connection with a text problem. The crucial point in this theory is that it is the situation that leads to the need for an information tool, and it is the nature of the help that is required that is essential for the description of the situation. The fact that information obtained by a user to fulfil his/her needs in a specific situation can also be used in another situation, is not relevant for the current concrete situation. It is at best a side effect (not a part of the situation) and not a demand for data presentation in the information tool currently used. In the same way it is not a relevant description of a specific user need if the user in another, perhaps similar situation, has learned something new or that the facts have changed. This is simply a new situation where the user would perhaps need another information tool. A tool is a tool. This means that a tool has a function according to the basic competence of the potential information tool user and his/her needs.

The function we know most about is the communicative function, or more precisely, the different communicative functions categorised as text reception, text production, translating and text correction (Bergenholtz and Tarp 2002; Bergenholtz and Tarp 2003; Tarp 2008b). These four main functions have to be divided according to different user competencies in different situations:

1. What is the users' native language?
2. At what level do users master their native language?
3. At what level do users master a foreign language?
4. How extensive is the users' experience of translating between the languages in question?
5. What is the level of the users' general cultural and encyclopaedic knowledge?

Other characteristics may be relevant for a particular dictionary, but the above are the most important for profiling a specific user group. Determining user characteristics is the first step that lexicographers should take to establish user needs, which are not abstract phenomena but related to concrete situations. Lexicographers should therefore identify these situations, distinguish them from one another and analyse them before establishing what the need of a specific type of user may be in a particular situation. The theory of lexicographic functions distinguishes between four main groups of situations of use. Communicative situations involve existing or planned written or oral communicative acts between two or more persons. In these instances, lexicographers intervene indirectly through their dictionaries when the interlocutors encounter some kind of communication problem that they attempt to solve by consulting a dictionary. Dictionaries that satisfy such needs are called **communicative dictionaries**.

For the kind of communicative dictionaries that are compiled to assist in the text reception process, one should first distinguish between general language and specialised language dictionaries. For the latter, one should distinguish between at least three levels of specialist: lay persons, semi-experts and experts. Some specialised language specialists however, distinguish between up to ten different kinds of specialist (Kalverkämper 1990). For each type one could describe a certain function and in the best-case scenario offer a special dictionary. This tool should also take into account whether the user has the explanation language as mother tongue or not. For translational tools one could differentiate much more with an indefinite number of functions and, thus, an indefinite number of needed monofunctional dictionaries. Such a solution is of course not practical. Well-known solutions with highly polyfunctional dictionaries can in our opinion not help the user accurately and quickly enough. In fact, in many cases they provide an information overload that leads to information stress or even information death, because the user gives up on finding what is needed from among the immense amount of available data. Instead of an offer that includes all kinds of bonus information, the meaning item in the communicative situation 'reception' is enough; all further detailed information is superfluous and disturbing. The real choice is whether a dictionary maker decides to provide only one monofunctional dictionary for all user types (e.g.

providing explanations in the mother tongue and intended for all users, including specialists, semi-experts and lay persons) or whether it is possible to compile different reception dictionaries for different user types. Although in both cases we refer to monofunctional dictionaries, it is clear that these are different levels of monofunctional dictionaries.

The following example shows results for three monofunctional dictionaries for fixed expressions taken from a single database (the three dictionaries are MEANING OF FIXED EXPRESSIONS, USE OF FIXED EXPRESSIONS and KNOWLEDGE ABOUT FIXED EXPRESSIONS). The intended user type has Danish as mother tongue. If someone reads a text and has a reception problem with some kind of fixed expression (idiom, proverb, phrase verb, quotation, etc.) the user can search in a reception dictionary with a search string for the whole fixed expression or for a part of it. A potential user reads a text with the expression *klæde sig i sæk og aske* (to dress in sackcloth and ashes) and is not sure what the expression means. He/she can use *i sæk og aske* as a search string and, by pressing the button 'understand an expression', find the entry below in the dictionary MEANING OF FIXED EXPRESSIONS. He/she gets only the meaning of the expression, nothing else:

Betydning udtryk for at vise ydre tegn på sorg eller anger	Meaning Expression for certain external signs of sorrow or regret
--	---

If the user knows the expression or has obtained it by a search as shown above, he/she may be in doubt as to how the expression can be used. He/she enters the search string *klæde i sæk og aske* (to put on sackcloth and ashes), presses the button 'Write a text' and finds the following entry in the dictionary USE OF FIXED EXPRESSIONS (this time, unlike in other entries, without collocation entries):

Faste vendinger iføre sig sæk og aske klæde sig i sæk og aske ligge i sæk og aske omvende sig i sæk og aske sidde i sæk og aske være iført sæk og aske	Fixed expression to put on sackcloth and ashes to wear sackcloth and ashes to lie in sackcloth and ashes to repent in sackcloth and ashes to sit in sackcloth and ashes to be dressed in sackcloth and ashes
Betydning udtryk for at vise ydre tegn på sorg eller anger	Meaning Expression for certain external signs of sorrow or regret
Grammatik nogen klæder sig i sæk og aske	Grammar Someone wears sackcloth and ashes
Eksempler Askeonsdag er den første af fastens	Quotations Ash Wednesday is the first of the 40

<p>40 dage. I middelalderen mødte de bodfærdige denne dag i kirke iført sæk og aske. Sammen med aflad skulle det syndige menneske også sone sin straf ved at få besked af præsten på at bede et vist antal Ave Maria — og evt. gå i sæk og aske og leve af vand og brød i en aftalt periode.</p>	<p>days of Lent. In the Middle Ages penitents went to church dressed in sackcloth and ashes on this day. Besides getting a letter of indulgence, the priest could also order the sinner to do penance by saying a number of Ave Marias — and if necessary to go dressed in sackcloth and ashes and to take nothing but water and bread for a set period.</p>
--	--

The above entry gives the grammatical information that the subject should be a person, as well as two examples that can be used for an own text.

However, the user may possibly want further information, for example, about the history behind the expression. He/she can then type in the same search string and press the button 'know more about a fixed expression'. In the dictionary KNOWLEDGE ABOUT FIXED EXPRESSIONS he/she then finds partly the same data as in the text production dictionary; additional data is, however, also provided about the type, style and history behind the expression:

type / stil idiom neutral	Expression type and style idiom neutral style
<p>Anmærkninger Udtrykket bruges både i Det Gamle og Det Nye Testamente, hvor personerne giver udtryk for stor sorg eller beklagelse, fx i Esters Bog 4.1: "Da Mordokaj fik at vide, hvad der var sket, flængede han sine klæder; han klædte sig i sæk og aske, gik rundt i byen og skreg højt og bittert". Man tager simple klæder på og strør aske på hovedet: "De fastede den dag og klædte sig i sæk og strøede aske på deres hoved og flængede deres klæder" (Første Makkabæerbog 3.47).</p>	<p>Note The expression occurs in the Old and the New Testaments when people express great sorrow or pity, e.g. in Esther 4.1: "When Mordecai perceived all that was done, Mordecai rent his clothes, and put on sackcloth and ashes, and went out into the midst of the city, and cried with a loud and a bitter cry". One wears simple clothes and scatter ashes on one's head: "They fasted that day and put on sackcloth and sprinkled ashes on their heads and tore their clothes" (1 Maccabees 3.47).</p>

The examples we have shown are from a quite specialised dictionary. The principles illustrated in these examples can be applied in exactly the same way in all other kinds of information tools, not only in (specialised) dictionaries.

5. Integrated information tools

From the preceding theoretical and practical examples it is clear that it is optimal and feasible to present the user with exactly the required information needed to solve an information-related problem. Care should therefore be taken to ensure that the user is not overwhelmed by an overload of information and that his/her information need is solved in a specific situation with exactly the required information.

The dictionary examples focused on the communication process, took into account both text reception and text production information needs, and also provided an example of a cognitive need. Information needs in communicative situations are best understood, especially in terms of how these relate to the use of dictionaries. This is evident from many of the publications of the Centre for Lexicography at the University of Aarhus (Centlex) and others within the framework of the function theory of lexicography. Finer or more detailed calibrations of cognitive situations and the nature of the information tools required to satisfy cognitive information needs within the framework of the function theory have not been researched in any detail. The following two sections will focus on the nature of information tools that can be used in any of the four situations. They will take the characteristics of information and users as discussed in section 2 as departure point.

To recap briefly:

- Information has specific characteristics that relate to complexity, detail and language.
- Users, from lay persons to experts, can be placed on a continuum.
- Users need exactly the information that solves their information need — not more and not less.

From the examples in section 4 it is clear that it is feasible to structure data in a dictionary environment and to provide (monofunctional) tools in such a way that these ideals can be met. Is this also feasible in the broader range of e-information tools as defined in section 1?

The rest of section 5 will focus on the nature of such integrated information tools, while section 6 will discuss how the data are to be organised and presented to provide these tools.

In the dictionary examples in section 4 the three dictionaries use one database. This implies that the lexicographer selected what data are necessary in any given situation to solve the user's information need and provided individualised tools for this purpose. For example, in the text reception case, all data not related to text reception are not presented to the user and the user is provided with a very limited view of the information in the database. Put differently: a filter has been applied to the database to ensure that only informa-

tion related to the specific search string and the specific information need is displayed. In other words, a user need with the characteristics [search string, text reception] is matched with data with the characteristics [text reception]. The same applies to a text production situation: a user need with the characteristics [search string, text production] is matched with data with the characteristics [text production]. Language attributes could also be added, for example to indicate that the language of the data to be displayed should be L1 or L2.

The task of the lexicographer or information specialist is to design tools that enable this matching process to take place.

In a communicative situation, finer granularity of information needs has been identified, namely text reception and text production. In the case of a cognitive situation no such finer granularity has yet been discussed and no such tools have been developed to date. For example, in the Centlex dictionaries as published by Ordbogen.com (*Ordbogen.com* 2010), there is simply a single button for cognitive functions, implying (to a certain extent) that the characteristics of information required in all cognitive situations are identical. In all cases, selecting a cognitive search option provides the user with everything that is in the database. However, our arguments in section 2 have made it clear that information has different characteristics, *inter alia* complexity and detail, in addition to language (in relation to the user's mother tongue, designated as L1/L2, etc.).

If the communicative function is subdivided, it is logical that the cognitive function should also be subdivided. (The same goes for the operative and interpretative functions, neither of which will be discussed in this article.) This process of subdivision and further refining of (sub) functions is also the task of the lexicographer and information specialist.

The technologies by means of which a user can access the information have already been hinted at. The first two include searching (and browsing) and filtering. A third technology, adaptive hypermedia, has to be added. Each of these will be discussed briefly.

Primary access to any information in a database occurs by means of searching and browsing and a combination of the two. This is discussed in more detail in Bothma (2011). Searching implies that the user enters a search string in a search box (with the required intellectual input to construct a search string), whereas browsing implies that the user navigates predefined links by simply clicking on the links. Currently, all e-information tools support searching and browsing. Many have an 'advanced search' option, but often this search option in e-dictionaries is not optimal and many of the functions of advanced search are not available (e.g. those found in e-journal platforms such as EbscoHost (*EBSCO Publishing* 2010) and ScienceDirect (*ScienceDirect* 2010)).

By executing a search in a standard search engine, all data that match the search string entered by the user are returned. In the Centlex dictionaries, however, the data are filtered — only data related to the specified function are returned. This means that the metadata of a data element are taken into account

to decide whether they should be displayed or not. In the Centlex dictionaries there is no explicit metadata, but metadata are implicit in terms of the fields in which the data occur. For example, when a text reception search is executed, data from a specific predefined set of fields are displayed; when a text production search is executed, data from a different predefined set of fields are displayed (even though there may be overlaps). This implies that the lexicographer or information specialist has decided that certain fields in a certain order are relevant for a text production information need and not for a text reception information need, and *vice versa*.

No such granularity, however, exists when a cognitive search is executed. If information required to satisfy a cognitive information need is placed in specific fields in a database, it would be possible to present the user with the information for the specific cognitive need only. If complexity and detail (the characteristics given in section 2) are used, it would result in information elements with nine characteristics, examples of which are given below:

Data element 1 [little detail, elementary]

Data element 2 [little detail, complex]

Data element 3 [comprehensive, elementary]

Data element 4 [comprehensive, complex]

Etc.

The same subdivisions can be made for any data element the lexicographer or information specialist includes in the database, for example historical comments, grammatical detail, etymology, multimedia, etc.

By providing a filtering mechanism it would be possible to provide the user in a cognitive situation with exactly the information he/she needs. However, at a practical level this may require too many predefined tools to be created — in the preceding example it would require nine tools. A more effective method to accomplish such granularity would involve providing the user with the tools to define the tool himself/herself by selecting specific check boxes that then display only the data that conform to the filter, as in the following:

Detail (select one)		Complexity (select one)	
Little	<input type="radio"/>	Elementary	<input type="radio"/>
Medium	<input type="radio"/>	Medium	<input type="radio"/>
Comprehensive	<input type="radio"/>	Complex	<input type="radio"/>

The user could then, by clicking two selections, specify one of nine possible combinations according to which the data would be filtered.

Filtering in this case would be on a per-case basis, i.e. each time the user uses the tool, he/she has to set the filter. It is also possible to set the filter as

part of a user profile and/or automated tracking of user behaviour, as explained in Bothma (2011).

Data are filtered in response to user needs and the content displayed for the user is adapted accordingly. The principles of adaptive hypermedia, as explained in Bothma (2011), are therefore applied.

Since user needs are not constant it is essential that the user remains in control of the information presented to him/her and he/she should be able to change the parameters of the search or profile or the nature of the adapted content at any stage.

In designing such an e-information tool, the lexicographer or information specialist can decide on a number of 'common' views of the database in line with the function theory, and provide additionally the possibility of a number of customised views. This would then enable the lexicographer and information specialist to conceive e-information tools "capable of meeting all the users' needs in specific types of situations" (Tarp 2009a: 292) by providing "dynamic articles [...] structured in different ways according to each type of search criteria", "articles that are especially adapted", allowing users to "define their own profile", resulting in "the 'individualization' of the lexical product, adapting to the concrete needs of a concrete user" (Tarp 2009b: 57, 59, 61). In the preceding quotation 'lexical product' can easily be replaced by 'e-information tool' to indicate that the principles apply much more widely than simply to dictionaries.

The principles discussed in relation to the subdivision of cognitive information needs can be expanded to include operative and interpretative information needs, and even further subdivisions of communicative information needs. Such subdivisions will not be attempted in this article.

The product of such development will be a set of information tools that can provide exactly the right information to a user in a given situation. Such a set of tools will be supported by one database; each tool will be a separate view of the database, either predefined by the lexicographer or information specialist or customised in terms of specific selections by the user. Each tool will be monofunctional, since the view that it provides on the data involves a single function at the level of granularity required by the user.

6. Data presentation

To provide such a customisable set of information tools requires a specific design of the database or set of databases. The database could physically be one database with a number of relational tables or a number of individual databases linked logically. For the user this distinction is irrelevant. However, at the programming level this distinction is important and requires very careful planning. The database could be a relational database with multiple tables and fields, or it could be an XML database. In a relational database the different attributes of the individual data elements will be reflected in the different fields that are used. In an XML database the attributes of individual data elements will be defined through metadata in an RDF schema (both concepts are briefly

discussed in Bothma (2011)). The underlying principles, however, remain the same.

One example of a physical single database or logical database (set of databases) will suffice. In the current dictionaries of Centlex published by Ordbogen.com (*Ordbogen.com* 2010), each of the 'main dictionaries' (*Ordbogen over faste vendinger*, *Musikordbogen* and *Regnskabsordbøgerne*) is supported by its own database, i.e. three physical databases. It is possible to search each of these databases individually, but it is also possible to do a search simultaneously over all of them (i.e. the three dictionaries mentioned as well as other dictionaries, depending on the user's subscription). For the sake of simplicity, reference will be made only to the afore-mentioned three dictionaries. In the second case the databases are the equivalent of a single logical database (even though they are not physically one database). It would be possible to integrate the three separate databases into one physical database, even though this will be a non-trivial and extremely complex venture.

From the user's side, however, there is no difference. When he/she logs into the site, he/she has the option to search 'all dictionaries' with a single search string (i.e. all databases) or to immediately filter the search by selecting a domain-specific dictionary. If the user decides on a domain-specific dictionary, he/she then has the option to filter the search further by deciding whether the data are required for (for example) text reception, text production or a cognitive function (as illustrated in section 4). The user therefore selects a further filter to provide a customised view of the database. These filters imply that the data are uniquely identifiable as relevant to a text reception, text production or cognitive need.

The structure of the database reflects the access methodologies envisaged by the lexicographer or information specialist. The levels of granularity that can be selected by the user must be matched by the attributes of the individual data elements in the database. This can be illustrated by the following example:

For a text production dictionary, the lexicographer or information specialist needs to decide what data elements are important for the user in a text production environment. These could include (but are not limited to) the meaning of an element, grammatical attributes, collocations and examples. Each of these will then be placed in a separate field or individually marked up with metadata in an RDF schema. Since this granularity is already reflected in the database, it is a trivial task to provide the user with a filter in the text production dictionary to see only meaning, grammatical elements, collocations or examples. However, it could be possible to provide even finer granularity, for example at the grammatical level, to put gender, form of plural, etc., in separate database fields (or to mark them up separately as metadata). A user could, in such a case, select to see only the gender of a noun when this is the only information that is needed in a text production task. This is currently partially possible in the *Base lexicale du français* (2010). The same principle applies to the other sub-

fields for text production, as well as to the cognitive, operative and interpretative information needs.

If the data for all these sub-functions are put into different database fields or are properly marked up by means of metadata, it is trivial to provide filters to show only these data. Interface design could, however, be more problematical. The screen could become quite cluttered and users may get confused about how to navigate between and select options. A possible solution would be to provide popup windows or dropdown menus. In both cases, the user could be given the choices between searching and getting more options, as in Figure 3 below.

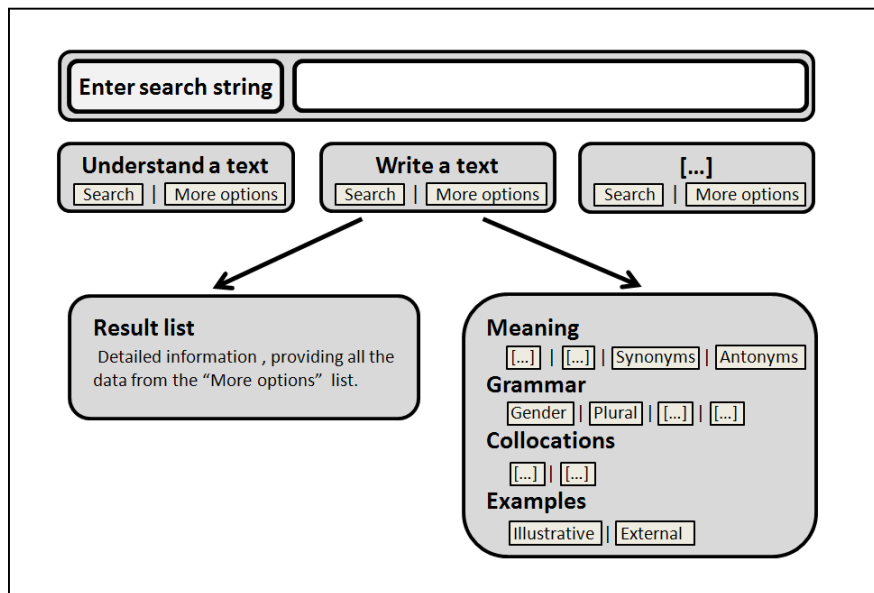


Figure 3: Popup or dropdown menus for customisation

In a text production environment, therefore, the user enters a search string, and then clicks on 'Search' to see all information regarding text production. He/she can also click on 'More options' and then select to see, for example, only 'Gender'. Designing detailed submenus for other functions is fairly trivial, but depends on the granularity that the lexicographer or information specialist adopted for the data.

Hence, by using a relational database structure with a high level of granularity or by marking up data in an XML database with detailed metadata it is possible to provide a highly customised view of the database. Such a view will satisfy the specific information need of the user in a given situation.

When in a specific work task situation a user knows that he/she will need (for example) only collocations, he/she could set this criterion as part of his/her profile at the beginning of a consultation session, thereby obviating the need to go through the selection process in each query. As indicated earlier, it remains essential that the user be able to reset this profile at any stage in his/her work.

The database structures discussed thus far imply that the lexicographic or information specialist team create all data in the database themselves. Currently it is very common in reference works to refer to secondary resources as well as to general web resources, as is generally the case in Wikipedia (2010) by means of links selected by the domain specialist. This principle can easily be incorporated in the structures outlined in this section, either in separate data fields or as part of any data field within which such a reference would be relevant. It is, however, not necessary for the authoring team to select all such cross-referenced data (such as a single encyclopaedia entry or YouTube video) themselves. Links can be provided to selected large open access data sets as well, as discussed in Bothma (2011). Such links are selected only when the user requires additional information to satisfy his/her information need in a specific situation (typically in a cognitive situation). Both these options may well lead to information overload, but it is the user's choice to decide whether to click on such external links or not.

The user therefore customises the information that he/she accesses in any given information need situation and remains in control of exactly how much or how little information is made available to him/her by the set of e-information tools.

7. Requirements for e-information tools

A good tool is one that is able to fulfil the needs of a certain user group by giving quick access to the data and by providing relevant and correct data in a way understandable to this user group. It does not send the user into information death by burdening him/her with a vast amount of unneeded data.

Such a tool must necessarily be a monofunctional tool prepared for a certain user group with a certain type of information need and in a certain type of user situation. However, most printed and electronic lexicographic information tools are not monofunctional, but highly polyfunctional. They are intended for a very broad user group with very different user needs and for quite different usage situations. One therefore often sees as a main thesis in many metalexicographical papers that the more data and the more different data types are provided (even though the user may not be looking for these data) the better the dictionary is (Varantola 2002: 31) — a direct contradiction of the principles of the function theory and what is proposed in this paper.

Lexicographers and information specialists have the possibility to prepare certain tools in such a way that the information type presented corresponds to the general knowledge about needs, e.g. one only needs definitions (semantic items) in a monofunctional tool for a certain user group with a need for help in reception problems. Having created only one database we can offer different monofunctional tools, e.g. the following from a database with 14 fields, searching in four of them for a monofunctional text production dictionary:

Fields searched + order of search	Field	Order in dictionary entry
1	1. Core field	
	2. Meaning	2
	3. Further meaning item	3
	4. Grammar	4
	5. Remark(s)	
	6. Internet link	
2	7. Fixed expression(s)	1
	8. Style	
	9. Classification of the fixed expression	
3	10. Collocation(s)	5
4	11. Example(s)	6
	12. Synonym(s)	7
	13. Antonym	8
	14. Association(s)	

Depending on their specific information needs in a given situation, some users may prefer another search order, searches in more fields, the data presented in another order or presented from different/additional fields. In the end there would be the possibility of an unlimited number of different dictionaries based on the customisation principles discussed above and the number of fields in the database. The question of the number of users making use of such advanced searches and presentations is not really an important point of argumentation. Whether it is 2% or 20% of the users does not matter. It would simply be an unnecessary technical possibility if they were not to be used by any user at all.

However, the different conceptions with the described technical possibilities are really important if we are dealing with information tools with relevant, correct and up-to-date data. A good tool must be completed within a few years. If not, the content is not up to date when the dictionary is published. In Molecular Biology, for example, we have seen that the amount of knowledge doubles within 18 months. It may be argued that development is not so rapid in all fields. In the Accounting dictionary project, however, that Aarhus Uni-

versity is undertaking in cooperation with colleagues from Valladolid, there are so many national and international changes in terminology and in the official rules for using the terms that a 3- or 4-year old dictionary will in many cases be misleading. A dictionary is of course never finished in the sense that one can stop working on it to refine and update the data. If one were to stop updating the data, the dictionary would become obsolete and not be usable except as a documentation tool of an earlier period. In some cases (depending on the subject area) this would occur after only a few years and in others the dictionary might stay current for perhaps more than 20 years.

An information tool with updated, relevant and correct data is unfortunately seldom produced in the required short time or presented as a really helpful monofunctional tool.

8. Conclusion

One of the reasons for the lack of innovation in e-lexicography is that lexicography is usually treated as a part of linguistics and lexicographic tools are primarily compiled by specialists with a linguistic background. Our main thesis, namely that *lexicography is not a part of linguistics but a part of information science* does not support this line of thought. We need different kinds of experts for different tools. For example, for a music dictionary one needs to involve experts in at least three fields: firstly an expert in lexicography, secondly an expert in lexicographical databases, and thirdly an expert in music theory and history of music. For a general language text production dictionary for L1-speakers one also needs an expert in lexicography and an expert in lexicographical databases, but then obviously a linguist. Thus, different kinds of experts are needed for different tools, and this does not necessarily always include a linguist. We do, however, need tools for special problems, not poly-functional tools. The expert called lexicographer could be called a special kind of information science expert.

Our argument does not imply that the administrative affiliation of all lexicographers at universities must be changed. Some lexicography units could — if they deal mostly with lexicography for communicative needs — stay in language departments. But normally and in most cases it would be more natural for lexicography to be treated as a part of information science. In our opinion, the lexicographer could be seated in an interdisciplinary centre that belongs at the same time to the faculty of science, the faculty of social sciences and the faculty of humanities. What we need in the information society (also at universities) is flexibility.

References

- ACM Digital Library. 2010. [Online]. Available: <http://portal.acm.org>. Accessed October 2010.
Amazon.com. 2010. [Online]. Available: <http://www.amazon.com>. Accessed October 2010.

- Base lexicale du français*. 2010. [Online]. Available: <http://ilt.kuleuven.be/blf>. Accessed October 2010.
- Bergenholtz, H.** 2010. Access to and Presentation of Needs-adapted Data in Monofunctional Internet Dictionaries. Presentation at the Acceptance of an Honorary Doctorate at the University of Valladolid, Valladolid, Spain, 17 June 2010.
- Bergenholtz, H. and I. Bergenholtz.** 2007. Kvaliteten af de leksikografiske definitioner i ordbøger, leksika og encyklopædier. *LexicoNordica* 14: 11-34.
- Bergenholtz, H. and I. Bergenholtz.** 2011. A Dictionary is a Tool, a Good Dictionary is a Monofunctional Tool. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *E-Lexicography*. London/New York: Continuum.
- Bergenholtz, H. and R.H. Gouws.** 2007. Korrek, volledig, relevant. Dít is die vraag aan leksikografiese definisies. *Tydskrif vir Geesteswetenskappe* 47(4): 568-586.
- Bergenholtz, H. and R.H. Gouws.** 2010. A New Perspective on the Access Process. *Hermes, Journal of Language and Communication Studies* 44: 1-25.
- Bergenholtz, H. and S. Tarp.** 2002. Die moderne lexicographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen. *Lexicographica* 18: 253-263.
- Bergenholtz, H. and S. Tarp.** 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 171-196.
- Bergenholtz, H. and S. Tarp.** 2005. Verteilungsstrukturen in Wörterbüchern. Barz, I., H. Bergenholtz and J. Korhonen (Eds.). 2005. *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*: 119-126. Frankfurt am Main: Peter Lang.
- Bothma, T.J.D.** 2011. Filtering and Adapting Data and Information in the Online Environment in Response to User Needs. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *E-Lexicography*. London/New York: Continuum.
- Cosijn, E. and P. Ingwersen.** 2000. Dimensions of Relevance. *Information Processing and Management: An International Journal* 36(4): 533-550.
- Cosijn, E.** 2006. Relevance Judgements within the Context of Work Tasks. 2006. *Proceedings of the 1st International Conference on Information Interaction in Context, IliX*: 20-29. New York: ACM.
- EBSCO Publishing*. 2010. [Online]. Available: <http://www.ebscohost.com>. Accessed October 2010.
- Gouws, R. and P. Leroyer.** 2009. Verhoogde leksikografiese toeganklikheid in die oorgang van 'n toeristewoordeboek na 'n toeristegids as naslaanbron. *Tydskrif vir Geesteswetenskappe* 49(1): 145-159.
- Haas, M.R.** 1962. What Belongs in a Bilingual Dictionary? Householder, F.W. and S. Saporta (Eds.). 1962. *Problems in Lexicography*: 45-50. Bloomington: Indiana University.
- Ingwersen, P.** 2001. Cognitive Information Retrieval. *Annual Review of Information Science and Technology* 34: 3-51.
- Ingwersen, P.** 2007. Context in Information Interaction — Revisited 2006. Bothma, T.J.D. and A. Kaniki (Eds.). 2007. *ProLISSA 2006: Proceedings of the Fourth Biennial DISSAnet Conference. Farm Inn, Pretoria, 2-3 November*: 13-23. Pretoria: Infuse.
- Ingwersen, P. and K. Järvelin.** 2004. Context in Information Interaction. Bothma, T. and A. Kaniki (Eds.). 2004. *Progress in Library and Information Science in Southern Africa (ProLISSA): Proceedings of the Third Biennial DISSAnet Conference*: 301-310. Pretoria: Infuse.
- Ingwersen, P. and K. Järvelin.** 2005. *The Turn: Integration of Information Seeking and Retrieval in Context*. Dordrecht: Springer.

- Järvelin, K. and P. Ingwersen.** 2010. User-oriented and Cognitive Models of Information Retrieval. Bates, M. (Ed.). 2010. *Encyclopedia of Library and Information Sciences*: 5521-5534. London: Taylor & Francis.
- Kalverkämper, H.** 1990. Der Begriff der »Fachlichkeit« im modernen Italienischen. Lexikalische Systematik und textuelle Integration. *Quaderni di Semantica. Rivista internazionale di semantica teorica e applicata* 11: 79-115.
- KNOWLEDGE ABOUT FIXED EXPRESSIONS = Henning Bergenholtz under medvirken af Esben Bjærg. 2010. *Betydning af faste vendinger*. Database and Design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Leroyer, P.** 2009. Balancing the Tools: The Functional Transformation of Lexicographic Tools for Tourists. Nielsen, S. and S. Tarp. 2009b. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: 103-122. (Terminology and Lexicography Research and Practice, Volume 12). Amsterdam: John Benjamins.
- LexisNexis 2010 International Workplace Productivity Survey*. 2010. [Online]. Available: <http://www.multivu.com/players/English/46619-LexisNexis-International-Workplace-Productivity-Survey>. Accessed October 2010.
- Lyman, P. and H.R. Varian.** 2003. *How Much Information?* 2003. [Online]. Available: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003>.
- MEANING OF FIXED EXPRESSIONS = Henning Bergenholtz under medvirken af Esben Bjærg. 2010. *Betydning af faste vendinger*. Database and Design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Nielsen, S.** 2009. Reviewing Printed and Electronic Dictionaries. A Theoretical and Practical Framework. Nielsen, S. and S. Tarp 2009b. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: 23-42. (Terminology and Lexicography Research and Practice, Volume 12). Amsterdam: John Benjamins.
- Nielsen, S. and S. Tarp.** 2009a. Introduction: Nothing is More Practical Than a Good Theory. Nielsen, S. and S. Tarp. 2009b. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: ix-xi. (Terminology and Lexicography Research and Practice, Volume 12). Amsterdam: John Benjamins.
- Ordbogen.com*. 2010. [Online]. Available: <http://www.ordbogen.com>. Accessed October 2010.
- ScienceDirect*. 2010. [Online]. Available: <http://www.sciencedirect.com>. Accessed October 2010.
- Tarp, S.** 2007. Lexicography in the Information Age. *Lexikos* 17: 170-179.
- Tarp, S.** 2008a. The Third Leg of Two-legged Lexicography. *Hermes, Journal of Language and Communication Studies* 40: 117-131.
- Tarp, S.** 2008b. *Lexicography in the Borderland Between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learners' Lexicography*. (Lexicographica. Series Maior 134). Tübingen: Max Niemeyer.
- Tarp, S.** 2008c. Revival of a Dusty Old Profession. *Hermes, Journal of Language and Communication Studies* 41: 175-188.
- Tarp, S.** 2009a. Reflections on Lexicographical User Research. *Lexikos* 19: 275-296.
- Tarp, S.** 2009b. Reflections on Data Access in Lexicographic Works. Nielsen, S. and S. Tarp. 2009b. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: 43-65. (Terminology and Lexicography Research and Practice, Volume 12). Amsterdam: John Benjamins.

- Tarp, S.** 2011. Lexicographical and Other e-tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *E-Lexicography*. London/New York: Continuum.
- Tono, Y.** 2010. A Critical Review of Lexicographical Functions. *Lexicon* 40: 1-26.
- USE OF FIXED EXPRESSIONS = Henning Bergenholtz under medvirken af Esben Bjarge. 2010. *Brug af faste vendinger*. Database and Design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Varantola, K.** 2002. Use and Usability of Dictionaries: Common Sense and Context Sensibility. Corréad, M.-H. (Ed.). *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*: 30-44. Göteborg: EURALEX.
- Verlinde, S., P. Leroyer and J. Binon.** 2009. Search and You Will Find. From Stand-alone Lexicographic Tools to User-driven Task and Problem-oriented Multifunctional Leximats. *International Journal of Lexicography* 23(1): 1-17.
- Wikipedia.** 2010. [Online]. Available: 2010. <http://en.wikipedia.org>. Accessed October 2010.

Voorbeelde en ooradressering in tweetalige woordeboeke*

Herman L. Beyer, *Departement Taal- en Literatuurstudie, Universiteit van Namibië, Windhoek, Namibië en Departement Afrikaans en Nederlands, Universiteit Stellenbosch, Stellenbosch, Suid-Afrika (hbeyer@unam.na)*

Opsomming: Adresseringsekwivalensie geld in 'n woordeboekartikel wanneer al die vertaal-ekwivalente in die vertaalekwivalentparadigma neerslag vind in die doeltaalvoorbeelde in die koteksafdeling en elke doeltaalvoorbeeld 'n vertaalekwivalent bevat wat in die vertaalekwivalentparadigma voorkom. Die versteuring van hierdie ekwivalensie vind plaas deur prosedures van onderadressering (wanneer nie al die vertaalekwivalente in die vertaalekwivalentparadigma in doeltaalvoorbeelde in die koteksafdeling neerslag vind nie) en ooradressering (wanneer doeltaalvoorbeelde vertaalekwivalente bevat wat nie in die vertaalekwivalentparadigma voorkom nie). In 'n polifunksionele tweetalige woordeboek verteenwoordig adresseringsekwivalensie die ideale adresseringstruktuur tussen vertaalekwivalentparadigma en koteksafdeling. In Afrikaanse tweetalige woordeboeke is adresseringsekwivalensie egter skaars, maar dit blyk in tweetalige woordeboeke van ander tale ook die geval te wees. Hoewel gevalle van onderadressering relatief maklik geneutraliseer kan word, bevind 'n dieper ondersoek na gevalle van ooradressering dat die meeste doeltaalvoorbeelde wat simptome van ooradressering blyk te wees, wel tot kommunikatiewe ekwivalensie bydra ten spyte van 'n oënskynlik gebrekkige adresseringstruktuur. Dit is omdat hierdie doeltaalvoorbeelde nie as ware voorbeelde (wat die optrede van vertaalekwivalente illustreer) funksioneer nie, maar in der waarheid optree as kotekstuele spesifiseerders wat 'n funksionele prosedure van kontekstuele vertaalekwivalentuitsluiting verteenwoordig, waardeur die teikengebruiker gewys word op die nie-bruikbaarheid van die aangebode vertaalekwivalente in die vertaalekwivalentparadigma in bepaalde kontekste. Hierdie bevinding noop 'n herwaardering van adresseringstrukture in die algemeen en die beginsel van adresseringsekwivalensie in die besonder.

Slutelwoorde: ADRESSERINGSEKWIVALENSIE, ADRESSERINGSTRUKTUUR, BRONTAALKOTEKSINSKRYWING, DIREKTE ADRES, DOELTAALKOTEKSINSKRYWING, EKWIVALENTVERHOUDING, FUNKSIE, INDIREKTE ADRES, KOMMUNIKATIEWE EKWIVALENSIE, KONTEKS, KONTEKSTUALISERENDE INSKRYWINGS, KONTEKSTUELE VERTAALEKWIVALENTUITSLUITING, KOTEKS, KOTEKSAFDELING, KOTEKSINSKRYWING, KONTEKSTUELE ILLUSTRATOR, KOTEKSTUELE SPESIFISEERDER, LEMMA, ONDERADRESSERING, OORADRESSERING, POLIFUNKSIONELE TWEETALIGE WOORDEBOEK, PRIMÊRE ADRES, SEKONDÊRE ADRES, SKYNADRESSERING, VERTAALEKWIVALENT, VERTAALEKWIVALENTPARADIGMA, VOORBEELD, WARE VOORBEELD

* Hierdie artikel is 'n verwerking van 'n gedeelte van 'n hoofstuk uit 'n D.Litt.-proefskrif, 'n *Metaleksikografiese ondersoek na kontekstleiding in Afrikaanse vertalende woordeboeke*, wat in April 2006 deur die Universiteit Stellenbosch, Stellenbosch, Suid-Afrika, aanvaar is.

Abstract: Examples and Overaddressing in Bilingual Dictionaries. Addressing equivalence prevails in a dictionary article when all the translation equivalents in the translation equivalent paradigm are represented in the target language examples in the cotext section and each target language example contains a translation equivalent that appears in the translation equivalent paradigm. The disruption of this equivalence takes place as a result of the procedures of underaddressing (when not all the translation equivalents in the translation equivalent paradigm are represented in the target language examples in the cotext section) and overaddressing (when target language examples contain translation equivalents that do not appear in the translation equivalent paradigm). In a polyfunctional bilingual dictionary addressing equivalence represents the ideal addressing structure between translation equivalent paradigm and cotext section. In Afrikaans bilingual dictionaries, addressing equivalence is rare, but this seems to be the case in bilingual dictionaries of other languages also. Although instances of underaddressing can be neutralised relatively easily, a deeper investigation into overaddressing has found that most target language examples that seem to be symptomatic of overaddressing indeed contribute to communicative equivalence in spite of a would-be deficient addressing structure. This is because these target language examples do not function as true examples (that illustrate the behaviour of translation equivalents), but are in fact cotextual specifiers that represent a functional procedure of contextual translation equivalent exclusion, through which the target user is advised of the non-usability of the translation equivalents offered in the translation equivalent paradigm in certain contexts. This finding requires a re-appreciation of addressing structures in general and the principle of addressing equivalence in particular.

Keywords: ADDRESSING EQUIVALENCE, ADDRESSING STRUCTURE, COMMUNICATIVE EQUIVALENCE, CONTEXT, CONTEXTUAL TRANSLATION EQUIVALENT EXCLUSION, CONTEXTUALISING ENTRY, COTEXT ENTRY, COTEXT SECTION, COTEXT, COTEXTUAL ILLUSTRATOR, COTEXTUAL SPECIFIER, DIRECT ADDRESS, EQUIVALENT RELATION, EXAMPLE, FALSE ADDRESSING, FUNCTION, INDIRECT ADDRESS, LEMMA, OVERADDRESSING, POLYFUNCTIONAL BILINGUAL DICTIONARY, PRIMARY ADDRESS, SECONDARY ADDRESS, SOURCE LANGUAGE COTEXT ENTRY, TARGET LANGUAGE COTEXT ENTRY, TRANSLATION EQUIVALENT PARADIGM, TRANSLATION EQUIVALENT, TRUE EXAMPLE, UNDERADDRESSING

1. Inleiding

Volgens artikelstruktuur kan die tipiese woordeboekartikel (*wa*) in 'n tweetalige woordeboek in twee onmiddellike stukke uiteenval, naamlik die vormkommentaar (*vk*) en die semantiese kommentaar (*sk*) (vgl. o.m. Wiegand 1988, 1996, 2000). Hierdie tweedeling kan geïllustreer word aan die hand van die woordeboekartikel van die lemma **sedert** uit PHAROS wat deur die voorafgaande afkortings geannoteer is:

- (1) $_{wa}[_{vk}[\mathbf{se} \cdot \mathbf{dert} \text{ prep., conj.}]_{vk} \text{ sk}[\text{since; for; } \sim \text{ die oorlog since the war; } \sim \text{ lank for a long time past, for ages.}]_{sk}]_{wa}$

Die vormkommentaarstuk bevat data oor die vorm van die lemma: In die geval in (1) word die ortografie, hoofklem, lettergreepverdeling en woordsoortkate-

gorieë waartoe die lemma behoort, weergegee. Die semantiesekommentaarstuk bevat die semantiese data oor die lemma en kan in twee verdere komponente verdeel word, soos hieronder aangetoon:

- (2) $sk_{[vep][since; for;]}_{vep} ka[\sim die oorlog since the war; \sim lank for a long time past, for ages.]_{ka} sk$

Die eerste komponent is die vertaalekwivalentparadigma (*vep*), wat die aangebode (twee) vertaalekwivalente bevat. Die tweede komponent is die koteksafdeling (*ka*), wat vier koteksinskrywings bevat: twee brontaalkoteksinskrywings en twee korresponderende doeltaalkoteksinskrywings.

2. Koteksinskrywings

Koteks verwys na die tipiese linguistiese omgewing waarin die lemma of 'n bepaalde vertaalekwivalent in werklike taalgebruik voorkom en word in die vorm van koteksinskrywings in woordeboekartikels weergegee (Gouws en Prinsloo 2005: 127). Koteksinskrywings val in twee klasse uiteen, naamlik *voorbeeld* en *kollokasies*. In hierdie artikel word op voorbeelde gefokus; vergelyk Gouws (1989: 227-228), Van Niekerk (1992: 256 e.v.) en Cop (1990, 1991) vir besprekings oor kollokasies in woordeboeke.

3. Voorbeelde

Jacobsen et al. (1991: 2782) en Rademeyer (1992: 5-6, 9) wys op die uiteenlopende terme wat ten opsigte van voorbeelde in die literatuur voorkom: Landau (2001: 115) gebruik die term *illustratiewe frases*; Al-Kasimi (1977: 88) maak melding van *illustratiewe voorbeelde*; Kromann et al. (1991: 2721, 2722) gebruik *idiosinkratiese konstruksies* en *vry frases*, terwyl *vry konstruksies* in Kromann et al. (1991a: 2770) voorkom; laasgenoemde term word ook deur Rademeyer (1992: 9) en Beyer (2009: 18) gebruik; by Zgusta (1971: 140) en Cop (1990: 37) word die term *vry kombinasies* aangetref. Gouws (1989: 228) en Wiegand (1988: 582) gebruik *voorbeeldsinne* en *voorbeeld* (in volsinvorm) onderskeidelik. Die term *voorbeeld* word in hierdie studie aangeneem, met die klas voorbeelde wat op sintagmatiese vlak in die twee subklasse *voorbeeldsinne* en *voorbeeldfrases* uiteenval.

3.1 Funksies

Uit die literatuur en woordeboekdata kan minstens tien funksies van voorbeelde in standaard verklarende en vertalende woordeboeke geïdentifiseer word. Volgens Lombard (1992: 149) het voorbeelde in 'n woordeboek in die breë 'n semantiese, sintaktiese en pragmatiese funksie: Die semantiese funksie behels die weergee van voorbeelde "in soverre dit die gebruiker met die dekodering van

betekenisinligting kan help"; die sintaktiese funksie is om "aspekte van die lemma se optrede in sinsverband te illustreer"; die pragmatiese funksie is "om taalgebruik so te illustreer dat die gebruiker gehelp word om self met die taal om te gaan" (Lombard 1992: 149-150). Hierdie driedeling strook met 'n kommunikatiewe perspektief op inligting (vgl. Van Cuilenburg et al. 1992: 12) en voorsien 'n nuttige raamwerk waarbinne die geïdentifiseerde funksies van voorbeelde geklassifiseer kan word. Die volgende funksies is naamlik geïdentifiseer:

(a) Semantiese funksie

Voorbeelde help die gebruiker om die betekenis van die lemma en/of vertaalekwivalente te agterhaal en ondersteun dus ekwivalentdiskriminasie (Zgusta 1971: 142; Al-Kasimi 1977: 90; Jackson 1988: 58; Manley et al. 1988: 291; Wiegand 1988: 580; Gouws 1989: 229; Kromann et al. 1991a: 2771; Lombard 1992: 150; Nesi 1996: 204; Szende 1999: 199; Prinsloo en Gouws 2000: 139).

(b) Sintaktiese funksies

(i) Voorbeelde bied die gebruiker sintaktiese illustrasie tydens teksresepsie, teksproduksie en vertaling uit en in die bron- of doeltaal (Zgusta 1971: 263; Al-Kasimi 1977: 90; Jackson 1988: 58; Wiegand 1988: 580; Gouws 1989: 229; Kromann et al. 1991a: 2771; Lombard 1992: 150; Nesi 1996: 204; Szende 1999: 199; Prinsloo en Gouws 2000: 154; Landau 2001: 115).

(ii) Voorbeelde kan kollokasies illustreer (Gouws 1989: 232).

(iii) Voorbeelde kan morfosintaktiese relasies illustreer, byvoorbeeld die verbuiging al dan nie van 'n adjektief in attributiewe optrede; vergelyk die volgende artikel uit PHAROS, waarin die kotekstuele inskrywings illustreer dat hoewel die vertaalekwivalent *sanitary* nie in attributiewe optrede verbuiging ondergaan nie, die leksikale item wat deur die lemma verteenwoordig word wel in sodanige optrede verbuiging ondergaan:

sanitêr [...] sanitary; ~e *doekie* sanitary pad/towel; ~e *tegniek* sanitary engineering; ~e *tegnikus* sanitary engineer.

Hierdie funksie sal 'n uitgebreide rol speel in 'n vertalende woordeboek waarin 'n analitiese taal met 'n flekterende taal gekoördineer word, soos in 'n Afrikaans–Oshindonga-woordeboek.

(c) Pragmatiese funksies

(i) Die gebruiker kan na analogie van voorbeelde sy/haar eie konstruksies in die bron- en/of doeltaal vorm (Zgusta 1971: 267; Jackson 1988: 58;

Lombard 1992: 150; Szende 1999: 200). Zgusta (1971: 267) beskryf hierdie funksie as die generatiewe krag van voorbeelde.

(ii) Voorbeelde kan die stilistiese waarde van 'n leksikale item in taalgebruik illustreer (Al-Kasimi 1977: 91; Landau 2001: 208; Gouws 1989: 231).

(iii) Voorbeelde kan kulturele data bevat (Al-Kasimi 1977: 89; Gouws 1989: 232; Szende 1999: 218).

(iv) Voorbeelde kan die gebruiker aan die bestaan van lokusies en clichés herinner (Szende 1999: 199).

(v) Voorbeelde kan aan die gebruiker leiding bied in leestekengebruik (Gouws 1989: 231).

(vi) Voorbeelde kan bewys dat 'n leksikale item of 'n spesifieke betekenis van 'n leksikale item bestaan (Al-Kasimi 1977: 89).

Die vervulling van hierdie funksies deur voorbeelde sal die gebruiker help om die relevante leksikale item te *ken* (Prinsloo en Gouws 2000: 144).

3.2 Adresseringstruktuur

Vir voorbeelde om hulle funksies ten volle te vervul, is dit noodsaaklik dat hulle sistematies en konsekwent in die woordeboek aangewend word. Dit veronderstel dat daar relasies geld tussen (a) die korresponderende bron- en doeltaalvoorbeelde in die koteksafdeling, (b) die brontaalvoorbeelde en die lemma en (c) die doeltaalvoorbeelde en die vertaalekwivalentparadigma. Hierdie soort relasie in die woordeboek word beskryf in terme van die konsep *adresseringstruktuur*: die woordeboekinterne stelsel waarvolgens een mikrostrukturele item volgens bepaalde prosedures aan 'n ander mikrostrukturele item gerig of geadresseer word (Gouws en Prinsloo 2005: 135). Wat relasie (a) betref, word doeltaalvoorbeelde aan hulle korresponderende brontaalvoorbeelde geadresseer aangesien hulle in wese doeltaalvertalings van die brontaalvoorbeelde is. Wat relasie (b) betref, word brontaalvoorbeelde aan die lemma geadresseer aangesien hulle die optrede van die lemma in werklike taalgebruik illustreer. Wat relasie (c) betref, word doeltaalvoorbeelde oënskynlik aan onderskeie vertaalekwivalente in die vertaalekwivalentparadigma geadresseer aangesien hulle die optrede van die betrokke vertaalekwivalente in werklike taalgebruik illustreer. Die adresseringstruktuur wat in die woordeboekartikel in (1) geld kan skematies aangetoon word aan die hand van die annotering van die artikel in (3) en die daaropvolgende tabel:

- (3) $wa[sk[se:dert^{lemma} prep., conj.]]_{vk} sk[vep[since^{ve1}; for^{ve2};]_{vep} ka[\sim die oorlog^{bvb1} since the war^{dwb1}; \sim lank^{bvb2} for a long time past^{dwb2}, for ages^{dwb3}.]_{ka}]_{sk}]_{wa}$

Tabel 1: Die adresseringstruktuur van die woordeboekartikel in (3)

Gerigte inskrywing	Adresinskrywing	Lemma as adres?
since (<i>ve1/vep</i>)	sedert (<i>lemma/vk</i>)	+
for (<i>ve2/vep</i>)	sedert (<i>lemma/vk</i>)	+
sedert die oorlog (<i>bvb1/ka</i>)	sedert (<i>lemma/vk</i>)	+
sedert lank (<i>bvb2/ka</i>)	sedert (<i>lemma/vk</i>)	+
since the war (<i>dvb1/ka</i>)	sedert die oorlog (<i>bvb1/ka</i>)	-
since the war (<i>dvb1/ka</i>)	since (<i>ve1/vep</i>)	-
for a long time past (<i>dvb2/ka</i>)	sedert lank (<i>bvb2/ka</i>)	-
for a long time past (<i>dvb2/ka</i>)	for (<i>ve2/vep</i>)	-
for ages (<i>dvb3/ka</i>)	sedert lank (<i>bvb2/ka</i>)	-
for ages (<i>dvb3/ka</i>)	for (<i>ve2/vep</i>)	-

Elkeen van die inskrywings in die "Adresinskrywing"-kolom in Tabel 1 tree op as bewerkingsseenhede in eie reg waaraan ander inskrywings geadresseer is, wat impliseer dat die lemma nie die enigste bewerkingsseenheid in 'n woordeboekartikel hoef te wees nie. Gevolglik kan onderskei word tussen lemmatiese en nie-lemmatiese adressering in die adresseringstruktuur van 'n woordeboek (vgl. Gouws en Prinsloo 2005: 134-138).

3.3 Adresseringsekwivalensie

In die artikel in (3) tree die lemmaties verteenwoordigde leksikale item (*sedert*) in elke brontaalkoteksinskrywing op, en elke vertaalekwivalent tree in minstens een doeltaalvoorbeeld op, terwyl geen doeltaalvoorbeeld 'n vertaalekwivalent bevat wat nie in die vertaalekwivalentparadigma voorkom nie. Tussen die vertaalekwivalentparadigma en die koteksafdeling in hierdie artikel geld *adresseringsekwivalensie* (vgl. Gouws 2000: 43). Die vereistes vir adresseringsekwivalensie word soos volg deur Gouws (2000: 43) gestel:

- (4) Adresseringsekwivalensie impliseer nie noodwendig dat daar altyd 'n een-tot-een-koördinasie tussen die koteksitems [d.i. doeltaalvoorbeelde — HLB] en die vertaalekwivalente moet wees nie. Een vertaalekwivalent kan meer as een ondersteunende koteksinskrywing hê, terwyl ander slegs een kan hê. Maar in 'n woordeboekartikel wat 'n ekwivalentverhouding van divergensie, en veral polidivergensie, vertoon, mag geen vertaalekwivalent geïsoleer word deur nie ook in die koteksafdeling 'n optrede te hê nie. Geen vertaalekwivalent mag in die koteksafdeling aangebied word as dit nie reeds in die vertaalekwivalentparadigma verstrekk

is nie. Elke koteksitem moet 'n vertaalekwivalent wat in die vertaalekwivalentparadigma opgeneem is, as adres hê. Elke vertaalekwivalent moet die adres van minstens een koteksitem wees.

Indien alle vertaalekwivalente in die vertaalekwivalentparadigma nie in die koteksafdeling in doeltaalkoteksinskrywings optree nie, geld *onderadressering* tussen die vertaalekwivalentparadigma en die koteksafdeling; indien doeltaalkoteksinskrywings vertaalekwivalente verstrek wat nie in die vertaalekwivalentparadigma verskyn nie, bevat die koteksafdeling 'n optrede van *ooradressering* (Gouws 2000: 43).

Om die beginsel van adresseringsekwivalensie te toets is die eerste 2 000 woordeboekartikels in die artikeltrajek van S in PHAROS (artikeldeeltrajek **s,S – selfbegrip**) nagegaan. Daar is bevind dat die artikel in (3) die enigste artikel van 'n polisemiese lemma is waarin adresseringsekwivalensie geld. In al die ander artikels van polisemiese lemmata in hierdie artikeldeeltrajek waarin voorbeelde wel voorkom, is adresseringsekwivalensie afwesig. Dit kan suggereer dat adresseringsekwivalensie relatief skaars in PHAROS voorkom. Vergelyk die volgende verteenwoordigende artikel uit dié woordeboek:

- (5) **greep** *grepe* clutch, grasp, grip, (hand)hold; coup; hilt; handle; pull (*of a door*); stock (*of a pistol, rifle, etc.*); cleat; (*comp.*) byte; →GREPIE; *iem. in jou ~ hê* have a stranglehold on s.o.; have a hold on/over s.o.; *in die ~ van ... wees* be in the clutches of ...; *grepe uit ...* snatches from ... (*history, s.o.'s life*); *'n ~ uit die lewe* a slice of life.

In die artikel in (5) geld 'n ekwivalentverhouding van polidivergensie tussen die lemma en die vertaalekwivalentparadigma (vgl. Gouws en Prinsloo 2005: 160-161), met agt polisemiese waardes wat daarin verteenwoordig word, een waarvoor vier vertaalekwivalente binne 'n doeltaalsinoniemparadigma aangebied word. Die koteksafdeling bestaan uit nege koteksinskrywings, waarvan vier brontaalkoteksinskrywings is. Slegs twee van die elf vertaalekwivalente in die vertaalekwivalentparadigma is die adresse van doeltaalkoteksinskrywings, naamlik *clutch* en *(hand)hold*.¹ In hierdie artikel ontbreek adresseringsekwivalensie duidelik, met die meeste vertaalekwivalente wat nie die adres van doeltaalkoteksinskrywings is nie (onderadressering) en doeltaalkoteksinskrywings wat vertaalekwivalente bevat wat nie in die vertaalekwivalentparadigma voorkom nie (ooradressering). Adresseringsekwivalensie (d.w.s. die neutralisering van onder- en ooradressering) kan in hierdie geval bereik word deur die vertaalekwivalente *stranglehold*, *snatch* en *slice* in die gepaste doeltaalsinoniemparadigmas binne die vertaalekwivalentparadigma te plaas en deur die koteksafdeling uit te brei sodat alle vertaalekwivalente in die vertaalekwivalentparadigma in die aangebode doeltaalkoteksinskrywings verteenwoordig word.

Uit die studie van ander polifunksionele woordeboeke blyk egter ook dat adresseringsekwivalensie nie so konsekwent voorkom as wat die beginsel in (4) sou wou hê nie. Vergelyk die volgende verteenwoordigende artikels — (6)(a)

uit COLLINS en (6)(b) uit HACHETTE:

- (6)(a) **Fluss** ▲ [flʊs] m -es, -e [ˈflʏsə] [a] (= *Gewässer*) river; **am** ~ by the river; **Stadt** on the river; **unten am** ~ down by the river(side); **den** ~ **aufwärts/abwärts fahren** to go upstream or upriver/downstream or downriver
 [b] no pl (Tech: = *Schmelzfluss*) molten mass; **im** ~ **sein** to be molten
 [c] (= *kontinuierlicher Verlauf*: von *Verkehr*, *Rede*, *Strom*, *Elektronen*) flow; (von *Verhandlungen* auch) continuity; **etw in** ~ (acc) **bringen** to get sth moving or going; **etw kommt** or **gerät in** ~ sth gets underway or going; (= *sich verändern*) sth moves into a state of flux; **im** ~ **sein** (= *sich verändern*) to be in a state of flux; (= *im Gange sein*) to be in progress, to be going on
- (b) **contrainte** nf [ɔ̃] (pression) pressure; (coercition) coercion; **user de** ~e à l'égard de qn to exert ou put pressure on sb; **par la** ~e by force ou coercion; **être empêché par la** ~e to be forcibly prevented; **sous la** ~e under duress; **tenir un peuple dans la** ~e liter to keep a nation in bondage; [2] (exigence) constraint; ~es **administratives/budgétaires** administrative/budgetary constraints; **les** ~es du **marché** the market restrictions; [3] (gêne) strain; **sans** ~e without restraint, freely; [4] Jur duress; ~e **par saisie de biens** distress, distraint; ~e **par corps** imprisonment for debt

'n Analise van die onderskeie koteksafdelings in (6) wys op die skynbare afwesigheid van adresseringsekwivalensie. Aangesien die artikels geïntegreerde mikrostrukture vertoon (vgl. Gouws 2003), is dit geredelik moontlik om die mate waarin adresseringsekwivalensie ten opsigte van elke polisemiese waarde (semantiese subkommentaar) geld, uit te druk. Die twee artikels in (6) bevat kollektief sewe semantiese subkommentare. In ses van dié subkommentare geld ooraadressing en in vier geld onderadressing. Slegs in die eerste semantiese subkommentaar van die artikel in 6(a) blyk adresseringsekwivalensie te geld.

Die probleem van onderadressing is relatief eenvoudig om op te los: Voeg tot elke semantiese subkommentaar koteksinskrywings waarin die niegeadresseerde vertaalekwivalente wel verstrekkend word. Dit is die verskynsel van ooraadressing wat 'n kompleksere gegewe blyk te verteenwoordig en wat in die res van hierdie artikel aandag geniet.

3.4 Ooraadressing

Ten einde die verskynsel van ooraadressing te verklaar, is die eerste 100 artikels in die artikeltrajek van *S* in PHAROS waarin ooraadressing voorkom, bestudeer. Hierdie 100 artikels word deur die artikeldeeltrajek *s,S-sinloos* geënkapsuleer, en bevat 'n totaal van 139 gevalle van ooraadressing. Uit die analise van hierdie 139 doeltaalkoteksinskrywings binne hulle verband met die korresponderende brontaalkoteksinskrywings is drie faktore geïdentifiseer wat adresseringsekwivalensie blyk teen te werk. Hierdie faktore word vervolgens uiteengesit.

3.4.1 Oorsake van ooraadressing

3.4.1.1 Ontoereikende bewerking

Die eerste faktor wat ooraadressing tot gevolg het, is gewoon ontoereikende bewerking. Vergelyk weer die artikel in (5) en die meegaande kommentaar.

Van die 139 gevalle van ooraadressering kan egter slegs 29 (20,86%) aan ontoereikende bewerking toegeskryf word.

3.4.1.2 Strukturele anamorfisme tussen bron- en doeltaal

Die tweede faktor wat skynbaar ooraadressering veroorsaak, is *die graad van strukturele anamorfisme tussen die bron- en doeltaal*. Volgens Zgusta (1971: 294) is die anamorfisme van tale die fundamentele probleem waarmee die vertalende leksikograaf te kampe het in die koördinerings van die brontaalitems met die doeltaalitems in sy/haar woordeboek. Die anamorfisme van tale kan op verskillende wyses in die taalgegewe gemanifesteer word, soos in die bestaan van kultuurgebonde leksikale items in 'n bepaalde brontaal wat in 'n doeltaal nie bestaan nie (referensiële leksikale gapings). Behalwe op die leksikale vlak, geld hierdie anamorfisme ook op die strukturele vlak, dit wil sê in terme van fonologie, morfologie en sintaksis. Waar twee tale genealogies nou aan mekaar verwant is, geld op strukturele vlak 'n laer graad van anamorfisme as wanneer die tale genealogies verder verwyderd van mekaar is. Dit volg dus dat Afrikaans en Engels, albei Wes-Germaanse tale, 'n laer graad van strukturele anamorfisme sal vertoon as byvoorbeeld Afrikaans en die "Afrika"-taal Oshindonga. Die Afrikaans-Engels-leksikograaf word derhalwe in hierdie verband met minder probleme ten opsigte van koteksinskrywings gekonfronteer as die Afrikaans-Oshindonga-leksikograaf, hoewel eersgenoemde leksikograaf nie belangrike uitdagings vryspring nie. Vergelyk die volgende artikel uit die data:

- (7) **selfvertroue** self-confidence, -reliance, -possession, -assurance, morale; *gebrek aan* ~ diffidence; *iets vol* ~ *doen* do s.t. self-confidently (or with assurance).

In die artikel in (7) is adresseringsekwivalensie volledig afwesig. Die koteksinskrywings in die artikel illustreer twee vlakke waarop strukturele anamorfisme van die brontaal en doeltaal ooraadressering veroorsaak. Ten opsigte van die brontaalkoteksinskrywing *gebrek aan selfvertroue* en die korresponderende doeltaalkoteksinskrywing *diffidence* is strukturele ekwivalensie afwesig, wat daarin manifesteer dat die lemmaties verteenwoordigde leksikale item *selfvertroue* in die brontaalvoorbeeld deel is van 'n konstruksie (nl. die setselgroep *aan selfvertroue* wat met die komplementerende substantief *gebrek* verbind)² wat in die geheel deur 'n enkele leksikale doeltaalitem (nl. *diffidence*) vertaal kan word. Dit volg dus dat die leksikale item *diffidence* nie as vertaalekwivalent van *selfvertroue* in die vertaalekwivalentparadigma as adres van die doeltaalkoteksinskrywing *diffidence* kan geld nie as gevolg van die afwesigheid van semanties-pragmatiese ekwivalensie. (Trouens, tegnies kan die inskrywing *diffidence* kwalik as 'n koteksinskrywing beskou word — dit is 'n geïsoleerde leksikale vertaalekwivalent sonder sintaktiese konteks, hoewel dit nie in die teenoorstaande deel van die woordeboek gelemmatiseer is nie.) Die probleem van ooraadressering sou in hierdie geval hanteer kon word indien die sinonieme frase

lack of self-confidence, wat strukturele ekwivalensie met die brontaalkoteksinskyrywing vertoon, naas die doeltaalitem *diffidence* geplaas kon word: *diffidence*, *lack of self-confidence*. Die plasing van dié konstruksie ná *diffidence* impliseer dan dat *diffidence* die meer gebruiklike vertaling verteenwoordig (indien dit wel die taalwerklikheid weerspieël). Hoewel hierdie prosedure die verskynsel van ooradressering sou hanteer, sou dit ooradressering nie neutraliseer nie, aangesien die doeltaalkoteksinskyrywing *diffidence* steeds in die koteksafdeling voorkom en geen direkte adres in die vertaalekwivalentparadigma het nie. Buitendien is so 'n prosedure nie in alle gevalle moontlik nie. Vergelyk die volgende verteenwoordigende artikel uit PHAROS:

- (8) **opgewonde** [...] excited, thrilled, enthusiastic; high (*infml.*); worked up, flutter; stimulated, aroused; animated, flurried, tumultuous, wild; *iets maak iem.* ~ s.t. excites s.o.; *oor iets ~ wees* be excited about s.t.; be wrought up over s.t.; *oor iets ~ raak/word* become/get excited about s.t.; get all worked up about s.t.

In die artikel in (8) geld ooradressering, met slegs die vertaalekwivalente *excited* en *worked up* wat in die koteksafdeling verskyn. Die doeltaalkoteksinskyrywing *something excites someone* verteenwoordig een geval van ooradressering, aangesien die vertaalekwivalent *excite* nie in die doeltaalkoteksinskyrywing figureer nie as gevolg van die anamorfisme tussen die bron- en doeltaalkotekste. Hier kan ooradressering egter nie geneutraliseer word deur die invoeging van 'n sinonieme doeltaalkoteksinskyrywing soos *something makes someone excited* nie, aangesien dit nie idiomatiese taalgebruik in die doeltaal verteenwoordig nie en derhalwe 'n disfunksionele doeltaalkoteksinskyrywing sou wees: Die geldende doeltaalkoteksinskyrywing het juis ten doel om die gebruiker te verhoed om so 'n onidiomatiese doeltaalkonstruksie te produseer en fasiliteer sodoende kommunikatiewe ekwivalensie (vgl. Gouws 1996). Die doeltaalkoteksinskyrywing *something excites someone* verteenwoordig in hierdie geval dus funksionele kontekstualisering (vgl. Beyer 2009) ten spyte van die verskynsel van ooradressering. Dit wil voorkom of hierdie stand van sake die eis van adresseringsekwivalensie, soos in (4) geformuleer, bevraagteken. Daar word aanstons op hierdie punt teruggekeer.

Die tweede vlak van strukturele anamorfisme tussen brontaal- en doeltaalkoteks word deur die brontaalkoteksinskyrywing *iets vol selfvertroue doen* en die korresponderende doeltaalkoteksinskyrywing *do something self-confidently* in die artikel in (7) geïllustreer. Die anamorfisme word hier op die spits gedryf deur die afwesigheid van woordsoortelike ekwivalensie tussen die brontaalkotekselement *selfvertroue* ('n substantief) en die doeltaalkotekselement *self-confidently* ('n bywoord). (Trouens, dit geld ook die vertaalekwivalent *excited* (adj.) en die doeltaalkotekselement *excites* (ww.) in die artikel in (7).) Dit beteken dat terwyl die brontaalkoteksinskyrywing aan die substantiwiese lemma **selfvertroue** geadresseer is, die doeltaalkoteksinskyrywing nié aan die substantiwiese vertaalekwivalent *self-confidence* in die vertaalekwivalentparadigma geadres-

seer is nie. Hier kan dus nie sprake wees van die funksionele kontekstualisering van die betrokke vertaalekwivalent nie, hoewel die vormlike ooreenkoms groot is — en misleidend kan wees. Vergelyk in hierdie verband ook die volgende artikel uit die data:

- (9) **siedend** =*dende seething*; ~ *kwaad*, ~ *van woede seething/fuming/boiling with rage*.

Die lemmaties verteenwoordigde adjektief *siedend* in (9) word van die adjektiewiese vertaalekwivalent *seething* voorsien. In die brontaalkoteksinskrywings *siedend kwaad* en *siedend van woede* tree die lemmaties verteenwoordigde adjektief op; dié inskrywings is aan die lemma geadresseer. Die doeltaalkoteksinskrywing *seething/fuming/boiling with rage* is egter nié aan die vertaalekwivalent *seething* in die vertaalekwivalentparadigma geadresseer nie, aangesien die leksikale item *seething* in die doeltaalkoteksinskrywing 'n werkwoord is en nie 'n adjektief soos die vertaalekwivalent nie. Die doeltaalkoteksinskrywing illustreer dus nié die optrede van die aangebode vertaalekwivalent nie. Aangesien ortografiese identiteit tussen die adjektiewiese vertaalekwivalent en die werkwoordelike leksikale item in die doeltaalkoteksinskrywing bestaan, kan 'n sintagmatiese verband gesuggereer word wat in der waarheid nie bestaan nie. Hier is sprake van die disfunksionele effek van *skynadressering*, wat vermy kan word deur bepaalde aanpassings aan die mikro- en toegangstruktuur.

Die strukturele anamorfisme tussen bron- en doeltaal verklaar 'n beduidende 53 (38,13%) van die 139 gevalle van ooradressering.

3.4.1.3 Konteksgeaktiveerde semanties-pragmatiese ekwivalensie

Die derde faktor wat ooradressering veroorsaak, is *semanties-pragmatiese ekwivalensie binne 'n bepaalde konteks wat nie sistematies voorspelbaar is nie*. Volgens Szende (1999: 201) kan leksikale items binne konteks bepaalde betekeniswaardes aanneem wat nie voorspel kan word deur bloot hulle denotatiewe betekenis buite konteks te bestudeer nie. Vergelyk die volgende verteenwoordigende artikel uit die data, waarin ooradressering geld:

- (10) **seeploos** =*lose soapless*; ~*lose wasmiddel synthetic detergent*.

Buite konteks sou die brontaalitem *seeploos* nie van die vertaalekwivalent *synthetic* voorsien word nie, aangesien daar nie sistematies voorspelbare semanties-pragmatiese ekwivalensie tussen dié items bestaan nie. Die spesifieke linguïstiese konteks waarin *seeploos* en *wasmiddel* kombineer om die frase *seeplose wasmiddel* te produseer, aktiveer andersins nie-voorspelbare semanties-pragmatiese ekwivalensie tussen *seeploos* en *synthetic* en vereis die doeltaalkoteksinskrywing *synthetic detergent* wat oënskynlik nie aan die vertaalekwivalent *soapless* in (10) geadresseer is nie. Ook hier geld funksionele kontekstualisering ter wille van kommunikatiewe ekwivalensie ten spyte van ooradressering.

'n Spesifieke pragmatiese konteks kan ook semanties-pragmatiese ekwivalensie tussen bron- en doeltaalitems aktiveer wat buite dié konteks geensins sou geld nie, soos wat die volgende verteenwoordigende artikel uit die data aandui:

- (11) **saaklik** =like businesslike, to the point, efficient, thorough; objective, impersonal; concise, succinct; essential, real; matter-of-fact; factual; pertinent, relevant; (~e) *onderpand*, (*jur.*) collateral security; ~e *reg.* (*jur.*) real right; ~e *servituut*, (*jur.*) predial servitude.

Uit die artikel in (11) word dit duidelik dat die vertaalekwivalent *collateral* slegs binne juridiese konteks voorspelbaar met die brontaalitem *saaklik(e)* gekoördineer kan word, en dat semanties-pragmatiese ekwivalensie tussen hierdie items andersins nie sistematies voorspelbaar is nie. Dieselfde geld die verhouding tussen die brontaalitem *saaklik* en die vertaalekwivalent *predial* in die koteksinskrywingspaar *saaklike servituut / predial servitude*. Ook hier geld funksionele kontekstualisering ter wille van kommunikatiewe ekwivalensie ten spyte van ooradressering.

Konteksgeaktiveerde semanties-pragmatiese ekwivalensie tussen brontaal- en doeltaalitems verklaar die oorblywende 58 (41,73%) van die 139 gevalle van ooradressering in die data.

3.4.1.4 'n Kombinasie van bogenoemde faktore

Die onderskeid tussen strukturele anamorfisme van die behandelde tale en konteksgeaktiveerde semanties-pragmatiese ekwivalensie as faktore wat ooradressering veroorsaak, is nie altyd so diskreet soos wat die artikels in (10) en (11) kan suggereer nie. Vergelyk die volgende artikel uit die data:

- (12) **seevaart** navigation; seafaring; voyage; *geskiedenis van die* ~ naval history.

In (12) kom 'n kombinasie van die twee faktore voor: Die doeltaalkotekselement *naval* is nie sintakties ekwivalent aan die lemmaties verteenwoordigde brontaalitem nie én in kombinasie met die element *history* geld die vertaalekwivalente *navigation* (*navigation history), *seafaring* (*seafaring history) en *voyage* (*voyage history) nie.

3.4.1.5 Ander faktore

Die faktore wat pas beskryf is, verteenwoordig nie noodwendig die enigste faktore wat tot ooradressering aanleiding gee nie. In die studie van die data is slegs op aspekte van strukturele ooreenkomste en verskille tussen die behandelde tale gefokus. Uiteraard word ander moontlike faktore, byvoorbeeld die aanduiding van semantiese seleksiebeperinge in die doeltaal as 'n moontlike funksie van 'n doeltaalkoteksinskrywing, nie negeer nie.

3.5 Adresseringsekwivalensie vs. funksie

Volgens Svensén (1993: 106) geld die anamorfisme tussen die behandelde tale en konteks-geaktiveerde semanties-pragmatiese ekwivalensie inderdaad as primêre motivering vir die gebruik van voorbeelde in vertalende woordeboeke:

Translated examples in bilingual dictionaries are usually given whenever a source-language phrase involving the headword cannot be translated word for word into the target language by means of equivalents of the headword in isolation.

Hierdie standpunt van Svensén word deur Kromann et al. (1991a: 2772) ondersteun:

It is obvious that constructions whose translation follows unproblematically from elementary grammatical (and semantic) rules should not be included in the dictionary.

Opvallend van die standpunte van Svensén en Kromann et al. is dat die beginsel van adresseringsekwivalensie implisiet nié daardeur ondersteun word nie. Beide die standpunte sowel as die beginsel van adresseringsekwivalensie behoort egter ondergeskik gestel te word aan die funksie en teikengebruiker van die betrokke woordeboek. Svensén en Kromann et al. se standpunte sou byvoorbeeld vir monofunksionele teksresepsie-woordeboeke kon geld, terwyl die beginsel van adresseringsekwivalensie as 'n vereiste kenmerk van 'n polifunksionele woordeboek beskou sou kon word.

3.6 Adresseringsekwivalensie hersien

Indien aanvaar word dat die faktore wat ooradressering veroorsaak, dit doen juis ter wille van funksionele kontekstualisering in belang van kommunikatiewe ekwivalensie, moet die vraag gevra word of adresseringsekwivalensie inderdaad afwesig is. In die artikel in (13) word die gebruiker deur die doeltaalkoteksinskrywing *naval history* daarop gewys dat die vertaalekwivalent *navigation* nié die lemma *seevaart* in die doeltaal kan vervang indien dit in die brontaalkonstruksie *geskiedenis van die seevaart* voorkom nie. Die funksie van sodanige doeltaalkoteksinskrywing kan *kontekstuele vertaalekwivalentuitsluiting* genoem word.³ Daar kan dus geargumenteer word dat die doeltaalkoteksinskrywing *naval history* in (13) indirek aan die vertaalekwivalent *navigation* geadresseer is. Dieselfde argument geld mutatis mutandis die woordeboekartikels in (8) tot (12). Indien 'n vertaalekwivalentparadigma meer as een vertaalekwivalent bevat, sou van *kollektiewe indirekte adressering* sprake kon wees, met al die lede van die vertaalekwivalentparadigma wat as kollektiewe indirekte adres van die betrokke doeltaalkoteksinskrywing optree. 'n Meer toepaslike mikrostruktuurtype, byvoorbeeld 'n geïntegreerde mikrostruktuur (soos wat geld in

die artikels in (6)), sal die verband tussen die betrokke doeltaalkoteksinskrywing en die nie-buikbare vertaalekwivalent(e) effektiewer lê as wat tans in PHAROS se gemengde primitiewe mikrostruktuur (vgl. Gouws 2003) die geval is, dit wil sê die *indirekte adres* van sodanige doeltaalkoteksinskrywing sal makliker deur die gebruiker geïdentifiseer kan word. Indien hierdie doeltaalkoteksinskrywings dan tesame met doeltaalkoteksinskrywings optree wat al die onderskeie vertaalekwivalente as *direkte adresse* het (d.w.s. wat die vertaalekwivalente bevat), geld adresseringsekwivalensie in die betrokke woordeboekartikels (gegeë dat onderadressering afwesig is).

Hiervolgens kan doeltaalvoorbeelde in twee klasse verdeel word:

- (13) (a) doeltaalvoorbeelde wat die *vervangbaarheid* van die lemma deur die *direk* geadresseerde vertaalekwivalent(e) in die relevante linguistiese en pragmatiese konteks *illustreer*; en
- (b) doeltaalvoorbeelde wat die *nie-vervangbaarheid* van die lemma deur die *indirek* geadresseerde vertaalekwivalent(e) in die relevante linguistiese en pragmatiese konteks *spesifiseer*.

Voorbeelde wat tot klas (13)(a) behoort, kan *kotekstuele illustrators* genoem word, terwyl na voorbeelde wat tot klas (13)(b) behoort as *kotekstuele spesifiseerders* verwys kan word. Gegee 'n toepaslike mikrostruktuurtype, sou hierdie twee klasse voorbeelde ook in aparte en duidelik deur struktuurmerker afgebakende soeksones geplaas kon word. Dit is by uitstek (maar nie uitsluitlik nie) die klas kotekstuele illustrators wat die funksies van voorbeelde soos vroeër aangetoon moet vervul en die verbandhoudende eienskappe moet vertoon. Hieruit volg dat kotekstuele illustrators dié kategorie koteksinskrywings verteenwoordig wat normaalweg as *voorbeelde* in die literatuur beskou word; daarom kan kotekstuele illustrators ook *ware voorbeelde* genoem word. Daarenteen illustreer kotekstuele spesifiseerders nie die optrede van vertaalekwivalente nie, maar spesifiseer eerder die nie-optrede daarvan in spesifieke konteks. As sodanig geld kotekstuele spesifiseerders ook as kontekstualiserende inskrywings (vgl. Beyer 2009). Wat hier van belang is, is dat kotekstuele spesifiseerders se optrede onafhanklik van die geldende ekwivalentverhouding tussen die lemma en vertaalekwivalentparadigma in die betrokke artikel fungeer, terwyl kotekstuele illustrators (ware voorbeelde) gewoonlik afhanklik daarvan fungeer. Vergelyk die volgende artikels uit PHAROS:

- (14) (a) **hoorbaar** =bare audible; *bo iets ~ wees* be audible above s.t.; *jou ~ maak* make o.s. heard; *hoorbare fluistering* stage whisper.
- (b) **konsertant** =tante concertante; *~e simfonie* sinfonia concertante.
- (c) **fisika** physics; *~ van die vaste toestand* solid state physics.

In die artikels in (14)(a) en (b) geld 'n ekwivalentverhouding van kongruensie (soos bepaal deur toetsing aan die omkeerbaarheidsbeginsel). Die veronderstelling is

normaalweg dat konteks geen rol speel in die aanbod van die vertaalekwivalent nie (vgl. Gouws 1989: 165). Tog tree koteksinskrywings in hierdie artikels op.

In (14)(a) verteenwoordig die eerste doeltaalkoteksinskrywing (*be audible above s.o.*) 'n kotekstuele illustrator (ware voorbeeld). Een rede vir die insluiting van hierdie ware voorbeeld is die afwesigheid van strukturele ekwivalensie tussen die bron- en doeltaalkoteksinskrywing: In die brontaal kan die lemma *hoorbaar* volgens die koteksinskrywing (ook) na die setsel *bo* optree; in die doeltaal moet die vertaalekwivalent *audible* voor die setsel *above* optree. Hierdie soort data is belangrik vir die gebruiker in die teksproduksiesituasie. In (14)(b) en (c) geld dieselfde soort omstandighede. (In (14)(c) geld 'n ekwivalentverhouding van absolute ekwivalensie, maar nie kongruensie nie.)

Die tweede en derde doeltaalkoteksinskrywing in (14)(a) (*make oneself heard* en *stage whisper*) verteenwoordig kotekstuele spesifiseerders wat die vertaalekwivalent *audible* kontekstueel uitsluit, ten spyte van die ekwivalentverhouding van kongruensie wat in die artikel geld. Kotekstuele spesifiseerders verteenwoordig dié doeltaalkoteksinskrywingstipe wat die meeste in artikels optree waarin 'n ekwivalentverhouding van kongruensie geld. Uiteraard tree hierdie tipe koteksinskrywing in artikels op ongeag die geldende ekwivalentverhouding.

Uit die voorgaande bespreking blyk dit dat 'n ekwivalentverhouding van kongruensie in 'n artikel die leksikograaf nie vrystel van die aanbod van koteksinskrywings (as 'n tipe kontekstualiserende inskrywing) nie, veral nie indien die anamorfisme tussen die twee behandelde tale 'n prosedure van kontekstuele vertaalekwivalentuitsluiting vereis nie. Die afwesigheid van kotekstuele spesifiseerders sou tot die disfunksionele effek van foutiewe teksproduksie of vertaling in die doeltaal deur die woordeboekgebruiker kon lei.

4. Gevolgtrekking

Die beginsel van adresseringsekwivalensie, as 'n funksie-afhanklike leksikografiese konvensie, mag nie die ewe funksie-afhanklike prosedure van kontekstuele vertaalekwivalentuitsluiting deur die aanwending van kontekstuele spesifiseerders verhinder nie.

Elke kotekstuele spesifiseerder het sowel 'n direkte adres (die korresponderende brontaalvoorbeeld) as 'n indirekte adres (die kontekstueel uitgesluite vertaalekwivalent(e)). Hierdie adresseringstruktuur suggereer die bestaan van 'n *adresseringshiërargie* wat potensieel vir alle tipes inskrywings wat aan meer as een adres gerig is, kan geld. Sodanige hiërargie bestaan daarin dat inskrywings wat aan meer as een adres gerig is, minstens 'n *primêre* adres moet hê wat ook 'n direkte adres moet wees, sowel as 'n *sekondêre* adres wat ook 'n indirekte adres sou kon wees.

Notas

1. Oor die funksionaliteit van hierdie koteksinskrywing word nie kommentaar gelewer nie.

2. Die feit dat die frase *gebrek aan* 'n kollokasie verteenwoordig, word nie negeer nie, maar dit speel hier nie 'n deurslaggewende rol nie.
3. Hierdie funksie word juis *kontekstuele* vertaalekwivalentuitsluiting genoem omdat die uitsluiting nie tot koteks beperk is nie, maar ook op die breër konteks betrekking kan hê, soos wat die artikel in (11) illustreer. Daar is in terme van funksie nie 'n waterdigte onderskeid tussen kontekstualiserende en koteksinskrywings te tref nie; laasgenoemde is eerder 'n subklas van eersgenoemde (vgl. Beyer 2009), aangesien (a) koteksinskrywings ook pragmatiese data kan bevat (vgl. 3.1), en (b) ander tipes kontekstualiserende inskrywings, soos diskriminators, ook koteks in verskillende grade kan weerspieël. Vergelyk ter illustrasie van laasgenoemde punt relevante uittreksels uit COLLINS se beskrywing van die vorm van glosse op p. xv (my kursivering — HLB):
 - 2.1.3 within verb entries, *typical subjects of the headword* [...]
 - 2.1.4 within noun entries, *typical noun complements of the headword* [...]
 - 2.2.1 in transitive verb entries, *typical objects of the headword* [...]
 - 2.2.2 in adjective entries, *typical nouns modified by the headword* [...]
 - 2.2.3 in adverb entries, *typical verbs or adjectives modified by the headword*

Bronnelys

Primêre literatuur (woordeboeke)

- COLLINS = **Terrell, P., V. Schnorr, W.V.A. Morris en R. Breitsprecher (Reds.)**. 1999⁴. *Collins German-English/English-German Dictionary*. Glasgow: HarperCollins.
- HACHETTE = **Ormal-Grenon, J. en N. Pomier**. 2001³. *The Oxford-Hachette French Dictionary: French-English/English-French*. Oxford: Oxford University Press.
- PHAROS = **Du Plessis, M. (Red.)**. 2005. *Pharos Afrikaans-Engels Engels-Afrikaans Woordeboek*. Kaapstad: Pharos.

Sekondêre literatuur

- Al-Kasimi, A.M.** 1977. *Linguistics and Bilingual Dictionaries*. Leiden: E.J. Brill.
- Beyer, H.L.** 2009. 'n Teoretiese basis vir kontekstualisering in tweetalige woordeboeke. *Lexikos* 19: 1-22.
- Cop, M.** 1990. The Function of Collocations in Dictionaries. Magay, T. en J. Zigány (Reds.). 1990. *BudaLEX '88 Proceedings. Papers from the 3rd International EURALEX Congress, Budapest, 4-9 September 1988*: 35-46. Budapest: Akadémiai Kiadó.
- Cop, M.** 1991. Collocations in the Bilingual Dictionary. Hausmann, F.J. et al. (Reds.) 1989-1991: 2775-2778.
- Gouws, R.H.** 1989. *Leksikografie*. Pretoria/Kaapstad: Academica.
- Gouws, R.H.** 1996. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 14-31.
- Gouws, R.H.** 2000. Doeltaalgerigtheid teenoor lemmagerigtheid in vertalende woordeboeke. *Tydskrif vir Geesteswetenskappe* 40(1): 39-47.
- Gouws, R.H.** 2003. Aspekte van mikrostrukturele verskeidenheid en inkonsekwentheid in woordeboeke. *Lexikos* 13: 92-110.

- Gouws, R.H. en D.J. Prinsloo.** 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PReSS.
- Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta. (Reds.)** 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*. Handbücher zur Sprach- und Kommunikationswissenschaft 5.1–5.3. Berlin/New York: Walter de Gruyter.
- Jackson, H.** 1988. *Words and their Meaning*. London/New York: Longman.
- Jacobsen, J.R., J. Manley en V.H. Pedersen.** 1991. Examples in the Bilingual Dictionary. Hausmann, F.J. et al. (Reds.) 1989–1991: 2782-2789.
- Kromann, H., T. Riiber en P. Rosbach.** 1991. Principles of Bilingual Lexicography. Hausmann, F.J. et al. (Reds.) 1989–1991: 2711-2728.
- Kromann, H., T. Riiber en P. Rosbach.** 1991a. Grammatical Constructions in the Bilingual Dictionary. Hausmann, F.J. et al. (Reds.) 1989–1991: 2770-2775.
- Landau, S.I.** 2001. *Dictionaries. The Art and Craft of Lexicography*. Tweede uitgawe. Cambridge: Cambridge University Press.
- Lombard, F.J.** 1992. Voorbeeldmateriaal in woordeboeke. *Lexikos* 2: 148-164.
- Manley, J., J. Jacobsen en V.H. Pedersen.** 1988. Telling Lies Efficiently: Terminology and the Microstructure in the Bilingual Dictionary. Hyldgaard-Jensen, K. en A. Zettersten (Reds.) 1988. *Symposium on Lexicography III. Proceedings of the Third International Symposium on Lexicography, May 14–16, 1986, at the University of Copenhagen*: 281-302. Tübingen: Max Niemeyer.
- Nesi, H.** 1996. The Role of Illustrative Examples in Productive Dictionary Use. *Dictionaries* 17: 198-206.
- Prinsloo, D.J. en R.H. Gouws.** 2000. The Use of Examples in Polyfunctional Dictionaries. *Lexikos* 10: 138-156.
- Rademeyer, L.** 1992. *Die funksies van verbale voorbeeldmateriaal in eentalige woordeboeke*. Ongepubliseerde M.A.-thesis. Stellenbosch: Universiteit van Stellenbosch.
- Svensén, B.** 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Oxford: Oxford University Press.
- Szende, T.** 1999. Problems of Exemplification in Bilingual Dictionaries. *Lexicographica. International Annual for Lexicography* 15: 198-228.
- Van Niekerk, A.E.** 1992. Kollokasies: 'n Leksikografiese perspektief. *Lexikos* 2: 254-264.
- Wiegand, H.E.** 1988. "Shanghai bei Nacht." Auszüge aus einem metalexikographischen Tagebuch zur Arbeit beim *Großen Deutsch-Chinesischen Wörterbuch*. Wiegand, H.E. (Red.). 1988. *Studien zur neuhochdeutschen Lexikographie* 6(2): 521-626. Germanistische Linguistik. Hildesheim: Georg Olms.
- Wiegand, H.E.** 1996. A Theory of Lexicographic Texts: An Overview. *South African Journal of Linguistics* 14(4): 134-149.
- Wiegand, H.E.** 2000. Adressierung in der ein- und zweisprachigen Lexikographie. Eine einführende Übersicht über die Forschungs- und Problemlage. *Lexikos* 10: 32-74.
- Zgusta, L.** 1971. *Manual of Lexicography*. Janua Linguarum. Series Maior 39. Den Haag: Mouton.

Equivalent Selection in Specialized e-Lexicography: A Case Study with Spanish Accounting Terms

Pedro A. Fuertes-Olivera, *University of Valladolid, Spain*
(*pedro@tita.emp.uva.es*)

Abstract: Interest in specialized lexicography has been propelled both by the development of LSP communication in academic circles and by the consolidation of function-based approaches to lexicography that have identified the existence of several user types, e.g., experts, semi-experts and interested laypersons, and use situations, typically cognitive-oriented and communicative-oriented (Bergenholtz and Tarp, 2003, 2004). This paper follows suit and elaborates on the selection of Spanish equivalents in a particular dictionary project: the *Diccionario Inglés-Español de Contabilidad*, one of the Accounting dictionaries. This dictionary aims to satisfy the needs of translators (primary user group), accountants and financial experts (secondary user group), as well as students of accountancy and translation, journalists, and interested laypersons (tertiary user group). It addresses the issue as a lexicographical problem and makes comments on the decisions taken by elaborating on three lexicographical principles that take into consideration the nature of lexicography, the technical options the Internet offers, and the defining characteristics of specialized discourse: *relevance*, *proscription* and *recreation*.

Keywords: E-LEXICOGRAPHY, LSP LEXICOGRAPHY, TERMINOLOGY, SPANISH, ENGLISH, EQUIVALENT SELECTION, ACCOUNTING DICTIONARIES, FUNCTION THEORY, LSP COMMUNICATION, TRANSLATION DICTIONARY

Opsomming: Ekwivalent-seleksie in gespesialiseerde e-leksikografie: 'n Gevallestudie met Spaanse rekeningkundige terme. Belangstelling in gespesialiseerde leksikografie is aangevuur deur sowel die ontwikkeling van TSD-kommunikasie in akademiese kringe as die konsolidasie van funksie-gebaseerde benaderings tot leksikografie. Laasgenoemde benadering het die bestaan van verskeie gebruikerstipes, bv. kundiges, semi-kundiges en geïnteresseerde leke, en gebruiksituasies, tipies kognitief-georiënteerde en kommunikatief-georiënteerde situasies, geïdentifiseer (Bergenholtz and Tarp, 2003, 2004). Hierdie artikel volg hierdie ontwikkelings na en wei uit oor die seleksie van Spaanse ekwivalente in 'n spesifieke woordeboekprojek: die *Diccionario Inglés-Español de Contabilidad*, 'n rekeningkunde-woordeboek. In hierdie woordeboek word beoog om te voorsien in die behoeftes van vertalers (die primêre gebruikersgroep), rekeningkundiges en finansiële kundiges (die sekondêre gebruikersgroep), en rekeningkunde-, vertaal- en joernalistiekstudente asook geïnteresseerde leke (die tersiêre gebruikersgroep). Hierdie kwessie word aangespreek as 'n leksikografiese probleem en kommentaar word gelever op die besluite wat geneem is deur uit te brei op drie leksikografiese beginsels wat die aard van die leksikografie, die tegniese opsies wat deur die Internet gebied word, en die definiërende eienskappe van gespesialiseerde diskoers in ag neem: *relevansie*, *proskripsie* en *herskepping*.

Sleutelwoorde: E-LEKSIKOGRAFIE, TSD-LEKSIKOGRAFIE, TERMINOLOGIE, SPAANS, ENGELS, EKWIVALENT-SELEKSIE, REKENINGKUNDE-WOORDEBOEKE, FUNKSIE-TEORIE, TSD-KOMMUNIKASIE, VERTAALWOORDEBOEK

1. Introduction: LSP Lexicography

Although dictionary-writing is an activity with a long history, the theoretical foundations on which such an activity are carried out were not much debated until the 1990s, specifically with reference to the presentation of lexicography as an independent science (see Gouws 2011 for a review of the three stages in lexicography he comments on). Similarly, interest in specialized lexicography was almost non-existent until the publication of the *Manual of Specialised Lexicography* (Bergenholtz and Tarp 1995), which introduced a lexicographical approach to the study of terminology, and the establishing of departments of LSP communication in academic institutions, demonstrating the enormous potential for this field of research. For instance, certain tenets of Wüster's General Theory of Terminology could not be substantiated when specialized texts and reference works were subjected to scrutiny, which emphasized the necessity for a better understanding of the lexical component present in LSP communication.

In particular, the lexicographical approach to terminology offered arguments for discarding the often-quoted distinction between specialized lexicography and terminology, which had contributed to the presenting of specialized reference works, usually referred to as glossaries, encyclopedias, knowledge bases, term bases, etc., as alien to lexicography (Hartmann 2009). Moreover, recent research (e.g., Bergenholtz and Nielsen 2006; Fuertes-Olivera and Arribas-Baño 2008; Fuertes-Olivera 2010; Fuertes-Olivera and Nielsen 2011) has assumed that specialized lexicography and terminology are two sides of the same coin, which lends support to the use of well-attested lexicographical principles when planning and compiling specialized reference works. Among these principles are the following:

- Specialized reference works must offer data for disambiguating the meaning and use of LSP terms in specific use situations, typically in cognitive-oriented and communicative-oriented. This means that the LSP dictionary should be all-inclusive, i.e., a type of dictionary that adds conceptual data to the linguistic data normally given.
- Specialized reference works can be descriptive or proscriptive. Bergenholtz and Gouws (2010), for example, advocate a proscriptive approach for text production and translation dictionaries, and a descriptive approach for text reception dictionaries.
- Specialized reference works typically address the information needs of experts, semi-experts and interested laypeople. In some use situations

(for example, with Spanish students of Business English), the above typology merits a more detailed classification, as shown in Fuertes-Olivera and Arribas-Baño (2008).

- The distinction between culture-dependent and culture-independent subject fields is becoming blurred due to the preponderance of English loan translations, borrowings and calques in the terminologies of most languages, a fact that can be explained due to economic globalization and the consolidation of English as the *lingua franca* of specialized communication.

In addition, current research into LSP lexicography is also paying attention to recent developments in user-based approaches to electronic lexicography. In particular, the future of LSP lexicography seems to be subject to the combined influence of technological innovations and theoretical considerations. For example, Bergenholtz, Nielsen and Tarp (2009: 9) claim that the appearance of electronic dictionaries has lent support to function-based lexicographic theories that are centered on the dictionaries and the users, with the aim of generating new dictionaries based on the functions they must fulfill. They add that these principles

[A]pply in particular to the technical options offered by databases, text corpora, Internet searches, and Internet-based dictionaries and encyclopedias, and also to theories of and the preparation of lexicographic tools that help users meet the needs that they may have in different situations, e.g. understanding texts, producing texts, translating texts or acquiring new knowledge.

The so-called Accounting dictionaries illustrate the workings of Bergenholtz et al's reflection on a specific lexicographical project, and signal the way ahead for LSP lexicography. The Accounting dictionaries are a set of specialized Internet dictionaries that aim to satisfy the needs of translators (primary user group), accountants and financial experts (secondary user group), as well as students of accountancy and translation, journalists, and interested laypersons (tertiary user group). Several recent papers have been published in order to describe the rationale for some of the lexicographical decisions taken during planning and compilation (Nielsen and Mourier 2007), to discuss technical aspects of the accounting database that allows users to retrieve on demand (Nielsen and Almind 2011), and to explain the dynamics of accounting terms in today's world (Fuertes-Olivera and Nielsen 2011).

In addition, more focused research has also been published in relation to each of the Accounting dictionaries. For example, the *Diccionario Inglés-Español de Contabilidad* (Nielsen et al. 2009) has merited the attention of Fuertes-Olivera and his colleagues regarding some of its current lexicographical characteristics and future possibilities: Fuertes-Olivera (2009) has elaborated on the construction of a systematic introduction that is currently underway. Fuertes-Olivera and Niño Amo (2011) have envisaged some lexicographical options for in-

creasing the reliability and systematicity of the dictionary. Finally, Fuertes-Olivera et al. (2010) have described how the concept of expert knowledge permeates the Spanish translation of the *English Accounting Dictionary* (Nielsen et al. 2010). Following suit, this paper elaborates on the application of a lexicographical approach to equivalent selection in specialized dictionaries, an issue that merits special attention considering that the process of equivalent selection is usually addressed from a linguistic and not a lexicographical perspective (Adamska-Salaciak 2010).

As has already been mentioned by Bergenholtz and colleagues, nothing is more practical than a good theory (Nielsen and Tarp 2009), and so the explanations discussed here are based on the tenets of the function theory of lexicography, as well as on claims concerned with the nature of LSP communication. By so doing we add weight to two principles that permeate this paper and much of the current debate on the nature of lexicography. Firstly, lexicography is an independent academic discipline that can benefit from the options introduced by the advent of the so-called information and knowledge society. This means that we have to offer a lexicographical explanation for the options chosen when planning and compiling (Internet) information tools.

Secondly, time is ripe for discussing whether lexicography is part of information science, a change of paradigm that "could spell the beginning of the end of the underestimation of lexicography" (Bergenholtz in press), and reinforce the focus on dictionaries as tools providing assistance in various situations rather than an exercise in linguistics; in addition, it could promote "the fundamental theories that will govern the implementation of a lexicographically designed needs-adapted information and data access in other types of information tools" (Leroyer 2011).

The above principles oppose the view of lexicography as a linguistic exercise, and therefore open the door for focusing on scientific endeavor to explain the process of decision-making in lexicography. In this paper, my interest lies in explaining the selection of Spanish equivalents in the *Diccionario Inglés-Español de Contabilidad*, one of the Accounting dictionaries that will briefly be presented in Section 3. Before this, however, Section 2 will review the concept of the specialized translation dictionary, as the *Diccionario Inglés-Español de Contabilidad* aims primarily to meet the needs of Spanish translators of English accounting texts. Section 4 constitutes the core of this paper, discussing the lexicographic rationale for taking real decisions on equivalent selection in specialized lexicography. This will help us to add support to our conviction regarding the necessity of upgrading theoretical discussions concerned with specialized lexicography. Finally, Section 5 will provide a summary of certain conclusions.

2. Specialized Translation Dictionaries

Specialized translation dictionaries are reference tools that aim to help users translate specialized texts. Dictionaries of this type are planned and compiled

to assist the user in the transfer of a message from one language into another within a particular subject field (Fata 2010; Nielsen 2010).

Specialized translation dictionaries are usually planned and compiled with the aim of helping users disambiguate the meaning of the headword term by including linguistic, conceptual and pragmatic data, e.g., definitions, equivalents, synonyms, antonyms, usage labels, domain tags and illustrations. In some lexicographic traditions, e.g., in Spain, these dictionaries are usually identified as dictionaries that offer source language lemmas on the left and their target language equivalents on the right.

However, the value and usefulness of the traditionally-conceived *bilingual specialized translation dictionary* became open to question when functional approaches to translation and lexicography started to gain momentum. Translation scholars such as Nord (1997) shifted the focus of attention from terms and words to larger units, making a case for defending the iterative nature of the translation process with several recursive processes that go "beyond the levels of terms and words" (Nielsen 2010: 71), and defending a target-language orientation, by emphasizing that the "target-language texts must belong to the same genre as the source text and the linguistic and textual material that is borrowed must have the same pragmatic functions as those in the source text." (Nielsen 2010: 71).

The above situation coincided with the publication of a new generation of LSP dictionaries, which were the result of a function-based lexicographical approach initiated at the Centre for Lexicography (Aarhus School of Business) (Nielsen 1994). The authors defended the argument that specialized translation dictionaries should be more than a repository of terms and words, and proposed taking into consideration factors that are relevant in the translation process, most of which had hitherto been paid no attention by specialized lexicographers: *text types* and *genres*, *use situation* (for example, a translator of specialized texts also needs conceptual data), and *culture*. The *Gene Technology Dictionary* (Kaufmann, Bergenholtz, et al.) is an example of the then new approach. In its Preface (p. 7) it informs users that this dictionary is a polyfunctional tool which assumes that the user "possesses some familiarity with scientific language usage both in English and Spanish and a minimum of knowledge of biology and chemistry." More importantly, this specialized dictionary is very innovative, as it simultaneously performs the functions of an equivalence, explanatory and defining dictionary. For example, in a translation situation users benefit from three lexicographical developments in this dictionary:

- it contains an encyclopedic section, usually called a *systematic introduction* (Bergenholtz and Nielsen 2006), which explains the basics of gene technology and hence can be used when a professional translator, who is not an expert or semi-expert in gene technology, needs help in order to understand the basics of the domain;

- the more central terms in the field are explained by means of translation equivalents and definitions, which help translators to place the term in its conceptual realm;
- it also contains some entries that "are included above all to help the user in a translation situation." (p. 8). For example, the inclusion of many multi-word expressions (e.g., **anion exchange column**) are relevant in a translation situation, considering that these expressions are difficult to translate and offer clues as to the degree of idiomaticity of the translated text.

In short, the functional approach to lexicography, as illustrated in the *Gene Technology Dictionary*, paved the way for planning and compiling specialized translation dictionaries that also included a comprehensive presentation of the subject field, lemmatized relevant multi-word expressions, and offered different, but complementary, ways of disambiguating meaning, especially in culture-dependent domains. These lexicographic decisions also permeate the planning and compiling of the Accounting dictionaries, as shown below.

3. The Accounting Dictionaries: *The Diccionario Inglés-Español de Contabilidad*

At an abstract level the relationship between lexicography and the Internet is mutually reinforcing, which lends support to the claims made on the inclusion of lexicography as part of Information Science (Leroyer 2011):

- Both lexicography and the Internet aim to provide quick and easy access to data from which users can retrieve the information needed in different types of social situations.
- The above aim is especially relevant in the current information and knowledge society, and hence both lexicography and the Internet contribute to this.
- Both lexicography and the Internet have to solve the problem of *data suffocation*, which occurs when users retrieve more data than needed.
- Both lexicography and the Internet target types of users in countless use situations, and not individual users in walled surroundings.
- Both lexicography and the Internet contain data that has been selected and prepared in the hope that potential users can convert this into information.
- Both lexicography and the Internet are currently analyzing options that will allow users to retrieve more focused hits.

A critical examination of most existing Internet accounting dictionaries shows that the majority of them are a long way from compliance with the above relationship (Fuertes-Olivera and Nielsen 2011). This relationship assumes the validity of two very general principles. The first supports the idea that we do not need a new theory of lexicography at the highest level of abstraction (Gouws 2011; Tarp 2011); the second maintains that what we need is an adaptation of lexicography to the technical options of the Internet (Bothma 2011). Both principles translate into lexicographical discussions that agree on presenting the online dictionary and other types of lexicographic tools as utility products; such products must be planned and compiled afresh, with the aim of meeting users' needs in the light of new technologies made available to practical lexicography (see Fuertes-Olivera and Bergenholtz (2011) for a detailed discussion on theoretical assumptions regarding e-lexicography and some practical applications).

An exception to the dismal situation reported on in Fuertes-Olivera and Nielsen (2011) is the planning and compilation of the Accounting dictionaries. At present, they comprise a set of two monolingual and three bilingual online dictionaries using Danish, English and Spanish (a Spanish dictionary and a Spanish-English dictionary are in the pipeline). The theoretical foundation underlying the project gives priority to lexicographic functions, i.e., the assistance these dictionaries can give users in specific types of situation where they require knowledge to resolve issues relating to accounting (Bergenholtz and Tarp 2003, 2004; Nielsen and Mourier 2007; Tarp 2007 and 2008). The dictionaries are designed to meet certain types of user needs through the careful selection of data and specific options for data access, so that the answers to users' questions match their needs in a number of different contexts.

The Accounting dictionaries, which were originally developed at the Centre for Lexicography from 2002 onwards, can be divided into two different groups. On the one hand, those that deal with the English and Danish languages (Nielsen et al., 2010 and 2011) have been adapted to the practical and theoretical foundations of online lexicography "in light of the electronic options available to produce targeted reference tools and the advances described in the literature" (Nielsen and Almind 2011). This means that users can retrieve data depending on the function they select. For example, in one of the monolingual dictionaries they can elect to search for the following kinds of assistance:

- Help to understand an accounting term.
- Help to produce an accounting text where the expression is known.
- Help to find a term where the meaning is known.
- Show all data.

Tarp (2009: 57-58) illustrates the lexicographical options available in relation to text reception and text production. For example, if a Danish speaker reading an

English text has a reception problem, the corresponding need for information may be provided in various ways depending on the characteristics of the reader. If the reader's proficiency level in terms of English accounting terminology and English language is high, then an English explanation would be sufficient in order to meet their needs. But if the reader's general and special English proficiency level is low, then their needs can only be satisfied by a Danish equivalent and/or explanation. Finally, if the proficiency level is more or less mid-way between the two former, then a combination of Danish and English would be the best way to ensure that the needs are met.

On the other hand, the accounting dictionary that deals with Spanish, *El Diccionario Inglés-Español de Contabilidad*, is only equipped with a kind of 'show all function', and consequently users retrieve all the data contained in the dictionary article they consult.

Examples (1), (2) and (3) illustrate the differences between the two groups of Accounting dictionaries. A user who consults **deemed cost** in the English Accounting Dictionary obtains example (1) when they elect to choose 'help to understand an accounting term', example (2) when they choose 'help to produce accounting texts where the expression is known', whereas they retrieve example (3) when consulting *El Diccionario Inglés-Español de Contabilidad*:

- (1) Definition as help to understand an accounting term

deemed cost <a, the, -s>

Definition

Deemed cost is an amount used instead of cost or depreciated cost at a specific date. Any following amortisation or depreciation is made on the assumption that the enterprise initially recognised the asset or liability at a cost equal to the deemed cost.

- (2) Data presented to give help to produce accounting texts where the expression is known

deemed cost <a, the, -s>

Definition

Deemed cost is an amount used instead of cost or depreciated cost at a specific date. Any following amortisation or depreciation is made on the assumption that the enterprise initially recognised the asset or liability at a cost equal to the deemed cost.

Collocations

fair value or revaluation as deemed cost
have a deemed cost of zero in the opening IFRS balance sheet
the deemed cost of goodwill
use an event-driven fair value measurement as deemed cost

Examples

Items of property, plant and equipment are measured at cost as deemed cost less accumulated depreciation and impairment losses.

(3) Data presented when the search word is **deemed cost**

deemed cost <a, the, -s>

definition

Deemed cost is an amount used instead of cost or depreciated cost at a specific date. Any following amortisation or depreciation is made on the assumption that the enterprise initially recognised the asset or liability at a cost equal to the deemed cost.

coste atribuido

synonyms

coste asignado

collocations

- fair value or revaluation as deemed cost
valor razonable o revalorización como coste atribuido
- have a deemed cost of zero in the opening IFRS balance sheet
tener cero como coste atribuido en el balance de apertura basado en la NIIF
- the deemed cost of goodwill
el coste atribuido del fondo de comercio
- use an event-driven fair value measurement as deemed cost
utilizar una medición del valor razonable de un acontecimiento inevitable como coste atribuido

examples

- Items of property, plant and equipment are measured at cost as deemed cost less accumulated depreciation and impairment losses.
Las partidas de edificios, instalaciones y equipos se valoran al coste como coste atribuido menos amortización acumulada y deterioros.

sources

IFRS1, Appendix A

Briefly, the data retrieved has been tailored to the user's needs in the use situation when they consult the *English Accounting Dictionary*, whereas this option is not yet available in the *Diccionario Inglés-Español de Contabilidad*. This English-Spanish dictionary is a polyfunctional dictionary that helps to read and understand and produce accounting texts, translate English accounting texts into Spanish, and acquire knowledge about accounting matters. In other words, the *Diccionario Inglés-Español de Contabilidad* does not yet benefit "from an overall and more stringent functional approach where the corresponding search mechanisms may lead their users to dynamic data (articles) that are especially adapted to each and every of the various types of users and situations covered by the dictionary in question." (Tarp 2009: 57-58).

With the aim of minimizing the risks associated with retrieving static data, the *Diccionario Inglés-Español de Contabilidad* has incorporated functions which are not usually found in other Internet accounting dictionaries. For example, it includes the search option *contains*, with which users can retrieve all the articles in which the search word is part of the word list. A search for the term **account-**

ing with *contains* in the scroll menu retrieves 92 articles. The short definition given in each article, as well as the data included in the article, allows users to gain some fundamental knowledge about this key accounting concept (Fuertes-Olivera and Tarp 2011). In addition, the lexicographers also adopted decisions that can be explained from a user-oriented lexicographic approach. For instance, the decisions concerning the selection of Spanish equivalents, which are discussed below, can be explained in a lexicographic way and are based on the characteristics of LSP lexicography.

4. A Case Study: Selecting Spanish Equivalents for English Accounting Terms in the *Diccionario Inglés-Español de Contabilidad*

As a research methodology, case study research has been extremely influential in shaping the way we do research in the Social Sciences (Casanave 2010). For example, in the field of education, Wolcott's study on general education (1973), Halliday's study of Nigel's first language acquisition (1975), and Schmidt's study in second language learning (1983) can be cited, as they "have shaped discussion and research in their respective fields of focus in forceful and productive ways." (Van Lier 2005: 195).

Case study research is a contextual type of investigation that has three main characteristics (Casanave 2010). First, its object of inquiry is unique and delimited, i.e., bounded. This means that the researcher's interest is in the particular rather than the general. In other words, a case study investigates one person, one group, one institution, one community, or a particular bounded phenomenon. Secondly, by choosing a case study tradition, researchers are committed to presenting in-depth analyses of particular phenomena, be they particular people, sites, groups, institutions, dictionary articles, teaching materials, etc. Thirdly, case studies are always situated or embedded in a particular context (Yin 2003).

The case study reported on in this paper is concerned with the selection of equivalents in the specialized Internet dictionary *El Diccionario Inglés-Español de Contabilidad*. Instead of focusing on quantitative methods, the case is delimited to explaining in detail the rationale for the decisions taken in the framework of functional approaches to specialized lexicography. Such decisions not only hypothesize the adequacy of deductive methodologies but also conclude that lexicographers' decisions must first of all be discussed in terms of the function(s) the dictionary aims to solve in specific use situations where users' needs can be met; secondly, they should be situated in the area of specialized language, whose characteristics and defining properties make it different from general language.

A review of the literature that examines equivalence shows that most researchers explain the concept of equivalence in terms of the abstract value it has in the system and its concrete presence in the text (i.e., interlingual or intertextual equivalence), as well as its corollary of *full equivalence*, *partial equiva-*

lence, and *zero equivalence* (Pedersen in Bergenholtz and Tarp 1995: 104-110; Fuertes-Olivera and Arribas-Baño 2008: 88-90; Adamska-Salaciak 2010: 387-409). This research examines equivalence in connection with bilingual lexicography, although researchers do not ignore the fact that the presence of translation equivalents does not make the dictionary bilingual (Gouws 2004), nor that all 'so-called' bilingual dictionaries can be used for translation (Tarp 2005). Nkomo (2008: 132) is right when he says that we have "to dispel some erroneous beliefs which result in dictionaries which provide translation equivalents as data categories being misused and hence unfairly evaluated."

Adamska-Salaciak (2010), for example, examines equivalence as an interdisciplinary concept, and defends an explanation of the concept of equivalence by looking outside lexicography, i.e., in the area of disciplines such as contrastive linguistics and translation. Regarding contrastive linguistics, she claims that it is convenient to distinguish two stages in the explanation offered, and adds that lexicographers can mostly benefit from the second stage, which appeared with the advent of electronic corpora. In particular, she urges the use of comparable and/or parallel corpora for extracting interlingual correspondences as an appropriate methodology for solving the practical problem of selecting an equivalent in lexicography. Regarding translation, she accepts Halverson's (1999 and 2002) suggestion that translation be viewed as a prototype concept that allows us to understand equivalence as "a broad spectrum of relations, from similarity to identity." (Adamska-Salaciak 2010: 403).

My view of specialized lexicography is rather different. I agree with her that electronic collections of texts can offer clues as to which equivalent must be selected. However, this is a practical activity that cannot be translated into offering a theoretical examination of the decisions taken in a real specialized lexicography project. For instance, corpora inform researchers about frequency but not about relevance, and they illustrate the existence of variation. Both issues, frequency and variation, run contrary to the nature of specialized communication. Similarly, prototypical equivalence is not adequate in specialized communication, in which, given that redundancy, widespread in general language, is typically avoided in specialized language, only full equivalence is a solution. For example, supposing we are dealing with an English International Accounting Standard (IAS) concept, we have to offer the exact full Spanish equivalent and not a proxy which could confuse investors, who might then ask for compensation if this confusion cost them money.

For specialized dictionaries, the concept of equivalence has to be addressed in lexicographical terms, and this implies taking into consideration the intended user of the dictionary, the intended use situation, the intended function(s) and the nature of specialized communication, especially when the subject domain can be described as culture-dependent (Mheta and Muhwati 2009). For example, if the dictionary aims primarily at meeting a user's needs when reading a text, the inclusion of an explanatory equivalent or a translational equivalent is more or less irrelevant. However, if the dictionary's main aim is to

meet the needs a user has when translating a text, it is very important to include a translational equivalent instead of, say, an explanatory one. Similarly, if the dictionary is concerned with producing texts, then the inclusion of synonyms is appreciated by the intended user, who will regard them as suitable equivalents and useful lexicographic structures in the communicative situation associated with text production.

The process of equivalent selection in LSP lexicography needs a lexicographic explanation that initially contemplates the true nature of lexicography as well as that of LSP communication in today's world. Contrary to what occurs in general lexicography, where new words and meanings need time to become standardized and accepted (usually because redundancy is widespread and meaning fluctuates before standardization), terms are accepted from the outset, and are then coined and adapted by experts; the latter then give these terms an unambiguous denomination, using them in professional genres and registers before they are noticed by terminology standardization committees and professional and/or amateur lexicographers. This implies adopting an approach to equivalent selection that acknowledges, first of all, experts' tendency to coin and use terms before they are sanctioned by terminological committees; secondly, that terms in widespread use cannot be eliminated; and, thirdly, that term formation in languages such as Spanish is very much influenced by economic reasons and the spread of English as the *lingua franca* of LSP communication. This occurs especially when dealing with disputed terms, as shown below.

Economic reasons are manifest in the fact that modern technology is based on the principle of the division of work, and that it does not function without verbal communication (Teubert 2005). This has brought to the fore the fact that terminology is also a commercial activity, with many companies paying large sums for acquiring multilingual terminologies that describe the products they sell, establishing their own terminology departments, and fighting to create, develop and standardize "the LSP they use themselves and want to force their competitors to use." (Teubert 2005: 97).

It is not only that technology rules our modern world, but that this technology is no longer self-explanatory. It needs instructions, leaflets for describing the product, the installation procedures, etc. In short, modern technology needs to improve communication strategies by which a given term denotes the same concept for producer (or developer) and user.

The spread of English in LSP communication, which is associated with economic globalization, is also making a deep impact on specialized lexicography. For example, it is blurring the difference between culture-dependent and culture-independent domains, as will be shown below when I explain the selection of International Accounting Standard/International Financial Reporting Standard (IAS/IFRS) terms in Spanish.

This general framework has influenced the process of equivalent selection in the *Diccionario Inglés-Español de Contabilidad*, which comprises both the selection of a proper equivalent and its synonym (where necessary), and which can

be explained as the result of three lexicographical principles that have permeated our lexicographic project: relevance, proscription and recreation. For the sake of simplicity, I will restrict the explanation to decisions taken when lexicographers face several possibilities, i.e., when equivalents are disputed and lexicographers cannot find a standardized full Spanish equivalent.

Fuertes-Olivera and Nielsen (2011) claim that the principle of *relevance* refers to the quality of being directly connected with the subject field in question, the function(s) of the dictionary, and the situation in which the dictionary is intended to be used. As the *Diccionario Inglés-Español de Contabilidad* is primarily directed at assisting Spanish translators of English accounting texts, the principle of relevance requires the selection of an insertable equivalent, whatever this may be called. Hence, in the *Diccionario Inglés-Español de Contabilidad*, there is always one, and only one, insertable equivalent per English lemma (see Example 3, above). This decision, which is based on the tenets of the function theory of lexicography (Bergenholtz and Tarp 2003 and 2004), follows Nielsen et al.'s practice in their planning and compilation of the Accounting dictionaries in Danish and English (2010 and 2011). Moreover, this practice agrees with the principle maintaining that more choices do not necessarily lead to greater customer satisfaction, and is a kind of 'proxy' device which is intended to eliminate the stress users suffer when a search retrieves much more data than needed (the so-called 'Google suffocation effect').

By including only one equivalent per lemma, the dictionary helps translators to use the right translation equivalent. For example, the English term "**balance**" has four dictionary articles, and therefore, four Spanish equivalents: *balance*, *saldo*, *deuda* and *saldar*. This practice constitutes a novelty in Spanish specialized lexicography, which is characterized by bilingual dictionaries that include as many equivalents as possible, usually without disambiguating them properly, as example (4) shows:

- (4) The entry "**balance**" in the *Diccionario Técnico Económico-Financiero-Actuarial. Inglés-Español* (Villalón and Martínez Barbeito 2008)

Balance: Saldo, diferencia. Balanza, equilibrio. Balance, resto.

The principle of relevance also explains the fact that lexicographers must opt for a one-to-one equivalence in specialized lexicography. Contrary to what has been claimed in general lexicography in the context of bilingualized lexicography, users of specialized dictionaries need lexicographic structures that offer them very precise meaning, not only because this is connected with the very nature of specialized discourse, but also because redundancy tends to be absent from specialized texts. In specialized lexicography users require meaning, therefore, to be precise. In the Accounting dictionaries this was achieved by also including a Cobuild-style definition per lemma. Hence, the inclusion of an English definition and a Spanish insertable equivalent makes the *Diccionario Inglés-Español de Contabilidad* a novelty in Spanish specialized lexicography,

offering an innovative lexicographic solution with which users can eliminate the problems associated with homonymy and polysemy in specialized texts. Furthermore, as users of specialized texts need lexicographic structures that make meaning very precise and unambiguous, this dictionary also has a well-conceived lexicographic layout that facilitates the disambiguation of homonymous and polysemous terms. These are differentiated by means of superscripts and Arabic numbers, respectively. In Figure 1 the superscripts ¹ and ² indicate that **balance** can be a noun or a verb, whereas the Arabic numbers 1 and 2 indicate that as a noun "**balance**" has two different meanings.

balance¹

<noun a, the, -s>

balance

1

definition

A balance is the net amount or amount remaining after part has been taken away, i.e. the difference between debits and credits in an account.

saldo

collocations

- a balance not written down
un saldo no amortizado
- the balance of accumulated profit or loss at the beginning of the period
el saldo de la cuenta de pérdidas y ganancias a principios del periodo

example

- The account showed a positive balance of £350.
La cuenta mostraba un saldo positivo de 350 libras.

synonyms

account balance

2

definition

A balance is an amount owed, i.e. an outstanding debt payable.

deuda

collocations

- a comprehensive analysis of related party balances
un análisis global de las deudas de las partes implicadas
- balances in foreign currencies

balance²

<verb -s, -d, has -d, balancing>

definition

To balance means to reconcile the debit and the credit side of an account to determine the balance.

saldar

collocations

balance off

saldar

Figure 1: Excerpt from *El Diccionario Inglés-Español de Contabilidad*

The inclusion of only one insertable Spanish equivalent connects the process of equivalent selection with the choice of descriptive, prescriptive, and proscriptive lexicography; these three key lexicographic concepts have been analyzed by Bergeholtz (2003), and Bergeholtz and Gouws (2010), with the aim of making them precise in a lexicographic sense. Considering the dictionary as a practical tool implies that the terms descriptive, prescriptive and proscriptive cannot be applied to the whole dictionary but to the lexicographic processes present in the planning and compiling of every dictionary article. In other words, when a lexicographer uses data from a linguistic survey and a text investigation in their practical lexicographic work, the process is an example of descriptive lexicography (Bergeholtz 2003); the same lexicographer is using a prescriptive approach when they state that a specific linguistic variant is explicitly prohibited, or that one or more linguistic variants are explicitly prescribed (Bergeholtz 2003); finally, the lexicographer is using a proscriptive approach when they (explicitly) recommend a variant and their recommendation is based on an in-depth analysis of the several options available (Bergeholtz 2003) (see Bergeholtz and Gouws 2010 for a detailed lexicographic discussion of the different types of prescription, description and proscription identified so far).

The proscriptive approach is extremely important when planning and compiling a specialized translation dictionary, especially in three specific lexicographic situations which oblige lexicographers to face a hard choice: (i) when a new concept is introduced, (ii) when there is variation among terms, and (iii) when the linguistic structure of a language does not favour the translation of the lemma in a new language (for example, the translation of English *-ing* terms into Spanish is always difficult, and sometimes impossible). This could be an example of what Bergeholtz and Gouws (2010: 38) identify as a choice among descriptive, prescriptive and proscriptive lexicography, which has to be solved by resorting to a "language and communication policy decision". In such a situation the proscriptive approach means that lexicographers offer an insertable equivalent. Their decision is mostly conditioned by the nature of LSP communication, especially the necessity of having a one-to-one correspondence between a concept and its linguistic representation, which makes it necessary for lexicographers to include full equivalents in the lexicographic tools they plan and compile.

A case in point occurs when Spanish lexicographers are forced to decide on the various Spanish accounting terms that correspond to the English IAS/IFRS terms. The IASs and IFRSs (International Accounting Standards and International Financial Reporting Standards, respectively) are examples of equal texts, i.e., texts that have been prepared in a language (usually English) and then translated into the different official languages of an organization such as the European Union. Equal texts conceptualize (new) realities that have been identified in the Accounting dictionaries as IAS/IFRS terms. These terms are widespread in accounting (In the *Diccionario Inglés-Español de Contabilidad* there are 377 English IAS/IFRS terms.) and are used in order to refer to normalized and standardized concepts, i.e., concepts that were defined by the writers of the

above-mentioned standards and are, therefore, obligatory when preparing official accounting documents. In other words, accountants must use these IAS/IFRS terms whenever they draw up accounting documents.

It may happen that translators of these equal texts do not know that a particular concept already has a well-known established Spanish term, or they may commit a translation mistake, which can easily lead to a conceptual mistake. Both situations occur with the Spanish IAS/IFRS terms. Examples of nonsensical Spanish renderings were found, as well as the introduction of unnecessary Spanish terms, i.e., the translators of the IAS/IFRS rules were unaware of the existing Spanish term and introduced a new one in the Spanish International Accounting Standards/International Financial Reporting Standards.

Both problems were resolved in the *Diccionario Inglés-Español de Contabilidad* by the inclusion of lexicographic notes (e.g., "Spanish accountants prefer *resultado de explotación* to the IAS/IFRS term *resultado operativo*"; or "the IAS/IFRS term is nonsensical in Spanish"), and by making frequent use of exact prescription, the term by means of which Bergenholtz and Gouws (2010) refer to the process of recommending one equivalent, although other equivalents may be mentioned. This resulted in the selection of traditional Spanish equivalents (for instance: *cuenta deudora*; *beneficio contable*; *producción agrícola*; *valor contable*; *tesorería*; *clasificado por tipo de gasto*; etc.) and the inclusion of the Spanish translated IAS/IFRS terms as synonyms (*cuenta a cobrar*; *ganancia contable*; *producto agrícola*; *valor en libros*; *efectivo*; *clasificado por naturaleza*; etc.)

Finally, the concept of *recreation* was used in *El Diccionario Inglés-Español de Contabilidad*. *Recreation* refers to the process of creating a Spanish equivalent whenever it is deemed necessary, either because there were fluctuations (variations) between competing terms, or because the existing Spanish term was not always insertable, perhaps because it could lead to misunderstanding due to the existence of different accounting traditions. The accounting experts who took part in the process of selecting Spanish equivalents referred to the necessity of using the process of *recreation* in two specific use situations, which are analyzed below.

The first one occurs when Spanish accountants typically use the English terms, usually as superordinates. A typical example occurs with "**rating**". Although traditionally-conceived English-Spanish dictionaries of business/economics include *calificación crediticia*, and *calificación de solvencia* as equivalents of English "**rating**", *rating* was maintained as the Spanish equivalent, and *calificación crediticia* and *calificación de solvencia* were included as synonyms.

The above decision is justified in terms of the external and internal subject classification made for selecting equivalents: this showed that *calificación crediticia* and *calificación de solvencia* could also be used in order to refer to the creditworthiness of a person, something that does not happen with "**rating**", which always refers to the creditworthiness of a country and/or organization. This difference is important in specialized lexicography, and could lead to misun-

derstandings if they were selected as equivalents. In other words, *calificación crediticia* and *calificación de solvencia* are not always full equivalents of English "rating", and might as a result lead to confusion and misunderstanding among users.

This problem disappears, however, if they are selected as synonyms, as this decision informs users that the terms are restricted to the exact meaning of "rating", and that they can be conveniently used when the text is unambiguous so as to comply with academic traditions requiring the use of synonyms in formal Spanish texts. To sum up, the lexicographers selected *rating* as equivalent because this term always shows *full equivalence*, whereas *calificación crediticia* and *calificación de solvencia* were selected as synonyms because both can also be used in other domains where they are no equivalents of English "rating".

The second situation involved selecting literal renderings of English terms as Spanish equivalents. This process was used whenever a literal translation was found that could be understood by Spanish speakers. The frequent use of recreation in this situation can be explained by referring to the status of English as the *lingua franca* of specialized communication, and by resorting to explanations based on the translation agent model, which aims to offer clues as to the cognitive processes undertaken by translators when translating (Chesterman 2009; Fuertes-Olivera submitted).

Lexicographers used to maintain that politics, economics and law are examples of culturally-dependent areas: they have developed their own specific features within a delimited geographical area. They also added that although a clear-cut distinction between culture-dependent and culture-independent domains cannot always be made, this problem had been solved by including partial equivalents and/or paraphrases rather than true equivalents. For example, some traditionally-conceived business/economics English-Spanish dictionaries, e.g., the *Diccionario Técnico Económico-Financiero-Actuarial* (Villalón and Martínez Barbeito 2008), often use paraphrase instead of a translation equivalent, as observed in example (5):

- (5) **Commodity pool:** Un fondo de productos que confía los fondos de muchos inversores a un fondo profesional de negociación y distribuye beneficios y pérdidas entre los participantes en proporción a sus intereses.

The above solutions are not appropriate in a translation dictionary: paraphrases are not insertable, and, moreover, partial equivalents can mislead users, as revealed in the discussion on **rating**. Instead, literal translations were produced, since this is what is typically done when experts come across an English term for the first time. This assumption agrees with the nature of specialized communication, i.e., with the necessity of finding a one-to-one correspondence between concept and term. Hence, it was deemed appropriate for specialized lexicography and consequently adopted, especially because this practice coincides with certain conclusions drawn in several empirical studies

showing that translators of specialized texts disambiguate and reproduce meaning more easily when the terms in the source and target language are conceptually related; this, for instance, occurs when they belong to the same conceptual scenario (Fuertes-Olivera and Nielsen 2011). This is particularly the case with multi-word terms, such as **business combination involving entities** or **businesses under common control**. For this English lemma the lexicographers selected the Spanish literal translation, *combinación de negocios entre entidades o negocios bajo control común*, as equivalents. In fact, such a decision is of considerable importance in this dictionary, considering that the number of lemmata similar to **business combination involving entities** or **businesses under common control** is quite large. For example, terms for accounting accounts, sub-accounts, classifications, sub-classifications, etc., reported in an accounting document, e.g., an **income statement** (see Table 1), were included as lemmas; these are conceptually and communicatively necessary. Table 1 shows the conceptual categories of an **Income Statement by Nature** in English and Spanish:

Income Statement by nature	Cuenta de resultados (cuenta de pérdidas y ganancias) por naturaleza
1. Continuing operations (<i>where appropriate</i>)	1. Operaciones continuadas (<i>si es adecuado</i>)
2. Revenue (Sales)	2. Ingresos por ventas
3. Other operating income	3. Otros ingresos de explotación
4. Changes in inventories of finished goods and work in progress	4. Variaciones de existencias de productos terminados y en curso de fabricación
5. Work performed by the company and capitalised	5. Trabajos realizados por la empresa para su activo
6. Raw materials and consumables	6. Materias primas y consumibles
7. Employee benefits expense	7. Gastos de personal
8. Other operating expenses	8. Otros gastos de explotación
9. Earnings before Interest, Tax, Depreciation and Amortisation (EBITDA)	9. Resultado antes de intereses, impuestos y amortizaciones (EBITDA)
10. Depreciation expense	10. Gasto por depreciación
11. Impairment of property, plant and equipment	11. Deterioro del inmovilizado
12. Earnings before Interest, Tax and Amortisation (EBITDA)	12. Resultado antes de intereses, impuestos y amortizaciones (EBITDA)
13. Amortisation expense	13. Gasto por amortizaciones
14. Impairment of goodwill	14. Deterioro del fondo de comercio
15. Earnings before Interest and Taxes (EBIT)	15. Resultado antes de intereses e impuestos (EBIT)
16. Finance costs, net	16. Gastos financieros netos
17. Share of profit of associates after tax	17. Participación en beneficios en asociadas después de impuestos
18. Profit before tax	18. Resultado antes de impuestos
19. Income tax expense	19. Impuestos sobre beneficios
20. Profit for the period	20. Resultado del ejercicio
Where appropriate distributed into:	Distribuido si es apropiado entre:
21. Profit for the period from continuing operations	21. Resultado del ejercicio procedente de operaciones continuadas
22. Discontinued operations	22. Operaciones interrumpidas

23. Profits for the period from discontinued operations 24. Profit for the period Attributable to: Equity holders of the parent Minority interests Earnings per share for profit* during the period - basic - diluted *profit attributable to equity holders of the parent	23. Resultado del ejercicio procedente de operaciones interrumpidas 24. Resultado del ejercicio Aplicable a: Matriz Intereses minoritarios Dividendo por acción durante el ejercicio - básico - diluido *resultado del ejercicio atribuible a la matriz
--	---

Table 1: The Income Statement by Nature in the *Diccionario Inglés-Español de Contabilidad*.

An analysis of Table 1 illustrates the rationale for the decisions taken when selecting the Spanish equivalents, which have been previously discussed and are summarized below. First, some of the equivalents are traditional Spanish equivalent terms that do not impose any burden on the lexicographer. These are bound to be encountered in real texts. They were found by means of an internal and external subject classification, as well as by the web being used as a corpus. For example, a consultation of official and relevant accounting texts, such as a Spanish *cuenta de resultado* and an English **income statement** or **profit and loss account**, offers clues about suitable standardized equivalents.

Secondly, some of the equivalents, for example, *cuenta de resultados* and *cuenta de pérdidas y ganancias*, fluctuate in Spanish: the former is the Spanish traditional term, whereas the latter is a literal translation that entered into Spanish through a translation. On such occasions the lexicographers selected the Spanish traditional term as an equivalent and the translated term as a synonym.

Thirdly, in some cases the original English term is adopted, i.e., as a borrowing from English, e.g., *EBITDA*. In these situations a Spanish term is usually included as a synonym, especially when the Spanish term is not an example of full equivalence. For instance, English accountants differentiate between **depreciation** and **amortization**, a differentiation that does not occur in Spanish traditional accounting texts, where the term *amortizaciones* is the only one used for both **depreciation** and **amortization**. As these two English terms differ slightly, using *EBITDA* as a Spanish equivalent minimizes the possibility of misunderstanding.

Fourthly, some of the Spanish equivalents are translations, usually literal ones, of English lemmas. For instance *variaciones de existencias de productos terminados y en curso de fabricación* is the Spanish translation of English **changes in inventories of finished goods and work in progress**.

5. Conclusion

This paper deals with specialized e-lexicography and maintains the necessity of

using a lexicographic approach for explaining the decisions taken when planning and compiling specialized dictionaries. Without disputing the (possible) advantages of the methods and procedures that are currently being used in general lexicography, this paper defends the view that LSP e-lexicography needs its own methods and practices, which take into consideration the nature of specialized communication and the technical options the Internet offers, and uses them within the tenets of a user-based approach, e.g., the function theory of lexicography.

The above idea is illustrated by an explanation of some of the decisions taken when selecting Spanish equivalents for the *Diccionario Inglés, Español de Contabilidad*, one of the Accounting dictionaries; these are a set of Internet specialized dictionaries that aim at satisfying the needs of translators (primary user group), accountants and financial experts (secondary user group), as well as students of accountancy and translation, journalists and interested laypersons (tertiary user group).

In particular, this paper elaborates on decisions concerned with selecting equivalents that are not yet well established, i.e., those that must refer to original English accounting terms, and which pose a lexicographic problem because they are new, difficult to translate, or fluctuate among several options. In such situations, we have explained the decisions as a result of three lexicographic principles that offer a rationale for most of the ones made in the dictionary here discussed: relevance, proscription, and recreation. The workings of these three principles and their rationale are explained as a case study, which implies a contextual form of research that is confined to an in-depth analysis of a bounded phenomenon; this was illustrated with a selection of dictionary entries revealing the workings of the three principles underlying lexicography that lend support to a lexicographic approach to terminology.

In conclusion, our exercise, i.e., the theoretical foundations on which fluctuating equivalents were selected, reinforces the condition of lexicography as a scientific endeavour. In this regard, it illustrates that equivalents were selected as a result of in-depth analyses of the nature of LSP lexicography and specialized communication, in addition to well-attested theoretical lexicographic considerations advocating the designing of specialized dictionaries planned and compiled to assist specific users to solve their information needs in specific use situations. Unfortunately, there are still many practitioners who continue planning and compiling dictionaries without taking into consideration the above reflections. This is what we typically found in traditional specialized translation dictionaries published in Spain (for example, the *Diccionario Técnico*). At best, such a dictionary is only useful for confirming the user's conjecture. It is to be hoped that the future of specialized lexicography, especially in places such as Spain, will reside in lexicographic principles and practices that are based on critical thinking and not on traditions, methods and practices taken for granted. This view is also shared by lexicographers at the Centre for Lexicography, who are now working with a new conceptual scenario: the building of lexicographic

data banks, usually called lexicographic databases, which are different from information tools. We should no longer view language as the main argument for classifying dictionaries, but rather the functions. Regarding the Accounting dictionaries, we have a single accounting database (Nielsen and Almind 2011), which can feed 23 totally monofunctional specialized e-dictionaries of accounting (Bergenholtz forthcoming): 13 of them (numbers 1 to 13 below), which have been prepared by Nielsen, Mourier, Bergenholtz and Almind (2011), are already in operation. The rest will be so in the near future:

Danish

1. Meaning of Danish accounting terms.
2. Use of Danish accounting terms, search with a known term.
3. Use of Danish accounting terms, search based on a known meaning.
4. Knowledge of Danish accounting terms, search with a known term.

English

5. Meaning of English accounting terms.
6. Use of English accounting terms, search with a known term.
7. Use of English accounting terms, search based on a known meaning.
8. Knowledge of English accounting terms, search with a known term.

English–Danish

9. Reception of English terms in English accounting texts.
10. Translation of English accounting texts into Danish.
11. Knowledge of English and Danish accounting terms.

Danish–English

12. Translation of Danish accounting texts into English.
13. Knowledge of English and Danish accounting terms.

Spanish

14. Meaning of Spanish accounting terms.
15. Use of Spanish accounting terms, search with a known term.
16. Use of Spanish accounting terms, search based on a known meaning.
17. Knowledge of Spanish accounting terms, search with a known term.

Spanish–English

18. Reception of English terms in English accounting texts.
19. Translation of Spanish accounting texts into English.
20. Knowledge of Spanish and English accounting terms.

English–Spanish

21. Reception of English terms in English accounting texts.
22. Translation of English accounting texts into Spanish.
23. Knowledge of English and Spanish accounting terms.

Acknowledgments

Thanks are due to Ministerio de Ciencia e Innovación (grant FFI2008-01703/FILO) and Junta de Castilla y León (grant VA039A09) for financial support. Thanks also go to Henning Bergenholtz for his comments on a draft of this paper. I myself am responsible for any remaining mistakes.

References

Dictionaries

- Kaufmann, U., H. Bergenholtz in cooperation with B. Stummann, S. Tarp, L. de la Rosa Marabet, G. la Serna Miranda and N. la Serna Torres. 1998. *Encyclopedic Dictionary of Gene Technology*. Volume I. English (with Spanish Equivalents). Toronto: Lugus.
- Nielsen, Sandro, Lise Mourier, Henning Bergenholtz, Pedro A. Fuertes-Olivera, Pablo Gordo Gómez, Marta Niño Amo, Angel de los Rios Rodicio, Angeles Sastre Ruano, Sven Tarp and Marisol Velasco Sacristán. 2009. *El Diccionario Inglés-Español de Contabilidad*. Available at: http://www.accountingdictionary.dk/regn/gbsp/regngbsp_index.php [01/09/2011]
- Nielsen, S., L. Mourier and H. Bergenholtz. 2010. Database and Design: R. Almind. *Danish Accounting Dictionary; Danish–English Accounting Dictionary; English Accounting Dictionary; English–Danish Accounting Dictionary*. URL: <http://www.ordbogen.com/ordboger/regn/index.php?dict=a007>. (Last accessed 20 January 2011.)
- Nielsen, S., L. Mourier and H. Bergenholtz. 2011. Database and Design: R. Almind. *Meaning of Danish Accounting Terms. Use of Danish Accounting Terms, Search with a Known Term. Use of Danish Accounting Terms, Search with a Known Meaning. Knowledge about Danish Accounting Terms, Search with a Known Term. Meaning of English Accounting Terms. Use of English Accounting Terms, Search with a Known Term. Use of English Accounting Terms, Search with a Known Meaning. Knowledge of English Accounting Terms, Search with a Known Term. Reception of English Terms in English Accounting Texts. Translation of English Accounting Texts. Knowledge about English–Danish Accounting. Translation of Danish Accounting Texts. Knowledge about Danish–English Accounting*. <http://www.Ordbogen.com>.

Villalón, J.G. and J. Martínez Barbeito. 2008. *Diccionario Técnico. Económico-Financiero-Actuarial. Inglés-Español*. Valladolid: Universidad de Valladolid.

Other Literature

- Adamska-Salaciak, A. 2010. Examining Equivalence. *International Journal of Lexicography* 23(4): 387-409.
- Andersen, B. and R. Almind. 2011. The Technical Realization of Three Monofunctional Phrasal Verb Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography. The Internet, Digital Initiatives and Lexicography*: 208-229. London/New York: Continuum.
- Barnbrook, G., P. Danielsson and M. Mahlberg (Eds.). 2005. *Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora*. London/New York: Continuum. [Research in Corpus and Discourse.]
- Bergenholtz, H. 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Bergenholtz, H. Forthcoming. Concepts for Monofunctional Accounting Dictionaries. *Terminology* 18.
- Bergenholtz, H. In press: Do We Need Another Theory of Lexicography? *Lexicon* 41.
- Bergenholtz, H. and Rufus H. Gouws. 2010. A Functional Approach to the Choice between Descriptive, Prescriptive and Proscriptive Lexicography. *Lexikos* 20: 26-51.
- Bergenholtz, H. and S. Nielsen. 2006. Subject-field Components as Integrated Parts of LSP Dictionaries. *Terminology* 12(2): 281-303.
- Bergenholtz, H. and S. Tarp (Eds.). 1995. *Manual of Specialised Lexicography*. Amsterdam/Philadelphia: John Benjamins. (Benjamins Translation Library 12.)
- Bergenholtz, H. and S. Tarp. 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes. Journal of Linguistics* 31: 171-196.
- Bergenholtz, H. and S. Tarp. 2004. The Concept of Dictionary Usage. *Nordic Journal of English Studies* 3: 23-36.
- Bergenholtz, H., S. Nielsen and S. Tarp (Eds.). 2009. *Lexicography at a Crossroad: Dictionaries and Encyclopedias Today, Lexicographica Tools Tomorrow*. Bern/Berlin/Bruxelles/Frankfurt am Main/New York/Oxford/Wien: Peter Lang. (Studies in Language and Communication 90.)
- Bothma, T. 2011. Filtering and Adapting Data and Information in an Online Environment in Response to User Needs. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography. The Internet, Digital Initiatives and Lexicography*: 71-102. London/New York: Continuum.
- Casanave, C. 2010. Case Studies. Paltridge, B. and A. Phakiti (Eds.). 2010: 66-79.
- Chesterman, A. 2009. The Name and Nature of Translator Studies. *Hermes. Journal of Language and Communication Studies* 42: 13-22.
- Fatá, I. 2010. The Bilingual Specialised Translation Dictionary for Learners. Fuertes-Olivera, Pedro A. (Ed.). 2010. *Specialised Dictionaries for Learners*: 83-106. Berlin/New York: De Gruyter.
- Fuertes-Olivera, Pedro A. 2009. Systematic Introductions in Specialised Dictionaries: Some Proposals in Relation with the Accounting Dictionaries. Nielsen, Sandro and Sven Tarp (Eds.). 2009. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: 161-178. Amsterdam/Philadelphia: John Benjamins.
- Fuertes-Olivera, Pedro A. (Ed.). 2010. *Specialised Dictionaries for Learners*. Berlin/New York: De Gruyter.

- Fuertes-Olivera, Pedro A.** Submitted. Translation Studies and The Literal Translation Hypothesis. Submitted for publication to *META* in 2012.
- Fuertes-Olivera, Pedro A. and Ascensión Arribas-Baño.** 2008. *Pedagogical Specialised Lexicography: The Representation of Meaning in English and Spanish Business Dictionaries*. Terminology and Lexicography Research and Practice (TLRP) 11. Amsterdam/Philadelphia: John Benjamins.
- Fuertes-Olivera, Pedro A., Pablo Gordo Gómez, Marta Niño Amo and Angel de los Rios Rodicio.** 2010. Expert Knowledge and Translation Equivalents in the *English-Spanish Accounting Dictionary*. Caballero Rodríguez, Rosario and María Jesús Pinar Sanz (Eds.). 2010. *Ways and Modes of Human Communication*: 1083-1090. Ciudad Real: Universidad de Castilla la Mancha.
- Fuertes-Olivera, Pedro A. and H. Bergenholtz (Eds.).** 2011. *e-Lxicography: The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum.
- Fuertes-Olivera, Pedro A. and Sandro Nielsen.** 2011: The Dynamics of Terms in Accounting: What the Construction of the *Accounting Dictionaries* Reveals about Metaphorical Terms in Culture-bound Subject Fields. Temmerman, Rita and M. van Campenhoudt (Eds.). 2011. *The Dynamics of Terms in Specialized Communication. An Interdisciplinary Perspective*. Special Issue of *Terminology. International Journal of Applied Issues in Specialized Communication* 17(1): 157-180.
- Fuertes-Olivera, Pedro A. and Marta Niño Amo.** 2011. Internet Dictionaries for Communicative and Cognitive Functions: *El Diccionario Inglés-Español de Contabilidad*. Fuertes-Olivera, Pedro A. and H. Bergenholtz (Eds.). 2011. *e-Lxicography: The Internet, Digital Initiatives and Lexicography*: 168-186. London/New York: Continuum.
- Fuertes-Olivera, Pedro A. and Sven Tarp.** 2011. Lexicography for the Third Millennium: Cognitive-oriented Specialised Dictionaries for Learners. *Ibérica. Journal of the European Association of Languages for Specific Purposes* 21: 141-162.
- Gottlieb, H. and J.E. Mogensen (Eds.).** 2007. *Dictionary Visions, Research and Practice*. Amsterdam/Philadelphia: John Benjamins. (Terminology and Lexicography Research and Practice 10.)
- Gouws, Rufus H.** 2004. Monolingual and Bilingual Learner's Dictionaries. *Lexikos* 16: 264-274.
- Gouws, Rufus H.** 2011. Learning, Unlearning and Innovation in the Planning of Electronic Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lxicography. The Internet, Digital Initiatives and Lexicography*: 17-29. London/New York: Continuum.
- Halliday, M.A.K.** 1975. *Learning How to Mean*. London: Edward Arnold.
- Halverson, S.L.** 1999. Conceptual Work and the 'Translation Concept'. *Target* 11(1): 1-31.
- Halverson, S.L.** 2002. Cognitive Models, Prototype Effects and Translation: The Role of Cognition in Translation (Meta)theory. *Across Languages and Cultures* 13: 21-43.
- Hartmann, R.R.K.** 2009. Keeping in Touch. A Survey of Lexicography Periodicals. *Lexikos* 19: 404-422.
- Hinkel, E. (Ed.).** 2005. *Handbook of Research in Second Language Teaching and Learning*: 195-208. Mahwah, NJ: Lawrence Erlbaum Associates.
- Leroyer, P.** 2011. Change of Paradigm in Lexicography: From Linguistics to Information Science and from Dictionaries to Lexicographic Information Tools. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lxicography. The Internet, Digital Initiatives and Lexicography*: 121-140. London/New York: Continuum.
- Mheta, G. and I. Muhwati.** 2009. The Use of Loan Translation as a Term-Creation Strategy in *Dura-mazwi reMimhanzi*. *Lexikos* 19 Supplement: 150-156.

- Nielsen, S.** 1994. *The Bilingual LSP Dictionary. Principles and Practice for Legal Language.* (Forum für Fachsprachen-Forschung 24.) Tübingen: Gunter Narr.
- Nielsen, S.** 2010. Specialised Translation Dictionaries for Learners. Fuertes-Olivera, Pedro A. (Ed.). 2010. *Specialised Dictionaries for Learners: 69-82.* Berlin/New York: De Gruyter.
- Nielsen, Sandro and R. Almind.** 2011. From Data to Dictionary. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography. The Internet, Digital Initiatives and Lexicography: 141-167.* London/New York: Continuum.
- Nielsen, S. and L. Mourier.** 2007. Design of a Function-based Internet Accounting Dictionary. Gottlieb, H. and J. E. Mogensen (Eds.). 2007: 119-135.
- Nielsen, S. and S. Tarp (Eds.).** 2009. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz.* Amsterdam/Philadelphia: John Benjamins.
- Nkomo, D.** 2008. Towards a Theoretical Model for LSP Lexicography in Ndebele with Special Reference to a Dictionary of Linguistic and Literary Terms. Master Thesis. University of Stellenbosh. Available at: <https://scholar.sun.ac.za/bitstream/handle/10019.1/1954/Nkomo.D.pdf?sequence=1> [Last accessed: January, 20, 2011.]
- Nord, C.** 1997. *Translating as a Purposeful Activity. Functionalist Approaches Explained.* (Translation Theories Explained 1.) Manchester: St. Jerome.
- Paltridge, B. and A. Phakiti (Eds.).** 2010. *Companion to Research Methods in Applied Linguistics.* London: Continuum.
- Schmidt, R.** 1983. Interaction, Acculturation, and the Acquisition of Communicative Competence. Wolfson, N. and E. Judd (Eds.). 1983: 137-174.
- Tarp, S.** 2005. The Pedagogical Dimension of a Well-conceived Specialised Dictionary. *Ibérica* 10: 7-21.
- Tarp, S.** 2007. Lexicography in the Information Age. *Lexikos* 17: 170-179.
- Tarp, S.** 2008. *Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography.* Lexicographica Series Maior 134. Tübingen: Max Niemeyer Verlag.
- Tarp, S.** 2009. Reflections on Data Access in Lexicographic Works. Nielsen, S. and S. Tarp (Eds.). 2009: 43-62.
- Tarp, S.** 2011. Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography. The Internet, Digital Initiatives and Lexicography: 54-70.* London/New York: Continuum.
- Teubert, W.** 2005. Language as an Economic Factor: The Importance of Terminology. Barnbrook, G., P. Danielsson and M. Mahlberg (Eds.). 2005: 96-106.
- Van Lier, L.** 2005. Case Study. Hinkel, E. (Ed.). 2005: 195-208.
- Wolcott, H.E.** 1973. *The Man in the Principal's Office. An Ethnography.* New York: Holt, Rinehart & Winston.
- Wolfson, N. and E. Judd (Eds.).** 1983. *Sociolinguistics and Second Language Acquisition.* Rowley, MA: Newbury House.
- Yin, R.K.** 2003. *Case Study Research: Design and Methods* Third edition. Thousand Oaks, CA: Sage Publications.

From Corpus to Dictionary: A Hybrid Prescriptive, Descriptive and Proscriptive Undertaking

Minah Nabirye, *Department of African Languages and Cultures,
Ghent University, Ghent, Belgium (minah.nabirye@UGent.be)*
and

Gilles-Maurice de Schryver, *Department of African Languages
and Cultures, Ghent University, Ghent, Belgium; and
Xhosa Department, University of the Western Cape, Bellville,
South Africa (gillesmaurice.deschryver@UGent.be)*

Abstract: Despite some heroic efforts over the past few years, Lusoga remains mostly underdeveloped. It is under continuous pressure from more prestigious languages, such as the neighbouring Luganda and especially the only official language in Uganda, English. Lusoga is undergoing rapid language shifts, with new concepts entering the language daily. Ironically, this process is taking place before Lusoga has even been properly reduced to writing. There is no single official orthography that is truly being enforced; people who do write, write as they think fit. Language data is needed for the production of reliable reference works. In the absence of a substantial body of published material in Lusoga, the researcher can resort to recording and transcribing the living language. This opens Pandora's box, in that spoken language (which is meant to be heard, and is typically less formal) is far more complex than written language (which is meant to be read, and is typically more formalised). Spoken and written variants are, by definition, different. And yet one wants to move the language forward, in a way, before the time is ripe. But then, with over two million speakers, how much longer can one wait? This article reports on the building of a new Lusoga corpus, nearly half of which consists of transcribed oral data. The writing problems encountered during the transcription effort are given detailed attention. Dealing with those writing problems in lexicography requires a multipronged approach. While most could be solved by laying down a norm, and thus through prescriptive lexicography, others need a more cautionary approach, and thus descriptive lexicography. Others still can only sensibly be solved when the lexicographer proposes certain options in defiance of existing norms and assumptions, at which point proscriptive lexicography needs to be called in.

Keywords: LUSOGA, UGANDA, ORTHOGRAPHY, SPELLING, CORPUS, ORAL, SPOKEN, TRANSCRIPTION, FULL WORDS, COMPOUNDS, MULTIPLE FORMS, LOANWORDS, BORROWINGS, FORMALITY LEVELS, CONCORDS, PRESCRIPTIVE LEXICOGRAPHY, DESCRIPTIVE LEXICOGRAPHY, PROSCRIPTIVE LEXICOGRAPHY

Obufunze: Okulondoola engeli Eitu ly'Olusoga bwe Linaatuusibwa mu Iwanika: Omutindo ogulaga Olulimi bwe luli, bwe luteekwa okuba oba bwe lube lutwalibwe. Empandiika y'Olusoga ekaali inhuma inho waile nga waliwooku obubonelo

obulaga enkola y'obuzila mu kugizimba. Olusoga lukaawagamiile wansi w'ennimi nga Oluganda n'olulimi olw'eiwanga, Olungeleza. Ensonga endala ku dhikaalemeisa Olusoga okwetengelela n'okuba nti lulina ebigambo ebiwingila buli olukeile. Eky'embu, enkyukakyuka ebigambo bino ebiyaaka by'eleetawo eli kwidha mu kiseela nga Olusoga lwene lukaali kufunilwaku mpandiika ntongole. Waile walifu dh'Olusoga buti dhiliwo kamaala, wazila ndala ku dho eli kugobelelwa mu kuwandiika kubanga abantu bakaawandiika nga bwe babona. Ebiwandiiko ebili ku Lusoga oba mu lulimi Olusoga byetaagibwa okulaga Olusoga bwe lulina okuba ela n'ebilina okwebuuzibwaku. Eibula ly'ebiwandiiko oti ni bino litegeeza nti omunoonheleza alina okwefunila entambi dh'Olusoga olwogelwa ela yeewanulila Olusoga oluli mu ntambi edho okusobola okutegeela engeli olulimi olwo bwe luli mu kiseela ekyo. Kino kileeseewo obuzibu obundi nti Olusoga olwogelwa tilutongoze ate lulimu emigote kamaala egyandibaile gilondoolwa okusenziila ku mutindo omusengeke singa lubailem mu buwandiike. Ekika ky'Olusoga oluli mu mbeela eyogelwa kya ndhawulo ku kili mu mbeela y'obuwandiike. Okuwanula Olusoga okuva mu ntambi kw'aba nti kugiililiile kwagayaga na kunoga kibala kikaali kwenga bukalamu. Aye engeli ye kili nti Olusoga lulina aboogezi abaswika mu bukaile obubili, abanoonheleza ku Lusoga baalisaine kutandiika li? Olupapula luno lulambulula enzimba y'eitu ly'Olusoga nga ekitundutundu ky'eitu lino kiviile mu Lusoga olwawanulwa okuva mu mbeela eyogelwa. Obuzibu obwayagaanibwa mu mpandiika y'ebigambo ebyawanulwa n'obusimbiibwaku eisila. Okulaga engeli y'okuzigula obuzibu bw'empandiika mu isomo ly'amawanika kwetaagisa enkola eteekubila inho ku nsonga ndala aye enoonheleza engeli esinga kugasa omutendela gw'obuzibu obulondoolebwa. Waile nga obuzibu obundi busobola okugondhoolwa okugila ku mutendela oguteebwawo, gwalaga olulimi nga bwe luteekwa okuba, obuzibu obundi bwetaagamu okwegendeleza nga wano mubaamu enkola y'okulaga nga olulimi bwe luli. Ate bwo obuzibu obundi bwandyetaagisa abawandiisi b'amawanika okuwaayo obudhulizi obusenziilwaku endowooza dhaibwe edhitagobeela mitendela giliwo nga balaga olulimi bwe lube lutwalibwe waile nga kino kyandiba nga kikontana n'amateeka g'olulimi agaliwo.

Ebigambo ebikulu: OLUSOGA, UGANDA, WALIFU, EMPANDIIKA, EITU LY'OLUSOGA, ENDHOGELA, OKUWANULA OLUSOGA, EBIGAMBO EBILAMBA, EBIGAMBO EBIGAITILILE, ENNAMULA KU BIGAMBO EBYAWUKANA, EBIGAMBO EBYEYAZIKE, EMITENDELA GY'OBUTONGOLE, ENNHUNGA YA NALIINA, OLULIMI BWE LUTEEKWA OKUBA, OLULIMI BWE LULI, OLULIMI BWE LUBE LUTWALIBWE

1. The Ugandan mother-tongue education policy

In September 2005, the Ugandan Parliament passed the teaching of nine regional indigenous languages in primary schools. Implementation officially started in February 2007. The languages concerned are Runyoro-Rutooro, Runyankore-Rukiga, Luganda, Lusoga, Rukonjo, Lugbara, Acholi, Ateso, and Karamojong (see NCDC 2006 and 2006a).

Elementary considerations required to prepare the introduction of all these languages into the mother-tongue education system are, however, pending. This is mostly detrimental for languages like Lusoga which are proposed as a medium of instruction for the very first time. The most crucial requirements for the successful implementation of this policy are the training of

mother-tongue educators and the production of reference works for the target audience.

Nabirye and De Schryver (2010) identified two target audiences for Lusoga: the primary-school teachers and their pupils. For these two audiences, Lusoga is supposed to be used both as the medium of instruction and as a subject in primary 1-3, but only as a subject in primary 4-7. Primary teachers do not learn Lusoga throughout their education and they are not Basoga trained to teach Lusoga as a subject either. As teachers they are nonetheless expected to conduct lessons in and on Lusoga, and to grade the pupil's Lusoga exercises.

2. The available Lusoga data

To date, no more than a handful of textbooks, spelling guides, and dictionaries exist for Lusoga. The available literature is limited to a few dozen booklets with stories only, as well as some religious publications. Unfortunately, in all this material, a variety of spellings is found. Moreover, given very few people currently write in Lusoga, there are numerous word forms in speech which have not been codified in any of the existing reference works. Word forms that lack a full conceptualization to merit a specific writing format are considered to be especially problematic. This state of affairs thus poses an additional problem when the aim is to uplift the status and use of Lusoga. Merely building a corpus of the language, where that corpus is internally consistent, proves to be a challenge in itself. In today's linguistic world which is driven by the use of electronic text corpora — with which a range of pedagogically sound reference works could and should be compiled — this poses problems for the corpus builder.

Should the corpus builder unify the various orthographies? Or is it better to reflect and thus keep the various, original spellings? If a single orthography is to be chosen, which one, or rather, whose orthography? Conversely, if a variety of orthographies is (or even if *all* are) kept, which one(s) should be used when transcribing new (i.e. oral) material? Answering these questions is not as trivial as it may seem. For example, while linguists and lexicographers may want to standardize the spelling, they run the risk to 'overcorrect' their original primary sources, making changes which end up masking the true processes at hand. But then, can one really handle a corpus which contains a variety of spellings for what are basically the same words?

In De Schryver and Nabirye (2010) it was shown how such a heterogeneous corpus can indeed be queried, and how, with it, one can undeniably acquire greater insights into the structure of the language. Teachers will — in contrast to linguists — need more definite guidance, however, and even though one of the orthographies (viz. Namyalo et al. 2008) has now been approved as the standard for writing Lusoga, being aware of the range of outstanding writing problems is a first step towards solving them (assuming they are all solvable, which, as we will see, is not always the case).

Over the past few years, an organic Lusoga corpus has been built,¹ the composition of which is currently (July 2011) as shown in Table 1.

Table 1: Composition of the Lusoga Corpus (July 2011)

Year	Genre	Breakdown of the data	Tokens	%
1998	Religious texts	The New Testament, ...	199 853	18.20
1998-1999	Newspapers (now defunct)	<i>Kodh'eyo, Ndiwulira</i>	187 393	17.07
1998-2003	Short stories	<i>Kintu, Ababita Ababiri, ..., Ebindi kw'idembe ery'Obw'omuntu mu nsi Yoonayoona</i>	150 560	13.71
2008-2009	Reports	Busoga clan leadership, Private sector, Academia, ...	24 166	2.20
2008-2011	Transcribed recordings	Songs, Traditional ceremonies, Speeches, Sermons, Interviews, Radio broadcasts, ...	413 827	37.69
2009-2010	Translations out of English	PEAP (<i>Poverty Eradication Action Plan</i>), ICEE (<i>International Centre for Eye Education</i>), FIDA/PLAN (inheritance laws), Home Structure, Public Health (questionnaire for baseline survey on institutional deliveries), JLOS (village imprisonment laws), Private Sector Uganda, Commercial advertisements, Message addressed to the Busoga clan governance	19 814	1.81
2010-2011	Electronic texts	E-mails (Personal and Mailing lists), Facebook postings, Material in electronic form, Various abstracts, ...	102 365	9.32
			1 097 978	100.00

As may be seen from Table 1, the total size of the Lusoga corpus now stands at 1.1 million running words or tokens. With well over 400 thousand tokens, the oral component of this Lusoga corpus is as high as 38%. Although all the transcriptions in the oral component have been made by the same person² — dili-

gently transcribing and adding material on a daily basis over a period of four years now — it has proven near impossible to be 'consistent'. This is mainly so because the transcription effort brought new and unusual constructions to the fore, which have not been documented in any of the existing orthography booklets (Byandala 1963, Kajolya 1990, LULANDA and CRC 2004, Namyalo et al. 2008), nor in any of the existing vocabulary lists or dictionaries (Korse 1999a, Gonza 2007, Nabirye 2009). Future writing guides and dictionaries (which ought to include extra-matter sections covering the orthography in depth) will thus do well to take cognizance of the problematic cases noted.

Reformulated: the transcriptions in the Lusoga corpus tapped into the most vibrant part of the language, namely the oral part, and enabled a realization of the unusual constructions that now need attention. The problem cases noted during the transcription exercise thus constitute the spelling issues to be addressed. Results from this study will patch gaps in the existing reference works (both writing guides and dictionaries), and will enable Lusoga primary teachers to improve their knowledge of writing Lusoga; a factor that is required to enable the judgement of the right and wrong usage of Lusoga in the Lusoga lesson exercises. As will be shown below, however, there are also cases where there is no right or wrong, at which point the lexicographer can resort to *proscriptive* lexicography (i.e. 'proposing' that a certain approach to writing Lusoga be used, rather than insisting on either 'prescribing' a norm, or merely 'describing' everything seen in the language; cf. Bergenholtz 2003, and Bergenholtz and Gouws 2010).

3. Writing problems noted

In this section each group of examples chosen for discussion represents a different type of problem noted during the transcription of the oral data. All examples are authentic, in that they have been taken from the transcriptions. They are also representative, in that many more similar cases have been encountered. They are called 'problematic' because none of the existing reference works includes guidance on how to reduce the spoken to the written form in these cases.

3.1 Full words

The existing orthographic specifications are for example challenged by speech forms that have not yet fully been conceptualized and passed on into the writing format. Moreover, results from the dictionary testing carried out in Nabirye (2008) showed that all respondents failed to demarcate word boundaries. The conceptualization and formalization of undocumented oral constructions are thus a challenge. Guidance is for instance needed in order to be able to distinguish between forms like 'today' and 'the day of today', as in (1a-b). In our tran-

scriptions, we differentiated between the two by writing the first conjunctively, the second disjunctively. The *leelo* 'today' in example (1b) is a variable formative and can be replaced with other words in the construction, such as 'Sunday' in (1c), whose lexicalization is independent from (1a).

- | | | | |
|---------|-------------------|------------------|--|
| (1) (a) | today | (o)lwaleelo | (full construction written conjunctively) |
| (b) | the day of today | (o)lwa leelo | (separate constructions written disjunctively) |
| (c) | the day of Sunday | (o)lwa Saabbiiti | (as previous, with capitalization) |

In the lexicalization of the Lusoga names of the months, as in (2a), we find that *Ogwokusatu* 'March' is a compounded form which does not require any further information to fulfil its function. It is a full, self-standing construction and a proper name which is capitalized. *Ogw'okusatu* 'a third' as in (2b), however, requires qualification to specify the subject of 'a third'. In other words, *omwezi ogw'okusatu* 'a third month' is not the same as 'March'. Or still, the subject of third as in for example *omulundi ogw'okusatu* 'a third time' is a dependant construction and requires contextual analysis to arrive at its full meaning.

- | | | | |
|---------|---------|--------------|---|
| (2) (a) | March | (O)gwokusatu | (full construction, capitalized, no abbreviation or complement) |
| (b) | a third | ogw'okusatu | (not capitalized, needs a noun complement) |

A distinction between structures such as (2a) and (2b) may be achieved through the use of capitalization and abbreviation. This is further evident in (3a) where *Gwakubili* 'February' is specified by the object (the underlined part). Without a specifier, as in (3b), we lack assurance that the fifth day also refers to 'Friday'. Here, it is an indefinite fifth day. In (3c), however, though the time sequence lacks an object, the context in which the term is used is sufficient to come to the conclusion that it is not referring to the beginning of just any sequence but specifically to the book of Genesis. Cases (3a) and (3c) are proper names which are by convention capitalized; cases such as (3b) are not.

- | | | |
|---------|---|---|
| (3) (a) | lit. month second time <u>2001</u> | <i>mwezi Gwakubili <u>nkumi ibili</u></i>
<i>ikumi na mulala</i> |
| (b) | lit. day fifth | <i>olunaku lwakutaanu</i> |
| (c) | lit. before we go in first | <i>nga tukaali gya mu Lubelyebelye</i> |

When forms such as (3c) are preceded by a possessive concord, as in (4a), the resulting forms may denote an independent concept. A similar example is shown in (4b). Even though such forms have not been entered or described in

any of the existing reference works, it is clear that they should and that they should be written as one full word, seeing that they refer to a single concept.

- (4) (a) *lit.* (something in cl. 7) of **first** > primary 1 *ekyolubelyebelye*
 (b) *lit.* (something in cl. 7) of **daylight** > lunch *ekyomusana*

The examples in (5) are cases of homography and reduplication. In our transcriptions (5b) is only a full construction when it is written conjunctively, while in (5c) the form can be repeated as many times as the emphasis allows, say as *boona boona boona boona*, but the meaning is not the same as in (5a). Distinguishing between the different lexical forms in (5) is dependent on establishing the function each form is intended to perform.

- (5) (a) also *boona* (single lexical construction)
 (b) all *boonaboona* (single lexical construction, written conjunctively)
 (c) all (+ emphasis) *boona boona ...* (compounded construction, written disjunctively)

Some constructions are difficult to specify because they sound twisted and are not only a puzzle in speech but also in writing. The writing problems are founded on similarities which make it difficult to establish boundaries in the entire construction. An example is given in (6c), which is best approached in successive steps, as in (6a) and (6b), leading to (6c).

- (6) (a) grandparents *badhaadha*
 (b) grandparents of the grandparent *badhaadha ba dhaadha*
 (c) grandparents of the grandparent's grandparent *badhaadha ba dhaadha ba dhaadha*

A thorough understanding of the Lusoga grammar is required to crack the puzzle in such constructions. The current absence of a proper grammatical description of Lusoga calls for extra caution during the transcription exercise to appreciate the different parts in the construction.

3.2 Compounds

Compounding is not problematic in speech where word boundary considerations are unconscious, but it may be problematic in writing where the process needs to be applied consciously. The specification of possible compounding procedures is thus required.

3.2.1 Compounding with prefixes

For example, the *Eiwanika* (i.e. the monolingual Lusoga dictionary, Nabirye 2009) defines *muna-* as a prefix used to link a subject to the object intended. A person of the journalist trade will for example be called *munamawulile*. At the time of compiling the *Eiwanika*, the usage of this prefix was not fully ascertained and seemed to crop up once or twice only, so that only the most dominant usages were given as examples. There are however new ways the same prefix is being exploited in the corpus, still serving the purpose defined but linking to a wider spectrum of objects not earlier conceived possible, as in (7a).

- | | | |
|---------|----------------------|--------------------|
| (7) (a) | someone from Bilumba | <i>Munabilumba</i> |
| (b) | people from Mayuge | <i>Banamayuge</i> |
| (c) | someone from Masese | <i>Munamasese</i> |

Such compounds clearly still need to be fully conceptualized as single concepts. In the e-mail and Facebook sections of the corpus, for example, (8i) is also found as (8ii-iii).³

- | | | | |
|-----|----------------|-------|------------------------|
| (8) | Ugandan | (i) | <i>(O)munauganda</i> |
| | also found as: | (ii) | <i>*(o)muna Uganda</i> |
| | also found as: | (iii) | <i>*(o)munaUganda</i> |

An analogous prefix is also being exploited in more elaborate ways than those primarily intended. The prefix *(o)mwise-* (sg.)/*(a)baise-* (pl.) usually denotes the belonging to a restricted context of a clan as in (9). The context restriction is however being relaxed and made to cover contexts other than the clan, as in (10).

- | | | |
|---------|----------------------|----------------------|
| (9) (a) | of the Menha clan | <i>Omwisemenha</i> |
| (b) | of the Igaga clan | <i>Abaiseigaga</i> |
| (c) | of the Mususwa clan | <i>Abaisemususwa</i> |
| (10) | supporter of the NRM | <i>mwiseNRM</i> |

This usage innovation puts both the prefix *(o)muna-* (sg.)/*(a)bana-* (pl.) and the prefix *(o)mwise-* (sg.)/*(a)baise-* (pl.) on the same footing with regard to their function. It is therefore essential to realize that all the cases noted in (7), (8i), (9) and (10) are correct forms, and where the derived compounds are proper names they should be capitalized, except for (10) where the subject referred to is not definite but could be any supporter of the political party NRM. Given all the examples represent a single concept in the construction, they should all be written conjunctively.

3.2.2 Compounding with independent word forms

Compounding in the formation of proper names provides a more familiar indication of the Lusoga compounding system. In (11a-c) the formatives on the left have independent denotative meanings, while they acquire connotative meanings and mutate into proper names conceptualized as single lexical units on the right. Being proper names they are capitalized, unlike in (11d) where the reference is indefinite.

- | | | | |
|----------|------------------------------------|---|---------------------------|
| (11) (a) | <i>mwene + iloboosi</i> | > | <i>Mweneiloboosi</i> |
| | <i>lit. owner + voice</i> | > | name of a specific person |
| (b) | <i>gaali + maka</i> | > | <i>Gaalimaka</i> |
| | <i>lit. it was + home</i> | > | name of a specific person |
| (c) | <i>mmele + ya + nkya</i> | > | <i>Mmeleyankya</i> |
| | <i>lit. food + of + morning</i> | > | name of a specific person |
| (d) | <i>ki + mpwitu</i> | > | <i>kimpwitu</i> |
| | <i>lit. someone + stubbornness</i> | > | stubborn person |

3.2.3 Compounding and noun gender

In the examples (7) through (11), both singular and plural forms were extracted from the corpus. When person and number come into play, rules for capitalization have to be laid down. If one for example refers to several people called *Gaalimaka*, it is suggested that the plural class prefix be written conjunctively with the compound, and the first letter only be capitalized, as in (12a). Similarly, a form like (12b) takes a word-initial capital letter, while (12c) doesn't.

- | | | | |
|----------|--------------------|--------------------------------|--------------------|
| (12) (a) | <i>Gaalimaka</i> | + number + capitalization | <i>Bagaalimaka</i> |
| (b) | <i>Omunauganda</i> | + number + capitalization | <i>Abanauganda</i> |
| (c) | <i>kimpwitu</i> | + number + null capitalization | <i>bakimpwitu</i> |

3.3 Multiple forms

The mapping between spoken and written forms is unfortunately not always unambiguous in Lusoga. There are two types, each the reverse of the other.

3.3.1 Different words for the same sound

Nabirye (2008) proposed a number of changes to the orthography. Some of these were also carried through to both Namyalo et al. (2008) and to the *Eiwanika*. In most cases only the 'safe' changes were implemented because the writing format for Lusoga was still new and only the cases that were considered to be really important were attended to. As a case in point, /j/ was intro-

duced for use with foreign words with that sound, as in *Janwali* 'January', *jahaadi* 'jihad', the place name *Jinja*, etc. With /j/ added to the alphabet, a frequent rendering like for example *Jinja* could now be accepted as part of the lexicon (in addition to the indigenous *Idhindha*).

While /j/ rather than /dh/ was applied throughout the lexicon, /j/ rather than /gy/ was regrettably not applied consistently, which is especially unfortunate as /j/ and /gy/ are homophones. Existing Lusoga texts as well as early versus more recent transcriptions therefore contain both spellings seen in (13a), as well as both seen in (13b).

- | | | | | |
|----------|--|------|----------------|------|
| (13) (a) | (woman who has made a pilgrimage to Mecca) | (i) | <i>Hagyati</i> | /gy/ |
| | | (ii) | <i>Hajati</i> | /j/ |
| (b) | (pilgrimage to Mecca) | (i) | <i>higya</i> | /gy/ |
| | | (ii) | <i>hija</i> | /j/ |

In the *Eiwanika*, *Hagyati* was unfortunately entered rather than *Hajati*, yet at the same time *hija* was (correctly) entered and not *higya*. The existing reference works thus give confusing signals, with multiple spellings for the same words as a result, and this is reflected in the corpus, including the transcribed parts.

Another case of multiple spellings for the same words is the result of a type-writer limitation to representing the velar nasal on the keyboard. The solution was to use /ng/ rather than /ŋ/, a 'solution' which became a tradition with time, even when the limitation had ceased to exist. Nabirye (2008) argued for the reintroduction of /ŋ/ in the orthography, and it was introduced in Namyalo et al. (2008). It was also used in the *Eiwanika*. However, in the written sections of the corpus as well as the early transcriptions, the velar nasal is represented with /ng/, rather than /ŋ/, resulting in multiple forms such as those seen in (14).

- | | | | | |
|----------|----------------|------|----------------|--------|
| (14) (a) | (a place name) | (i) | <i>Ngando</i> | / ng / |
| | | (ii) | <i>Nyando</i> | / ŋ / |
| (b) | relatives | (i) | <i>enganda</i> | / ng / |
| | | (ii) | <i>enyanda</i> | / ŋ / |

3.3.2 Different sounds for the same word

There are sounds which are dying out and others are taking their place, resulting in spelling variations as shown in (15a) and (15b).

- | | | | | |
|----------|----------------------|------|-----------------|-------------|
| (15) (a) | Jinja | (i) | <i>Idinda</i> | |
| | | (ii) | <i>Idhindha</i> | [dying out] |
| (b) | nation/country/tribe | (i) | <i>eiwanga</i> | |
| | | (ii) | <i>eighanga</i> | [dying out] |

Although some sounds may be on the wane, as indicated in (15), a single person may still use the various alternatives, even in the same sentence. It is not so that one pronunciation is 'better' Lusoga than the other(s), or that one is 'better' from a linguistic point of view.

3.4 Loanwords

In the absence of any operational language regulatory body for Lusoga, it often seems as if the language spirals out of control, hampering all attempts to streamline and to formalize the writing. When transcribing oral data, foreign adoptions are especially ambiguous to deal with. The linguist's intuition tells her or him to rid the language of these 'intrusions', but then, they are found — often in high numbers — so have clearly become part of the language.

3.4.1 Borrowings from neighbouring languages

The current Lusoga orthography strictly specifies that the combination /ny/ is not part of the language; rather, the combination /nh/ is used in Lusoga where neighbouring languages like Luganda have /ny/. However, Luganda forms such as the one shown in (16a) are used so often in Lusoga — more often even than their Lusoga counterparts — that when transcribing oral data one cannot simply ignore or 'correct' these forms. Similarly, in (16b), the prefix is dropped in the spoken language, rendering the Luganda version rather than Lusoga.

(16)		Lusoga	Luganda
	(a) madam	<i>inhabo</i>	<i>nnyabo</i>
	(b) sir	<i>isebo</i>	<i>ssebo</i>

More confusing is the use of foreign words to refer to place names, even when referring to places in Busoga. In (17i) the common Lusoga spellings of two place names are shown, as well as the reasons why these spellings are problematic from a Lusoga point of view. Using a Lusoga(ized) spelling, (17ii) is obtained. A brief etymology is also provided in (17ii).

(17) (i)	<u>Wanyange</u>	/ny/ does not exist in Lusoga
	<u>Makerere</u>	/r/ is contested in the Lusoga alphabet
(ii)	<u>Wanhange</u>	named after a bird known as <i>ennhange</i>
	<u>Makeelele</u>	name derived from a Luganda idiom ' <i>amagezi makeelele</i> '

When transcribing, it is not simply a matter of choosing either series, say only the commonly-found renderings as in (17i), or the Lusoga(ized) versions as in (17ii). When carefully listening to how these words are actually pronounced, the underlined forms in (17) are used.

3.4.2 Borrowings from the religious sphere

With religious indoctrination not only came new concepts (and a new faith), but also new, previously unseen, sound combinations. Tussling denominations also gave their own twist to word-final vowels. Both these aspects are illustrated in (18a-b).

(18) (a) Christ	(i) Catholics	<i>Kristu</i>	but not	* <i>Kulisitu</i>
	(ii) Protestants	<i>Kristo</i>	but not	* <i>Kulisito</i>
(b) sacrament	(i) Catholics	<i>sakramentu</i>	also as	<i>saakalamentu</i>
	(ii) Protestants	<i>sakramento</i>	also as	<i>saakalamento</i>
(c) democracy		<i>demokrasiya</i>	also as	<i>dimokulasiya</i>

Bantu words normally show a CVCV structure, but forms such as *Kristo/Kristu* are now so well entrenched that they are not only written like that, but also pronounced as they are written (and thus not as **Kulisito*/**Kulisitu*, which would have been in accordance with a CVCV structure). Religious publications 'faithfully' stick to using the combination /kr/ throughout, for all loanwords with this foreign sound combination, whether religious (e.g. *sakramento*, as in (18b)), or not (e.g. *demokrasiya*, as in (18c)). When these same words are used in a non-religious context, and/or by non-believers, speakers and writers often resort to variants which do adhere to the CVCV structure however (thus e.g. *saakalamento* and *dimokulasiya* here). An accurate transcription will reflect these differences.

3.4.3 Lexicalization of borrowed abbreviations and acronyms

The wish by speakers of Lusoga to adhere to a CVCV structure is often so strong (i.e. intuitive), that when abbreviations are borrowed wholesale, as in 'FM', some pronounce it as /fa ma/, rather than /ef em/. At that point the transcriber has at least three options, as shown in (19).

(19) Frequency Mode	(i) <i>FM</i>	/ef em/
	(ii) <i>Ef Em</i>	/ef em/
	(iii) <i>Fa Ma</i>	/fa ma/

Abbreviations and acronyms are not only cited as they are formally known but also made to carry characteristics of person and number as in (20).

(20) (a) Resident District Commissioners	<i>RDC + number</i>	<i>baRDC</i>
(b) District Security Officers	<i>DISO + number</i>	<i>baDISO</i>
(c) District Police Commanders	<i>DPC + number</i>	<i>baDPC</i>
(d) District Health Officers	<i>DHO + number</i>	<i>baDHO</i>

The new abbreviated forms represent a single concept of indefinite office occupants, hence requiring no word-initial capitalization, but they should be written conjunctively. Although the characteristics marked on plurals (i.e. the use of prefixes) are also plausible for singular, the singular form doesn't require additional lexicalization processes to infer the meanings.

3.4.4 Borrowings from English

The hardest borrowings a transcriber has to contend with are not those from neighbouring languages (which at least have a Bantu structure), nor those from the religious sphere (as they are mostly limited to terms from a restricted domain), nor the abbreviations and acronyms (where the spelling issues are minor), but the wholesale borrowings of concepts and words from English that are literally 'dropped' into the language, and on which the full Bantu morphological and morphophonological apparatus is unleashed (including possible phonologization, marking of person and number, prefixation, the addition of verbal extensions, enclitization, etc.).

Example (21) shows a trivial case where some writers will adapt the foreign word to the structure of Lusoga (21i), while others will keep the English spelling intact yet still pluralize it according to the Lusoga morphology rather than to use the English plural suffix *-s* (21ii).

- | | | | | |
|------|---------|-----|-----------------|-----------------------|
| (21) | lawyers | i) | <i>balooya</i> | number + phonologized |
| | | ii) | <i>balawyer</i> | number + English root |

The fully phonologized version (*balooya*) and the version in which the English root is merely prefixed with the cl. 2 plural class prefix (*balawyer*), have the exact same pronunciation. When transcribing Lusoga recordings it is advisable to stick to one spelling only, so either (21i) or (21ii), with our preference going to (21i).

Depending on the noun class a particular borrowed word ends up in, both singular and plural forms may of course also look the same; compare (22a-c) with (22d-e) in this respect.

- | | | | | |
|------|-----|---------------|---|--------------|
| (22) | (a) | hello(s) | <i>alo</i> (sg.)/ <i>baalo</i> (pl.) | gender 1a/2a |
| | (b) | district(s) | <i>district</i> (sg.)/ <i>madistrict</i> (pl.) | gender 9/6 |
| | (c) | government(s) | <i>gavumenti</i> (sg.)/ <i>dhigavumenti</i> (pl.) | gender 9/10 |
| | (d) | bye(s) | <i>bbaayi</i> (both sg. and pl.) | gender 9/10 |
| | (e) | septrine(s) | <i>septrine</i> (both sg. and pl.) | gender 9/10 |

Furthermore, any type of prefixes (thus also other than noun class prefixes) may precede the borrowed material, as in (23).

- | | | | |
|------|--------------------------------------|--------------------------------|---------------------|
| (23) | pertaining to the post of a Minister | <i>obwa-</i> + <i>Minister</i> | <i>obwaminister</i> |
|------|--------------------------------------|--------------------------------|---------------------|

Usages of such constructions have a considerable occurrence in the oral part of the corpus. A specification of how to address these forms in writing is thus warranted. The argument of full word conceptualization as a single lexical form lends itself easily here, while the non-capitalization seen is in harmony with similar cases in the lexicon.

Once a word is borrowed into Lusoga, it may also be moved around the noun classes, and for example be made to take on degree assessments, such as the diminutives seen in (24).

- (24) (a) a bit of airtime *kaairtime* gender 12/14
 (b) short programs *buprograms* gender 12/14

In the process, speakers of Lusoga may combine two grammatical systems, as in (24b), where one notices a double plural marking: one from Lusoga, i.e. the cl. 14 plural noun prefix *bu-*, and one from English, i.e. the plural suffix *-s*. A truthful transcription will reflect such idiosyncrasies.

Words from word classes other than nouns in English may also end up being nominalised in Lusoga, as seen in (25).

- (25) (a) small plus *kaplus* gender 12/14
 (b) *extensives *ebyo biextensive ebyo* gender 7/8

(25a) shows an example of the nominalization of a conjunction/adjective, while (25b) shows an example of the nominalization of an adjective/adverb. The task of the transcriber is again to record all such instances, leaving their analysis for a later phase.

Lastly, English verbs too may be borrowed wholesale. They are typically adopted in their infinitive form, and then 'Lusogaized' by means of the standard verbal morphology of Lusoga. As such, Lusogaized verbs may take on verbal extensions, as the applicative in (26a), or the perfective forms reflected on the verb ending vowels (including sound changes) in (26b) and (26c), they may include reflexive markers as in (26d), they may mark aspect as in (26e) and (26f), they may accommodate a subject concord and an enclitic as in (26g), etc. Cases of elision, which are quite dominant in the Lusoga corpus, also occur with the Lusogaized word forms, as seen in (26h).

- (26) (a) and planning for her/him *n'okumuplaningila*
 (b) they pronounced *bapronouncinze*
 (c) I did not monitor *tyamonitorinze*
 (d) it comes *neecominga*
 (e) they still lack *bakaalackinga*
 (f) it/which will always subsidize *enaasubsidizinga*
 (g) she/he categorized them *yaadhicategorizingamu*
 (h) I am petitioning *ndiipetitioninga*

Language purists may prefer to ignore such borrowings, but given their increasingly frequent use, any description of the current language cannot ignore them.

3.5 Formality levels

Spoken language is by and large less formal than written language. This difference is also apparent in the transcriptions.

3.5.1 The use of the apostrophe

As it turns out, the grammatical function of the use of the apostrophe to indicate shortening in Lusoga has remained unspecified in the existing reference works on Lusoga. For example, (27i) and (27ii) are cases that should not be mistaken for (2a) and (2b), as they do not result in any semantic change but should be considered as a distinction between formal and informal usages.

- (27) as a person (i) *nga omuntu* (formal representation)
(ii) *ng'omuntu* (informal representation)

3.5.2 The reduplication of noun and verb stems

Stem reduplication, too, has remained unspecified in the existing reference works on Lusoga. Various examples have been extracted from the corpus and are shown in (28), where the parts between brackets are the reduplicated parts. From these examples we see that stem reduplication typically occurs twice or three times, and with both noun and verb stems.

- (28) (a) numerous small wounds *obubwa(bwa)(bwa)*
(b) somewhere in the middle *wagati(gati)*
(c) somehow of black colour *obwilugavu(ilugavu)*
(d) I have redistributed them *mbakutwile(kutwile)mu*
(e) they just drop in *basuula(suula)mu*
(f) small amounts (of liquids) *matono(tono)*
(g) they do not repeatedly fall sick *tibalwala(lwala)*
(h) numerous small envelopes *obubbaasa(bbaasa)*
(i) *lit.* small eye eye *kaliiso(liiso)*

Carefully listening to the recordings moreover reveals that there are environments where the vowel between the reduplicated parts is lengthened. See for instance (29c-f) vs. (29a-b).

- (29) (a) *lit.* we close close *tuwumba(wumba)*
(b) *lit.* to frighten frighten them *okubatiisa(tiisa)*

(c) <i>lit.</i> to try try	<i>okuwadaa(wada)</i>
(d) <i>lit.</i> someone trick tricks	<i>abuzaa(buza)</i>
(e) <i>lit.</i> to look look after	<i>okubudaa(buda)</i>
(f) <i>lit.</i> to suffer suffer	<i>okubonaa(bona)</i>

3.5.3 The use of enclitics

Enclitics slightly modify the meanings of expressions, as in (30). They can be repeated and/or combined two or three times, as in (31), respectively (32). The recordings also reveal that any additional enclitic on a word brings about the doubling of the vowel before the onset of the next enclitic. In some cases (underlined in (31) and (32)) the verb ending vowel is equally doubled before the onset of the enclitic(s), a case similar to that noted in (29c-f). In (33a) the enclitic *-ku* is further preceded by the conjunctive *na*, and in (33b) the enclitic *-yo* is preceded by the aspect feature *nga*.

(30) (a) we also have there	<i>twalinayo</i>
(b) a bit up	<i>wanguluku</i>
(c) a bit down	<i>wansiku</i>
(d) a bit near/nearby	<i>kumpiku</i>
(31) (a) <i>lit.</i> to bathe a bit a bit	<i>okunaabakuu(ku)</i>
(b) <i>lit.</i> we have some bit bit	<i>tubaayookuu(ku)</i>
(c) <i>lit.</i> to converse a bit bit	<i>okuwaya<u>amuu</u>(ku)</i>
(d) <i>lit.</i> you are of a bit bit number	<i>muwelaawelamuu(ku)</i>
(e) <i>lit.</i> s/he still has a remaining bit bit	<i>afiisiizaa<u>amuu</u>(ku)</i>
(f) <i>lit.</i> let us make it a bit bit smoother	<i>tutandaaze<u>ewoo</u>(ku)</i>
(32) (a) <i>lit.</i> we respond to some bit bit bit	<i>twilangamuu(kuu)(ku)</i>
(b) <i>lit.</i> you prevent some bit bit bit	<i>waaziyiza<u>akuu</u>(kuu)(ku)</i>
(33) (a) I enable them to get at least a bit	<i>mbafunhise(na)(ku)</i>
(b) let's proceed for the moment	<i>tweyongele(nga)(yo)</i>

Forms with multiple enclitics and imbedded formatives only occur in the oral part of the corpus, which leads to the assumption that the reduplication of enclitics is more regulated in the written language. In other words, it seems as if the use of multiple enclitics is exploited in the spoken language to bring about meanings that best serve informal contexts.

3.6 Concords

A core feature of all Bantu languages is the existence of a noun class system with linked concordial agreement. This basically means that nouns, verbs, ad-

jectives, and many other parts of speech, are all 'in harmony' with one another on phrase level. Thus, if a noun belongs to cl. 4, then the adjective modifying it will use an adjective concord (AC) of cl. 4, the verb referring to it will use a subject concord (SC) of cl. 4, etc. Concordial agreement is thought to be inviolable. In the oral part of the corpus, however, this system is occasionally defied, as may be seen from the selection shown in (34).

(34)

#	Phrase	Gender	Concord mismatch
(a)	<i>omwana gw'endiga</i> child of a lamb	1/2	for cl. 1, PC cl. 3 <i>gwe-</i> is used not cl. 1 <i>wa-</i>
(b)	<i>Mbeebaziiza emilimo dhe mukoze</i> I thank you for the work you have done	3/4	for cl. 4, RC cl. 10 <i>dhe-</i> is used not cl. 4 <i>gye-</i>
(c)	<i>emilimo edhindi</i> some other jobs/work	3/4	for cl. 4, SC cl. 10 <i>dhi-</i> is used not cl. 4 <i>gi-</i>
(d)	<i>n'emiti Katonda</i> <i>n'eyadhitonda</i> <i>lit.</i> even trees God created them	3/4	for cl. 4, SC cl. 10 <i>dhi-</i> is used not cl. 4 <i>gi-</i>
(e)	<i>emikolo dhaatebwa</i> celebrations were held	3/4	for cl. 4, SC cl. 10 <i>dhi-</i> is used not cl. 4 <i>gi-</i>
(f)	<i>emyoyo dhaife</i> our souls	3/4	for cl. 4, PC cl. 10 <i>dhi-</i> is used not cl. 4 <i>gya-</i>
(g)	<i>mwino akuwa gy'owa</i> <i>lit.</i> you give to the one who gives you	1/2	for cl. 1, RC cl. 4 <i>gye-</i> is used not cl. 1 <i>gwe-</i>

In (34a), for example, *omwana* 'child' is a noun in cl. 1, with its plural *abaana* 'children' in cl. 2, thus together gender 1/2. In the phrase, the possessive concord (PC) that follows should thus also be the one from cl. 1, yet here the PC from cl. 3 is used. In (34b), *emilimo* 'jobs' is a noun in cl. 4, yet the relative concord (RC) that follows and which should be in accordance with it is the one from cl. 10 rather than cl. 4. And so on for the other examples.

It is of course not so that the entire concordial agreement system is broken when mother-tongue speakers speak, but the instances shown in (34) are not isolated examples: similar cases are found in the speech of different speakers, in recordings made across different regions. Actually, there were likely more instances than those captured which were 'corrected' in error.

Thus is the burden of the transcriber: While reducing the spoken to the written form of a language, all existing orthography rules need to be applied, mapping undocumented writing problems onto it, so as to be as consistent as possible. However, consistency does not imply that one can intervene to 'correct' what is *thought* to be wrong. Imposing too rigid a structure may even create more structure than there actually is in a language.

4. Implications for Lexicography

AFRILEX members were introduced to the notion of 'proscriptive lexicography' during the keynote address of Henning Bergenholtz at the *7th International AFRILEX Conference*, in July 2002. Although Gregory James questioned the choice of the term itself during question time (and suggested 'praeterscriptive lexicography', cf. Bergenholtz 2003: 80), and although similar doubts have periodically been raised over the years, amongst others during a workshop on "Proscription, Prescription and Description" at the *14th International AFRILEX Conference*, in July 2009, the concept and the need for it are not in dispute.

The most powerful single example of what proscriptive lexicography entails has been provided by Bergenholtz himself. Consider the following entry in a hypothetical Danish LSP dictionary (Bergenholtz 2003: 78):

kraftvarmeværk noun <et; -et, -er, -erne>

Other spellings with hyphens are possible: *kraft-varmeværk* or *kraft-varme-værk*.

They are not recommended. They are quite rare in language use, e.g. by special field experts. The Danish Language Council allows only the spelling with two hyphens: *kraft-varme-værk*.

Note, here, that even though the Danish Language Council allows the use of one and only one spelling for this word, namely *kraft-varme-værk* (i.e. with two hyphens), the lexicographer explicitly states that this form is "not recommended" (in lines 2 and 3 of the entry). Rather, the lexicographer recommends the spelling without any hyphens, thus *kraftvarmeværk*, as seen in the lemma sign, the argument being:

Contrary to prescriptive dictionary articles, [the dictionary user] is advised about language use similar to the normal language use in society.

Bergenholtz (2003: 78)

Indeed, both a linguistic survey and a Google search revealed an overwhelming preference (respectively 85.0% and 98.3%) for the form without any hyphens (Bergenholtz 2003: 69). In proscriptive lexicography, a *recommendation* is thus given, and that recommendation is based on true language usage, regardless of norms and regulations.

Contrast this with prescriptive and descriptive lexicography. Using a prescriptive approach, *kraft-varme-værk* would have been the lemma sign, with the user being told that all other spellings are disallowed. Using a descriptive approach, the user would have been provided with all the facts, and all the spelling options would also have been entered as lemma signs and cross-referred in the dictionary.

Proscriptive lexicography thus takes both data and actual language usage extremely seriously. With regard to the data available to the lexicographer who wants to pursue proscription, Bergenholtz distinguishes the following possibilities:

- (a) introspection,

- (b) analysis of a linguistic survey,
 - (c) involvement of descriptions in existing dictionaries, grammars, monographs, articles, etc.,
 - (d) analysis of a number of examples which have been randomly chosen from random texts (corresponding with the practice of dictionary making before the age of computers),
 - (e) analysis of a specifically constructed text corpus, and
 - (f) analysis of usage found in texts in the examined language in all available web-sites on the Internet.
- Bergenholtz (2003: 77)

During the compilation of the *Eiwanika*, possibilities (a) through (d) were used (cf. Nabirye 2008, 2009a, and Nabirye and De Schryver 2010), and since then possibilities (e) and (f) have been added to the mix (cf. De Schryver and Nabirye 2010, as well as Table 1 in the present article). Lusoga lexicography is thus in the unique position where so-called 'total proscription' is not only a possibility, but also where it is actually put in practice.

This unique position could, in theory, lead to the perfect dictionary, or rather, "the perfect *set* of dictionaries". Indeed, given a certain type of problem, a certain type of user in a certain type of user situation (cf. Bergenholtz 2003: 68), will fair best when consulting a *monofunctional* dictionary:

[A] monofunctional dictionary contain[s] as much data as necessary but as little as possible to guarantee a rapid access that is not impeded by unnecessary data or that leads to information stress or even information death. [... However:] The default approach of many lexicographers is that they are producing a polyfunctional dictionary which should assist the user in satisfying at least a cognitive function, a text reception function and a text production function.
Bergenholtz and Gouws (2010: 41, 43)

With regard to the latter two dictionary functions, Bergenholtz and Gouws (2010) have made a case that (text) reception dictionaries need to be *descriptive*, but that for (text) production, *description* is not viable if more than one variant prevails, that *prescription* could be viable, and that *proscription* is likely the best option.

While all of this makes perfect sense *in abstractum*, it is instructive to re-analyze the various transcription problems enumerated in section 3 above, and to see how they could be handled in a database from which a variety of monofunctional dictionaries could subsequently be extracted.⁴

The first type is the easy one: A monofunctional reception dictionary is ideally descriptive throughout. Each of the grouped examples in (1) through (34) can indeed be described in utmost detail in a dictionary database: this form is correctly spelled and means X; this form has a more/lesser frequently used variant X; this form is used in domain X, for domain Y use ...; don't confuse

this form with X; note the capital letter in this form; note the sound change in this form; this form can be repeated, with the meaning/function X; this form is constructed as follows ...; this form is an obsolete spelling for X; this form is derived from language X; this form is actually language X, but used; this form is spelled wrongly (according to the official terminology), the correct one is X; the plural of this form is X; the diminutive of this form is X; although this is not according to the grammar, it is used; etc. At face value, the lexicographer did not introduce any value judgements here; in truth, however, both implicit prescription and implicit proscription are at work, amongst others because the most recent and now standard orthography for Lusoga was proposed by the same lexicographer.

For the second type, namely the monofunctional production dictionary, a hybrid approach has been put forward. While proscription is seen as the best option, this needs to be understood as a situation in which the lexicographer recommends throughout. At times, these recommendations may overlap with what is actually prescribed, and where there are no variants, one is in effect describing. Or in the words of Bergenholtz and Gouws:

In trying to satisfy a text production function, the lexicographer should pay careful attention to the application of an approach characterised by either description, prescription or proscription or a hybrid application in which more than one of these approaches can be combined. This decision should not be made in a haphazard way.

Bergenholtz and Gouws (2010: 47)

While our database contained 'mere' descriptions in the slots meant for the creation of a reception dictionary, the situation is truly hybrid in the database slots meant for the creation of a production dictionary. The latter is illustrated visually in the Addendum to this article.

By way of conclusion, then, the metalexigraphic explorations by Bergenholtz and his colleagues have found an application — and indeed confirmation — in the emerging field of dictionary making for Lusoga. Reducing a language to writing before it has even been standardized is a daunting undertaking, but knowing that after moving from oral data to transcription, there is a theoretical framework consisting of a hybrid use of lexicographic prescription, description and proscription, to subsequently move to a dictionary database and finally to a set of monofunctional dictionaries, is sufficient consolation to carry on with the work.

Endnotes

1. For the definition of 'organic corpus' (and the related issues of 'representativeness' and 'balance'), see Atkins et al. (1992: 1, 4, 6).
2. Namely the first author of this article, who is also a father-tongue speaker of the language.

3. Similarly, *Omusoga* 'Musoga' is also found as **omusoga* and **omuSoga* in the e-mail and Facebook sections of the corpus.
4. For the concept of "one database, many dictionaries", see De Schryver and Joffe (2005).

References

- Atkins, B.T.S., J. Clear and N. Ostler. 1992. Corpus Design Criteria. *Journal of Literary and Linguistic Computing* 7(1): 1-16.
- Bergenholtz, H. 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Bergenholtz, H. and R.H. Gouws. 2010. A Functional Approach to the Choice between Descriptive, Prescriptive and Proscriptive Lexicography. *Lexikos* 20: 26-51.
- Byandala, G.I. 1963. *The Lusoga Orthography*. Iganga.
- De Schryver, G.-M. and D. Joffe. 2005. One Database, Many Dictionaries — Varying Co(n)text with the Dictionary Application TshwaneLex. Ooi, V.B.Y., A. Pakir, I. Talib, L. Tan, P.K.W. Tan and Y.Y. Tan (Eds.). 2005. *Words in Asian Cultural Contexts. Proceedings of the 4th Asialex Conference, 1-3 June 2005, M Hotel, Singapore*: 54-59. Singapore: Department of English Language and Literature & Asia Research Institute, National University of Singapore.
- De Schryver, G.-M. and M. Nabirye. 2010. A Quantitative Analysis of the Morphology, Morphophonology and Semantic Import of the Lusoga noun. *Africana Linguistica* 16: 97-153.
- Gonza, R.K. 2007. *Lusoga-English Dictionary and English-Lusoga Dictionary*. (Revised edition of P. Korse's (1999a) dictionary.) Kampala: MK Publishers.
- Kajolya, J.B.N. 1990. *The Lusoga Orthography*. (Revised edition of G.I. Byandala's (1963) Lusoga orthography.) Jinja: Lusoga Ecumenical Committee.
- Korse, P. 1999a. *Dictionary Lusoga-English/English-Lusoga*. Jinja: Cultural Research Centre.
- LULANDA and CRC. 2004. *Empandiika y'Olulimi Olusoga Enkalamu/Standard Lusoga Orthography*. Jinja: Lusoga Language Authority.
- Nabirye, M. 2008. *Compilation of the Monolingual Lusoga Dictionary*. Unpublished MA dissertation. Kampala: Makerere University.
- Nabirye, M. 2009. *Eiwanika ly'Olusoga. Eiwanika ly'aboogezi b'Olusoga n'abo abenda okwega Olusoga* [A Dictionary of Lusoga. For speakers of Lusoga, and for those who would like to learn Lusoga]. Kampala: Menha Publishers.
- Nabirye, M. 2009a. Compiling the First Monolingual Lusoga Dictionary. *Lexikos* 19: 177-196.
- Nabirye, M. and G.-M. de Schryver. 2010. The Monolingual Lusoga Dictionary Faced with Demands from a New User Category. *Lexikos* 20: 326-350.
- Namyalo, S., L. Walusimbi, G. Bukenya, M.W. Masakala, M. Nabirye and F. Kiingi. 2008. *A Unified Standard Orthography of Eastern Interlacustrine Bantu Languages*. Monograph Series 68. Cape Town: The Centre for Advanced Studies of African Society.
- NCDC. 2006. *THEMA. The newsletter of the Thematic Primary Curriculum. Issue 1. August 2006*. Kampala: National Curriculum Development Centre.
- NCDC. 2006a. *THEMA News Letter. Issue 2. December 2006*. Kampala: National Curriculum Development Centre.

Addendum: Prescription (Pre), Description (De) and Proscription (Pro) in a Monofunctional Production Dictionary for Lusoga
 [with ✓ = the inclusion of the lemma sign; (✓) = the inclusion of the lemma sign, but with a treatment limited to a cross-reference to the main/correct variant only]

N #	English	Lusoga	Pre	De	Pro
1 a	today	(o)lwaleelo	✓		
b	the day of today	(o)lwa leelo	✓		
c	the day of Sunday	(o)lwa Saabbiiti	✓		
2 a	March	(O)gwokusatu	✓		
b	a third	ogw'okusatu	✓		
3 a	lit. month second time 2001	mwezi <i>Gwakubili nkumi ibili ikumi na mulala</i>	✓		
b	lit. day fifth	olunaku <i>lwakutaanu</i>	✓		
c	lit. before we go in first	nga tukaali <i>gya mu Lubelyebelye</i>	✓		
4 a	lit. (something in cl. 7) of first > primary 1	ekyolubelyebelye	✓		
b	lit. (something in cl. 7) of daylight > lunch	ekyomusana	✓		
5 a	also	boona	✓		
b	all	boonaboona	✓		
c	all (+ emphasis)	boona boona ...	✓		
6 a	grandparents	badhaadha	✓		
b	grandparents of the grandparent	badhaadha ba dhaadha	✓		
c	grandparents of the grandparent's grandparent	badhaadha ba dhaadha ba dhaadha	✓		
7 a	someone from Bilumba	Munabilumba	✓		
b	people from Mayuge	Banamayuge	✓		
c	someone from Masese	Munamasese	✓		
8 i	Ugandan	(O)munauganda	✓		
ii	also found as:	*(o)muna Uganda	(✓)		
iii	also found as:	*(o)munaUganda	(✓)		
9 a	of the Menha clan	Omwisemenha	✓		
b	of the Igaga clan	Abaiseigaga	✓		
c	of the Mususwa clan	Abaisemususwa	✓		
10	supporter of the NRM	mwiseNRM	✓		
11 a	lit. owner + voice	Mweneiloboosi	✓		
b	lit. it was + home	Gaalimaka	✓		
c	lit. food + of + morning	Mmeleyankya	✓		
d	lit. someone + stubbornness	kimpwitu	✓		
12 a	Gaalimaka + number + capitalization	Bagaalimaka	✓		
b	Omunauganda + number + capitalization	Abanauganda	✓		
c	kimpwitu + number + null capitalization	bakimpwitu	✓		
13 a_i	(woman who has made a pilgrimage to Mecca)	Hagyati			(✓)
a_ii		Hajati			✓
b_i	(pilgrimage to Mecca)	higya			(✓)
b_ii		hija			✓

14 a_i (a place name)	<i>Ngando</i>	(✓)		
a_ii	<i>Nyando</i>	✓		
b_i relatives	<i>enganda</i>	(✓)		
b_ii	<i>enjanda</i>	✓		
15 a_i Jinja	<i>Idinda</i>		✓	
a_ii [dying out]	<i>Idhindha</i>		(✓)	
b_i nation/country/tribe	<i>eitwanga</i>		✓	
b_ii [dying out]	<i>eighanga</i>		(✓)	
16 a madam	<i>inhabo > nnyabo</i>			✓
b sir	<i>isebo > ssebo</i>			✓
17 <i>Wanyange</i>	<i>Wanhange</i>			✓
<i>Makerere</i>	<i>Makeelele</i>			✓
18 a_i Christ (Catholics)	<i>Kristu / *Kulisitu</i>		✓/(✓)	
a_ii Christ (Protestants)	<i>Kristo / *Kulisito</i>		✓/(✓)	
b_i sacrament (Catholics)	<i>sakramentu / saakalamentu</i>		✓/(✓)	
b_ii sacrament (Protestants)	<i>sakramento / saakalamento</i>		✓/(✓)	
c democracy	<i>demokrasiya / dimokulasiya</i>		✓/(✓)	
19 i Frequency Mode	<i>FM</i>		(✓)	
ii	<i>Ef Em</i>		(✓)	
iii	<i>Fa Ma</i>		✓	
20 a RDC + number	<i>baRDC</i>	✓		
b DISO + number	<i>baDISO</i>	✓		
c DPC + number	<i>baDPC</i>	✓		
d DHO + number	<i>baDHO</i>	✓		
21 i lawyers (number + phonologized)	<i>balooya</i>	✓		
ii lawyers (number + English root)	<i>balawyer</i>	✓		
22 a hello(s)	<i>alo (sg.) / baalo (pl.)</i>		✓	
b district(s)	<i>district (sg.) / madistrict (pl.)</i>		✓	
c government(s)	<i>gavumenti (sg.) / dhigavumenti (pl.)</i>		✓	
d bye(s)	<i>bbaayi (both sg. and pl.)</i>		✓	
e septrine(s)	<i>septrine (both sg. and pl.)</i>		✓	
23 pertaining to the post of a Minister	<i>obwaminister</i>		✓	
24 a a bit of airtime	<i>kaairtime</i>		✓	
b short programs	<i>buprograms</i>		✓	
25 a small plus	<i>kaplus</i>		✓	
b *extensives	<i>ebyo biextensive ebyo</i>		✓	
26 a and planning for her/him	<i>n'okumuplanningila</i>			✓
b they pronounced	<i>bapronouncinze</i>			✓
c I did not monitor	<i>tyamonitorinze</i>			✓
d it comes	<i>neecominga</i>			✓
e they still lack	<i>bakaalackinga</i>			✓
f it/which will always subsidize	<i>enaasubsidizinga</i>			✓
g she/he categorized them	<i>yaadhicategorizingamu</i>			✓
h I am petitioning	<i>ndiipetitioninga</i>			✓
27 i as a person (formal representation)	<i>nga omuntu</i>	✓		
ii as a person (informal representation)	<i>ng'omuntu</i>	✓		
28 a numerous small wounds	<i>obubwa(bwa)(bwa)</i>		✓	
b somewhere in the middle	<i>wagati(gati)</i>		✓	
c somehow of black colour	<i>obwilugavu(ilugavu)</i>		✓	
d I have redistributed them	<i>mbakutwile(kutwile)mu</i>		✓	
e they just drop in	<i>basuula(suula)mu</i>		✓	
f small amounts (of liquids)	<i>matono(tono)</i>		✓	

	g	they do not repeatedly fall sick	<i>tibalwala(lwala)</i>	✓	
	h	numerous small envelopes	<i>obubbaasa(bbaasa)</i>	✓	
	i	<i>lit.</i> small eye eye	<i>kaliiso(liiso)</i>	✓	
29	a	<i>lit.</i> we close close	<i>tuwumba(wumba)</i>	✓	
	b	<i>lit.</i> to frighten frighten them	<i>okubatiisa(tiisa)</i>	✓	
	c	<i>lit.</i> to try try	<i>okuwadaa(wada)</i>	✓	
	d	<i>lit.</i> someone trick tricks	<i>abuzaa(buza)</i>	✓	
	e	<i>lit.</i> to look look after	<i>okubudaa(buda)</i>	✓	
	f	<i>lit.</i> to suffer suffer	<i>okubonaa(bona)</i>	✓	
30	a	we also have there	<i>twalinayo</i>	✓	
	b	a bit up	<i>wangulu^{ku}</i>	✓	
	c	a bit down	<i>wansiku</i>	✓	
	d	a bit near / nearby	<i>kumpiku</i>	✓	
31	a	<i>lit.</i> to bathe a bit a bit	<i>okunaabakuu(ku)</i>	✓	
	b	<i>lit.</i> we have some bit bit	<i>tubaayo^{ku}kuu(ku)</i>	✓	
	c	<i>lit.</i> to converse a bit bit	<i>okuwaya^{muu}kuu(ku)</i>	✓	
	d	<i>lit.</i> you are of a bit bit number	<i>muwelaawelamuu(ku)</i>	✓	
	e	<i>lit.</i> s/he still has a remaining bit bit	<i>afiisiza^{muu}kuu(ku)</i>	✓	
	f	<i>lit.</i> let us make it a bit bit smoother	<i>tutandaaze^{woo}kuu(ku)</i>	✓	
32	a	<i>lit.</i> we respond to some bit bit bit	<i>twilangamuu(kuu)(ku)</i>	✓	
	a	<i>lit.</i> you prevent some bit bit bit	<i>waaziyi^{zaakuu}kuu(kuu)(ku)</i>	✓	
33	a	I enable them to get at least a bit	<i>mbafunhise(na)(ku)</i>	✓	
	b	let's proceed for the moment	<i>tweyongele(nga)(yo)</i>	✓	
34	a	child of a lamb	<i>omwana gw'endiga</i>		✓
	b	I thank you for the work you have done	<i>Mbeebaziiza emilimo dhe mukoze</i>		✓
	c	some other jobs / work	<i>emilimo edhindi</i>		✓
	d	<i>lit.</i> even trees God created them	<i>n'emiti Katonda n'eyadhitonda</i>		✓
	e	celebrations were held	<i>emikolo dhaatebwa</i>		✓
	f	our souls	<i>emyoyo dhaife</i>		✓
	g	<i>lit.</i> you give to the one who gives you	<i>mwino akuwa gy'owa</i>		✓

The Compilation of Multilingual Concept Literacy Glossaries at the University of Cape Town: A Lexicographical Function Theoretical Approach*

Dion Nkomo, *School of Languages, Rhodes University, Grahamstown,
and Department of Afrikaans and Dutch, Stellenbosch University,
Stellenbosch, South Africa (deeouf@yahoo.co.uk)*

and

Mbulungeni Madiba, *Multilingualism Education Project, Centre for Higher
Education Development, University of Cape Town, Cape Town, South Africa
(mbulungeni.madiba@uct.ac.za)*

Abstract: This article proposes a lexicographical approach to the compilation of multilingual concept literacy glossaries which may play a very important role in supporting students at institutions of higher education. In order to support concept literacy, especially for students for whom English is not the native language, a number of universities in South Africa are compiling multilingual glossaries through which the use of languages other than English may be employed as auxiliary media. Terminologies in languages other than English are developed by translating English terms or coining new terms in these languages to exploit the native language competence of most students. The glossary project at the University of Cape Town (UCT) which was conceived under the auspices of the Multilingualism Education Project (MEP) is discussed. It is shown that the UCT glossaries are compiled using methods consistent with those employed in modern lexicography or proffered in lexicographical theory. The lexicographical function theory is specifically used to account for the glossaries and their production. It is suggested that modern lexicography can provide useful guidance for the production of glossaries, given that the earliest glossaries constitute the humble beginnings of lexicography.

Keywords: MULTILINGUAL CONCEPT LITERACY GLOSSARIES, GLOSSARY, DICTIONARY, TERMINOLOGY, TERMINOGRAPHY, LEXICOGRAPHY, SPECIALISED LEXICOGRAPHY, USERS, CORPUS, WORDSMITH

Opsomming: Die samestelling van veeltalige-konsep-geletterheidswoorde-

* Part of this article was presented as a paper at the *International Conference on Language Policy, Planning and Support in Higher Education — Challenges of Multilingualism*, organised by the Language Centre of Stellenbosch University and held at the Spier Hotel, Stellenbosch, 17-20 November 2009.

lyste by die Universiteit van Kaapstad: 'n Leksikografiese funksieteoretiese benadering. Hierdie artikel stel 'n leksikografiese benadering tot die samestelling van meertalige-konsep-geletterheidswoordelyste voor wat 'n baie belangrike rol kan speel by die ondersteuning van studente by instellings vir hoër onderwys. Om konsepgeletterdheid te ondersteun, veral vir studente vir wie Engels nie die moedertaal is nie, stel 'n aantal universiteite in Suid-Afrika meertalige woordelyste saam waardeur die gebruik van ander tale as Engels as hulpmedia aangewend kan word. Terminologieë in ander tale as Engels word ontwikkel deur Engelse terme te vertaal of nuwe terme in hierdie tale te skep om die moedertaalvaardigheid van die meeste studente te benut. Die woordelysprojek by die Universiteit van Kaapstad (UK) wat ontwerp is onder die beskerming van die Multilingualism Education Project (MEP) word bespreek. Daar word aangetoon dat die UK-woordelyste saamgestel word deur die aanwending van metodes ooreenstemmend met dié gebruik in die moderne leksikografie of voorgestel in die moderne leksikografiese teorie. Die leksikografiese funksieteorie word spesifiek gebruik om die woordelyste en hul totstandbrenging te verantwoord. Daar word aangevoer dat die moderne leksikografie nuttige leiding vir die totstandbrenging van woordelyste kan verskaf, gegee dat die vroegste woordelyste die nederige begin van die leksikografie uitmaak.

Sleutelwoorde: VEELTALIGEKONSEP-GELETTERHEIDSWOORDELYSTE, WOORDELYS, WOORDEBOEK, TERMINOLOGIE, TERMINOGRAFIE, LEKSIKOGRAFIE, GESPECIALISEERDE LEKSIKOGRAFIE, GEBRUIKERS, KORPUS, WORDSMITH

1. Introduction

Following South Africa's multilingual language policy in 1994, African language terminologists, who are all first language speakers of the various official languages have been employed to document African language terminology on a variety of subject fields (Alberts 2010: 610). This has resulted in the publication of technical dictionaries or glossaries under the auspices of the Department of Arts and Culture (DAC). Alberts (1999, 2010) provides an outline of the policies, procedures and principles that have guided such terminology work. Besides the terminology work conducted by the Terminology Coordination Section, i.e. the section that deals with terminology under the National Language Services of the DAC, several South African institutions of higher education have engaged in the compilation of multilingual concept literacy glossaries. These glossaries may also play a significant part in the development of terminology in South Africa (Alberts 2010: 616). This article argues for a lexicographical approach towards the compilation of such glossaries. It suggests that modern lexicography provides a theoretically sound and comprehensive scope that may facilitate the production of functional and user-friendly products. Firstly, common fundamental motivations between the compilation of the earliest glossaries and modern lexicographical practices are noted, thereby dismissing what have been purported to be major differences between glossaries on the one hand and dictionaries on the other. Secondly, the boundaries between lexicography, particularly specialised lexicography, and

terminography, have already been dismissed as fluid and practically irrelevant according to the lexicographical function theory (Bergenholtz and Tarp 1995, 2003; Bergenholtz and Nielsen 2006; Tarp 2000; Fuertes-Olivera and Arribas-Baño 2008). The lexicographical function theory is then critically applied to account for the multilingual concept literacy glossaries being compiled under the auspices of the Multilingualism Education Project (MEP) at the University of Cape Town (UCT). The glossaries seek to satisfy the pedagogical needs especially of first year learners at UCT, who are non-native speakers of English.

2.1 Disciplinary Claims for Subject-Field Glossaries

Which discipline is responsible for the production of glossaries? This may seem to be an irrelevant question. Probably it has not been raised before because glossary makers, i.e. glossarians, glossarists or glossators, as they are variously called (cf. Hartmann and James 1998), are more likely to be concerned about the final products, i.e. glossaries, rather than how their disciplines shape the outputs. However, the question is both theoretically and practically significant. It is rather more a question of quality than of quantity or availability, taking into cognisance the motivations behind the creation of the earliest glossaries and the developments having since taken place to inform the production of glossaries today. Once the influence of the various disciplines informing the production of glossaries is clear, the nature of the available glossaries may be better understood with a view to determining how they may be improved.

There are at least three established practical fields or disciplines to which the compilation of glossaries have been or may be attributed. These are *glossography*, *terminography/terminology* and *lexicography*. The theoretical and disciplinary status of each of these will be examined in order to determine how much influence each has had on the production of glossaries in the course of time. Furthermore, the relationships and lack thereof between them is also explored in order to demonstrate the viability of the lexicographical approach.

In addition to the three fields or disciplines mentioned in the previous paragraph, subject-field glossaries have also been, and continue to be, produced by experts in a variety of other scientific, professional and academic disciplines. Some of the glossaries have been published as independent hard copies; others have been integrated or appended to textbooks or practical manuals while others are available online. Such glossaries will not be discussed at length here. Suffice it to say that their compilation largely relies on the specialised knowledge possessed by the experts in the specialised subjects. Despite their accuracy as far as specialised knowledge is concerned, some of the glossaries fall short when it comes to the linguistic and communicative needs of the target users.

2.1.1 Glossography

According to the definitions from the *Free Online Dictionary, Thesaurus and Encyclopedia* (accessed on 6 November 2009), *glossography* may be defined in sim-

ple terms as the compilation of glosses and glossaries. However, the term seems not to be popular any more as it does not appear in many contemporary English dictionaries, including dictionaries of linguistics. Even Sauer (2008), who has written a book chapter about glosses and glossaries, does not use the term. However, Hartmann and James (1998: 63) have lemmatised it and treated it in a relatively comprehensive and insightful way. They define *glossography* as:

A complex of activities concerned with the compilation of glossaries from the Middle Ages to the Renaissance. Starting with glosses or notes on hard words in Latin texts, vocabulary lists gradually developed into context-independent compilations, laying the foundations for lexicography, monolingual as well as bilingual and multilingual, alphabetic as well as thematic, lexical as well encyclopaedic.

This article of Hartmann and James confirms that glossography is concerned with the compilation of glosses and glossaries. However, it seems to suggest that this complex of activities was time-framed, i.e. "from the Middle Ages to the Renaissance". This then makes the question regarding the disciplinary claim for the compilation of glossaries in the modern era relevant for determining the progress that has been made in practice. It is clear that it should no longer be glossography, which Hartmann and James (1998) describe retrospectively. More significantly, the article of Hartmann and James suggests that glossography provided the "foundations for lexicography". Then, given the progress that has been made by lexicography up to the present, to what extent has the production of glossaries incorporated modern developments in lexicographical theory and practice? This question, among others related to the similarities and differences between lexicography and the compilation of modern glossaries, will be addressed partly in Section 2.2.2 and thereafter in the remainder of the article.

2.1.2 Terminography (Terminology)

Terminography is another field which has produced subject-field glossaries through the documentation of terms (Sager 1996, Bowker and Pearson 2002). As a practical activity, terminography may be understood in the context of terminology. According to Sager (1996: 3), terminology has the following three senses:

- a vocabulary of a special subject field, i.e. terms;
- practices and methods used for the collection, description and presentation of terms;
- a theory, i.e. the set of premises, arguments and conclusions required for explaining the relationships between concepts and terms which are fundamental for a coherent activity.

Terminography is encapsulated in the second sense. Glossaries result from the collection, description and presentation of terms. In African languages, terminography work does not only involve the collection, description and presentation of terms. It also includes term-creation which is integral since equivalent terms are created for established terms in languages such as English. In South Africa, terminologists and terminographers have produced subject-field glossaries, e.g. the trilingual glossaries for subject fields such as Law, Psychology and Economics compiled by Pumlanzi Sibula at the Unit for IsiXhosa of the Language Centre of Stellenbosch University. However, terminological theories do not account fully for all the processes involved in the compilation of such glossaries. For instance, in the case of African languages, terminology development is not purely onomasiologically oriented as proffered in traditional terminological theories since in some cases term-creation begins with an English term rather than a concept. Furthermore, the issues of data presentation and accessibility in the final products may not be fully appreciated within the scope of existing terminological theories.

As a conscious practical activity and subject of intellectual enquiry, terminology may be traced back to the 1930s when Eugen Wüster, later to be regarded as the father of terminology, developed the General Theory of Terminology (GTT). Although the GTT was a milestone in the intellectual development of terminology as a field, it met severe criticism over time. Some of the issues continue to bedevil practical terminology as acknowledged in the *POINTER Project Final Report* (1996: 1):

A common problem of terminology work is that the importance and indeed the very nature of terminology is poorly understood. ... many people simply have no idea of what it is, while others, searching for an explanation of some sort, end up associating it with "thermal science" and hence radiators ... Related professions in the communications field, such as translation and technical writing, will often be aware of the word without having precise knowledge of what it entails.

The situation described here may be explained by interrogating the very foundations of terminology as a practical field. "Terminology ... was a science pioneered by subject experts" (Antia 2005). Wüster himself was an engineer and, as such, treated terms as engineered language, and not part of natural languages. Thus the initial focus of terminology was on concepts and their designations, terms, which were understood as units of knowledge, and not as lexical units. The relationship between concepts and terms was supposed to be fixed and univocal regardless of context. In quest of standardised communication, especially between experts, terms were strongly prescribed. Synonymy and variation were not tolerated. Terminography became extremely prescriptive to such an extent that it resulted in products which did not fully address cognitive and communicative needs, especially of laypersons.

However, successive scholars such as Sager (1984, 1996), Temmerman (2000) and Cabré (2000) have expanded the scope of understanding terminology. For example, Sager (1996) identifies three dimensions of terminology, namely the cognitive, the linguistic and the communicative. Temmerman (2000) and Cabré (2000) argue for a sociocognitive approach, with the latter developing an alternative terminological theory called the Communicative Theory of Terminology. Bowker and Pearson (2002) have incorporated developments from corpus linguistics in terminology development, particularly the compilation of glossaries. These later developments differ from the GTT in that they view terms as part of natural languages whose meanings may be dynamic, depending on their disciplines, communicative situations and contexts. They also acknowledge the fact that specialised communication does not only occur among experts but also between an expert and a semi-expert or an expert and a layperson, a user-categorisation which has been refined within the context of metalexigraphy (cf. Bergenholtz and Tarp 1995, 2003). In such situations, context plays an important role in the conceptualisation of the represented concept.

In the light of the foregoing, this article contends that the compilation of subject-field glossaries in South Africa needs to draw insights from modern theories and practices instead of being restricted by the tenets of the GTT. This has been shown elsewhere by scholars such as Madiba (2004, 2010) who demonstrates how corpus linguistics and its analytic tools may improve terminological work. However, this article argues that such developments in terminology may be incorporated into a lexicographical model and that such a model may improve the compilation and use of glossaries. In so doing, the article will dismiss some implied claims that lexicography lacks the dynamics facilitating the production of functional and user-friendly language tools (cf. Bowker and Pearson 2002; Moropa 2004).

2.1.3 Lexicography

Among a variety of its products, dictionaries are the primary concern of lexicography, with practical lexicography focusing on dictionary production and metalexigraphy focusing on the development of theories to support lexicographical practice. Glossaries have not received significant attention from metalexigraphy and any suggestion that attributes the production of glossaries to lexicography may meet with strong opposition, especially from those engaged in the production of glossaries. There are at least two reasons for this:

- Most of the available dictionaries have major limitations and barely satisfy the needs they are purported to serve. Consequently, glossaries have been seen as an alternative to dictionaries which unfortunately provide limited assistance (cf. Bowker and Pearson 2002; Moropa 2004).

- There is a general lack of awareness of the progress metalexigraphy has made to address some major limitations of existing dictionaries and the resultant potential practical lexicography has accrued over the years. Consequently, poor quality dictionaries continue to be used as general references to lexicography. In the light of this, scholars tend to view glossaries and other language resources such as corpora as alternatives to dictionaries without noting that these resources are not necessarily incompatible with dictionaries (cf. Bowker and Pearson 2002; Moropa 2004).

Because of reasons such as the above, the production of glossaries is justifiable. Glossaries are meant to serve certain needs which dictionaries *have failed* or *do not seem to have the capacity* to address. However, it will be argued that the compilation of glossaries *per se* is not a sufficient remedy given that dictionaries developed as modifications of glossaries (Sauer 2008). What really matters is the adoption of an approach and principles that would ensure that the products are not simply better than the available dictionaries but also much better by far than the earliest glossaries which laid the foundations for dictionary-making. A poor glossary would not be any better than a poor dictionary. Therefore it will be demonstrated that the production of useful glossaries cannot be inspired by having as a point of departure anti-lexicographical discourses which are not only based on superficial distinctions between a dictionary and a glossary but which also fail to demonstrate the advantages of the latter over the former. This will be done by highlighting a lexicographical approach which is not based on the compartmentalisation of knowledge, be it between glossography, terminography and lexicography itself.

2.2 Building Useful Glossaries: Breaking Unnecessary Disciplinary Barriers

There is a disciplinary and theoretical confusion which affects the practical production of glossaries for various purposes. This confusion emanates from the often unnecessarily sharp distinction between glossaries and dictionaries as products of distinct disciplinary practices, namely glossography or terminography and lexicography. In South Africa, some practitioners involved in the compilation of glossaries are regarded as translators simply because translation is one of the main methods of creating terms whereby equivalent terms in indigenous languages are created to meet punctual translation needs of and by translators (Madiba 2004). While the alleged distinction between glossography and lexicography has barely been investigated in an insightful way, the distinction between terminography and lexicography, especially specialised lexicography, has been effectively dismissed as of no practical value (cf. Bergenholtz and Tarp 1995, 2003; Bergenholtz and Nielsen 2006; Tarp 2000; Fuertes-Olivera 2010; Fuertes-Olivera and Arribas-Baño 2008). More attention will be given to the former distinction while a brief recapitulation of the cited

sources will be made on the distinction between the latter. In so doing, the article puts the compilation of glossaries within a historical perspective which, in turn, translates into a lexicographical perspective.

2.2.1 A Historical Perspective: From Glossaries to Dictionaries

A historical perspective makes it clear that glossography is the activity preceding lexicography, whose progress has added the dimension of "a field of study" to the meaning of the suffix *-graphy*, an Anglicisation of the French *-graphie*, inherited from the Latin *-graphia*, which is a direct transliteration of the Greek verb which means "to write" (*Free Online Dictionary, Thesaurus and Encyclopedia*, accessed on 6 November 2009). The dictionary article corresponding to the lemma *glossography*, taken from Hartmann and James (1998: 63) makes it clear that the practice started with glosses or notes on hard words in Latin texts. Vocabulary lists then "gradually developed into context-independent compilations, laying the foundations for lexicography, monolingual as well as bilingual and multilingual, alphabetic as well as thematic, lexical as well encyclopaedic". Glossography may therefore be regarded as the humble beginnings of lexicography.

In spite of these reflections, it cannot be ascertained whether the earliest glossarians, glossarists or glossators called themselves such, but it is unlikely given that they were mainly educators who considered it helpful to explain to their learners some of the hard words used by text producers. It is doubtful whether they would also have called their activity 'glossography', as it was not their main engagement. What is seen is the development of a practice whereby text producers use difficult words of foreign origins (Latin and French) and educators make efforts to explain them, first within texts and later outside their contexts as independent compilations. This is similar to the modern practices mentioned earlier where specialised subject experts identify difficult concepts and explain them in glossaries which are either published as independent texts or integrated into textbooks.

The real functions of glosses and glossaries are summarised by Sauer (2008: 19):

The main purpose of the glosses, as well as glossaries, thus must have a didactic one: interlinear glosses facilitate the understanding and possibly also the learning of the glossed Latin text. Glossaries help the acquisition of the Latin vocabulary (and probably also of the English vocabulary). Thus many of the glosses and glossaries must have been used for teaching purposes in schools, especially in monastic and cathedral schools.

As may be noted, it was a quest to solve communication problems and facilitate knowledge acquisition and dissemination that motivated the earliest glossaries. The problems could have emanated from the fact that the texts contained foreign language elements and also that they contained specialised lan-

guage vocabulary items. The second cause of problems, namely terminology, is also highlighted by Sauer (2008: 19) in the context of professional fields such as medical practice:

Specific glossaries (class glossaries) such as the plant name glossaries were ... intended for use of physicians and healers, to help them identify the plants in Latin (and later also in Anglo-Norman) and to prepare the proper medicines.

The historicisation of lexicography shows not only that glossaries were forerunners of dictionaries, as explained by Hartmann and James (1998) and Sauer (2008), but also that dictionaries have or ought to have exactly the same functions as glossaries (cf. Tarp 2008). However, as already stated, dictionaries have been criticised in favour of glossaries on the premise that they do not provide adequate assistance (cf. Bowker and Pearson 2002; Moropa 2004). For example, before arriving at such a conclusion, Bowker and Pearson (2002: 139) ask some questions:

How often do you have to consult more than one dictionary in order to find the information you are seeking? How often do you find that the information you are seeking is simply not there? How often do you choose a word in the dictionary without really knowing whether it is the right one? While dictionaries are an essential part of any language student's toolkit, they rarely provide the answer to all your questions.

Intriguing as these questions are and accurate as the conclusion is regarding many dictionaries, Bowker and Pearson fail, however, to construe a plausible argument for the preference of glossaries. That dictionaries contain more information than glossaries (Sauer 2008: 21), which is generally used as the main distinction, may not be their limitation except when such information serves no function or remains inaccessible. The progress made in lexicography, from the earliest glossary making up to the present, will also invalidate the claims that the limitations of dictionaries compared to glossaries is their inability to provide a specific type of information. For example, grammatical information, usage examples and sense discrimination are among other information types which can be found in modern dictionaries. This fact is ignored in Bowker and Pearson (2002). Sauer (2008: 22) correctly observes that "the difference between a glossary and a dictionary is one of degree and ... more by convention that the early collections are usually called glossaries and the later collections ... dictionaries". To this it suffices to add that the argument that no single dictionary will ever provide the information required by a student (Bowker and Pearson 2002: 140), also applies to some glossaries and does not make them any better than dictionaries in general. One product will be better than the other depending on how it has been conceived in view of its functions, not simply because one is a glossary and the other is a dictionary.

From the foregoing, it may not be wrong to suggest that in its reverse

mode, lexicography would go back to embody glossography. The earliest glossaries may then be regarded as the earliest lexicographical products. The question as to why dictionaries as they are known today developed from glossaries may be easily answered by pointing to their capacity to accommodate more data. Then why do glossaries continue to be produced alongside dictionaries? The answer is that the quality of dictionaries did not increase proportionally with the increase in the quantity of information they provide. That the quality of assistance provided by glossaries is generally better than that provided by dictionaries has so far not been adequately demonstrated. As such, the production of glossaries may still benefit from lexicographical theories and methods primarily focused on dictionaries.

2.2.2 A Lexicographical Perspective

The historical perspective which suggests that lexicography encapsulates the compilation of glossaries, or glossography, prompts the lexicographical approach advanced here. It is said that dictionaries, not in their current form but as glossaries, were first compiled over thousands of years (McArthur 1986; Tarp 2008). However, it was not until the second half of the 20th century that lexicography started to constitute itself as an academic discipline and professional practice. Prior to that, dictionary making was generally a part-time activity, starting with educators compiling glossaries for their students and then subject-field experts compiling subject-field glossaries, with linguists claiming the rest through their involvement in 'general language lexicography'. This became the greatest source of confusion regarding the disciplinary status of lexicography, with linguists generally regarding lexicography as part of so-called 'applied linguistics'.

That lexicography is not part of linguistics was first challenged by Wiegand (1984) who argued that it was a professional activity and a scientific practice whose object was not language but the production of dictionaries. He also argued that dictionaries were utility products with each having its own *genuine purpose*, thereby making the user an important variable in the whole enterprise. His ideas were adopted by practising and academic lexicographers at the Centre for Lexicography at the Aarhus School of Business, Denmark, who collectively developed what they now call the lexicographical function theory (Bergenholtz and Tarp 1995, 2003; Tarp 2000, 2002, 2008). However, the proponents of the lexicographical function theory did not uncritically adopt Wiegand's theories, especially his general theory of lexicography. Tarp (2000, 2002, 2008), for example, has criticised Wiegand for having a linguistic bias, before strongly arguing that lexicography is in fact an independent scientific discipline. This called for a development of lexicographical theories which would improve the lexicographical practice and its products. Linguistic training is no longer regarded as a prerequisite qualification and it

has been shown that the relevance of its theories and findings varies from one dictionary to another, with LSP lexicography requiring more than linguistic skills.

The lexicographical function theory and its application in lexicographical practice and dictionary criticism have been described in more detail by its proponents in several publications which include Bergenholtz and Tarp (1995, 2003), Bergenholtz and Nielsen (2006), and Tarp (2000, 2002, 2008). Its main argument is that every dictionary should be produced with a specific user in mind, devoting attention to questions such as the following (cf. Bergenholtz and Tarp 2003: 173):

- Who are the intended users of the prospective dictionary?
- What is their native language and how are they competent in it?
- What is their competence in a particular foreign language?
- What are the characteristics of the intended users in terms of their competence regarding specific languages and subject knowledge?
- In what situations are they likely to experience problems which may be solved by referring to a dictionary?
- What problems are they likely to encounter in the specific situations which may be solved by referring to a dictionary?
- What information needs may address the specific problems?
- What types of data need to be included in a dictionary which may provide users with certain information types required to solve specific problems encountered by specific users in specific situations?
- What lexicographical devices should be employed to ensure an unimpeded and successful access to data while reducing information costs?

Focusing on the glossary project of the University of Cape Town, Section 3 applies some ideas from the lexicographical function theory to demonstrate that the production of multilingual concept literacy glossaries may benefit from some lexicographical insights.

3. UCT's Multilingual Concept Literacy Glossaries

The UCT Multilingual Concept Literacy Glossaries project was initiated in 2007 as part of the implementation of the University's Language Policy (1999 revised in 2003) and the Language Plan adopted in 2003. The Language Plan requires that multilingual concept literacy glossaries be developed to support students for whom English is not their first language. Thus, the glossaries are aimed at concept literacy in the different content-learning areas, with the pilot project focusing on

developing glossaries for Statistics, Economics and Law. All these glossaries are based on the Special Language corpora constructed for this purpose. Drawing from the lexicographical function theory, the remainder of the article characterises the target users of the glossaries, defines the user situations, enumerates the envisaged assistance that the glossaries were conceived to provide to the target users and shows how all this informed the compilation of the glossaries.

3.1 User Characterisation

The glossaries being produced are targeted at first year students of various specialised subject fields at the University. The following screen shot shows a list of the disciplines that the project seeks to cover in the long term.

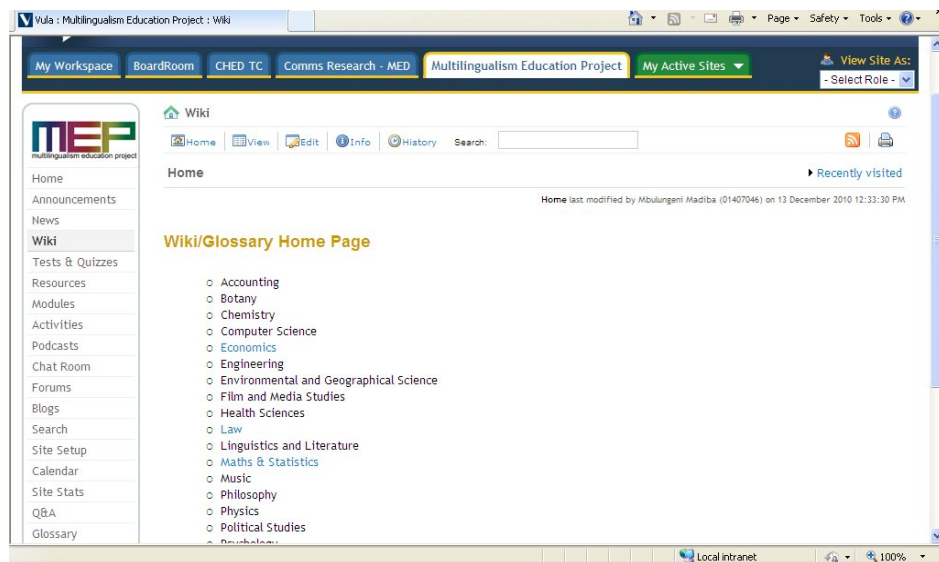


Figure 1: Subject fields for which glossaries are created at UCT

At first year stage, students have low levels of competence in the respective fields and low proficiency in English which is the primary language of instruction at UCT. Some of these students have high levels of competence in their home languages depending on the location of the schools in which they matriculated. In addition to language competences and competence in the specific subjects, another characteristic of the students that should be considered is their cultural backgrounds, which may have a bearing on their competence in subjects such as Law.

The focus on first year non-native English speaking students is very much

consistent with the requirement of the lexicographical function theory that every lexicographical product should be conceived with a specific user in mind. This requirement may be problematic when applied uncritically in developing countries where lexicographical resources and resources required to support lexicographical projects are very limited. In this regard, other potential users of the glossaries were not neglected at the conceptualisation stage. English native-speaking first year students may also use the UCT glossaries, seeing that general language competence in English does not automatically translate to LSP competence in the various subject fields. Furthermore, students at higher levels of university education may still consult the glossaries when they experience some of the typical challenges encountered by first year students for whom the glossaries are specifically produced. Even outside university, translators have also been considered as potential users of the UCT multilingual glossaries, while those who have participated in the project found it potentially useful. What is critical though is that first year non-native English speaking students are deliberately prioritised, while not neglecting other potential users.

3.2 User Situations

The situations in which the multilingual glossaries are envisaged to help students relate to their learning of the key concepts in the various subject fields, coupled with the mastering of the respective LSP varieties and registers. The learning situations may include lectures, tutorials, the writing of assignments and revision for examinations. In terms of the lexicographical function theory, these situations may be classified into communicative and cognitive functions (cf. Tarp 2008: 84). The glossaries were conceived to address some of the problems learners may encounter in such learning situations.

3.3 Problems

In the learning situations described in Section 3.5, students first encounter communication-oriented problems, primarily the reception of oral texts (lectures) as well as written texts (textbooks, course readers and assignment questions). They also experience text production problems like oral contributions and questions in lectures and tutorials, as well as text production in the form of writing assignments and examinations. Ultimately, this prevents them from grasping the key concepts and reaching the required performance levels in the specific subjects.

3.4 User Needs

Students have various needs to address both communicative and cognitive

problems in their specific learning situations. They need appropriate terminology and information regarding meaning, register and examples of its usage. This is necessary to enable them to receive and produce academic texts in the various subject fields. Those students whose home language is not English need translation equivalents of the terminology as well as translated definitions in order to address text reception problems. In addition to this, all students need special and encyclopaedic information about the concepts so that they may master the respective subject field.

3.5 Functions of the Glossaries

The MEP concept literacy glossaries seek to serve both cognitive and communicative functions.

The major cognitive functions are:

- Assisting university students at first year level with specialised knowledge of threshold concepts in various subject areas;
- Assisting university students at first year level with English LSP in various subject areas;
- Assisting students with special subject-field information in the form of key concepts and conceptual relations within the subject fields.

The communicative functions include:

- Assisting students with text reception in English;
- Assisting students with text production in English;
- Assisting lecturers and tutors with text production in indigenous languages;
- Assisting lecturers, tutors, learners and translators with the translation of texts from English to indigenous languages;
- Assisting students with text reception in English;
- Assisting students with text reception via their indigenous languages as auxiliary medium;
- Assisting students with text production in English.

In view of the above functions, the genuine purpose shared by the MEP glossaries is to facilitate concept literacy in English, the main language of education at UCT, and in indigenous languages as auxiliary media. The glossaries are thus conceived as multilingual pedagogical resources which facilitate communication and learning in the context of special subject fields. These

aspects guided the planning and compilation of the glossaries, for, following Tarp (2004: 312), "no data is included because of tradition or the practice of existing dictionaries". This is a central tenet of the lexicographical function theory.

3.6 The Compilation Process

The compilation of the multilingual concept literacy glossaries at UCT was mainly based on special language corpora. However, lecturers, tutors as well as translators and terminologists who are native speakers of the official languages of South Africa other than English were involved for their different kinds of expertise. The next subsections describe the special language corpora, the compilation of wordlists, the use of corpus data in defining and the online presentation of the glossaries.

3.6.1 Special Language Corpora

The special language corpora are used as the main conceptual bases for the glossaries. The corpora are based on generally accepted criteria for designing special language corpora (cf. Bowker and Pearson 2002; Madiba 2004). These criteria include size, text types, publication status, text origin, constitution of the texts, authorship, external and internal criteria. With regard to the first criterion of size, special language corpora tend to be much smaller than general language corpora as a result of their content and the compilation process (Lawson 2001: 293). Several studies recommend at least 30 000 words for a small special language corpus. Although the size of some of the UCT corpora is small, most important is the purposes for which they have been designed (Meyer and Mackintosh 1996).

The second criterion of text selection is important since not all texts are useful for special language corpora. Text selection involves choice of the language and type of the texts. The UCT corpora comprise English texts only since there are no texts available in the indigenous African languages at higher education level. Only written texts were selected and such texts are based on either module texts or full texts. To achieve a good balance in each corpus, texts are selected according to external as well as internal criteria. The external criteria are essential to maintain a good balance in the corpora. These criteria include the domains of individual texts, text genres, publication status and age, text origins, constitution, and persons who produce texts (authors). With regard to the domains of individual texts, specialised texts are considered, i.e. texts covering different genres developed for special purposes and specialised subject fields. In order to allow for the analysis of terms used in specialised subject domains and the compilation of specialised glossaries for these domains, it is important that domains be clearly identified from the start. Once the bounda-

ries of a domain have been determined, it was attempted to establish the truest possible representation of the delimited domain in the corpus. The chosen body of texts should cover all aspects of the domain, including subdomains and related domains, as equally as possible. As a rule of thumb, the focus should be on one domain at a time (cf. Madiba 2004).

The third criteria concerns text genres. A text genre is a text type with its own linguistic and pragmatic conventions (Meyer and Mackintosh 1996). Since different domains produce different texts, various texts occurring in a single domain were considered. These texts were divided into two general types, namely, *instructional texts* (texts used for pedagogical purposes such as books and tutorials) and *advanced texts* (abstracts of theses/dissertations and full theses/dissertations where necessary). With these genres the corpora should be able to achieve linguistic as well as pragmatic balance.

The fourth criterion concerns the selection of texts on the basis of their publication status or age. The UCT corpora comprise published and unpublished instructional materials. If unpublished, the authenticity of such texts is checked. With regard to age, the main focus is on current texts as the purpose is to establish synchronic terminological corpora focusing on the present state of terminology used in different domains.

Lastly, the constitution of the texts and their internal criteria were considered. The internal criteria have to deal with formal linguistic characteristics such as factuality and technicality. In compiling the corpora, it was attempted to include texts that are as factual as possible and thus to enrich the corpus linguistically and conceptually. The texts should provide an extensive range of terms generated by the discourse community, accompanied by a maximum number of usages of concepts in different contexts. The technicality of the texts should range from high to low. Highly technical texts usually involve communication between subject specialists whereas less technical texts relate to communication between non-specialists (e.g. tutorials). The different degrees of technicality is useful for the identification of terms and concepts used in different domains, that is, specialised usage and everyday usage (cf. Madiba 2004).

Regarding the criteria discussed in the foregoing, it may be noted that consideration was given to the texts students encounter in their respective subject areas. The texts have direct consequences for the learning situations as well as the problems and needs experienced by the learners. The UCT Glossary of Economics will be used to illustrate aspects of the previous discussion. This glossary is based on a small corpus consisting of prescribed books, study guides and tutorials in Economics. The size of the corpus is about 70 000 running tokens. Although the corpus is small in size, it contains relevant key concepts and their contexts because of the module approach used in collecting texts. As Sinclair (2001) rightly pointed out, the advantage of a small corpus is that it is unnecessary to wait until the corpus is complete, but instead, work can start from the first day.

3.6.2 Extraction of Terms and Compilation of the Basic Wordlist

The extraction of terms was carried out by using WordSmith Tools. WordSmith Tools does automatic term extraction on the basis of statistical analyses. However, it requires that the texts be prepared in text format first before term extraction is carried out. The following are examples of terms extracted from the Economics corpus by using WordSmith Tools:

N	Word	Freq	% Lemmas
1	the real cost	60	1.06
2	production possibilities curve	56	0.99
3	real cost of	51	0.90
4	cost of producing	49	0.87
5	the production possibilities	42	0.74
6	used to produce	27	0.48
7	the resources are	23	0.41
8	a production possibilities	22	0.39
9	the opportunity cost	22	0.39
10	opportunity cost of	21	0.37
11	data for a	18	0.32
12	to calculate the	16	0.28
13	land and labour	15	0.27
14	what is the	14	0.25
15	bundle of output	13	0.23
16	perfectly flexible between	13	0.23
17	that the resources	13	0.23
18	in figure #	12	0.21
19	increasing opportunity costs	12	0.21
20	next best alternative	12	0.21
21	of producing the	12	0.21
22	the second pencil	12	0.21
23	all its resources	11	0.20
24	cost of the	11	0.20
25	of the two	11	0.20
26	possibilities curve showing	11	0.20

Figure 2: A frequency wordlist generated from the Economics corpus using WordSmith Tools

A final wordlist was compiled not solely on the basis of the number of occurrences or frequencies. Having gone through the basic list, Economics lecturers and tutors were consulted since some common words would normally have a high frequency of occurrence in specialised texts while some terms would be of a relatively low frequency. After the compilation of the final wordlist, concordances were generated using WordSmith Tools and Multiconcord to identify the meanings or senses of the terms in different contexts. Figure 3 is an illustration of the concordances of the term *opportunity cost* generated by using WordSmith Tools

The concordances in the examples provide different contexts of the term *opportunity cost*. In WordSmith tools, these contexts can be expanded as in the screen shot shown in Figure 4 to provide a better understanding of the concept:

N	Concordance	Set	Tag	Word No.	File	%
1	est alternative forgone. For example, the opportunity cost of going to the beach will			80	port-1.txt	1
2	ity of training to be a medical doctor. The opportunity cost, or real cost, of training			328	port-1.txt	5
3	er to go to the beach. Why is there an opportunity cost for our every action? Thi			195	port-1.txt	3
4	ources are being used to produce pencils Opportunity cost and the production pos			2,530	port-1.txt	46
5	oose between different possibilities. .. Opportunity cost in production and cons			455	port-1.txt	8
6	ilities curve, in Figure 3.2, to calculate the opportunity cost of producing different qu			4,480	port-1.txt	82
7	ilities curve, in Figure 3.2, to calculate the opportunity cost of producing different qu			4,789	port-1.txt	88
8	ame next best alternative. Therefore, the opportunity cost for each of us is differen			135	port-1.txt	2
9	t to do any number of other things. The opportunity cost of any action is the next			66	port-1.txt	1
10	ity cost for each of us is different. The opportunity cost is also known as the re			144	port-1.txt	2
11	sed to produce one pencil. What is the opportunity cost of increasing our produc			4,892	port-1.txt	90
12	ur tomatoes is one pencil. What is the opportunity cost of producing one pencil			4,852	port-1.txt	89
13	sed to produce one pencil. What is the opportunity cost of increasing our produc			4,583	port-1.txt	84
14	ur tomatoes is one pencil. What is the opportunity cost of producing one pencil			4,543	port-1.txt	83
15	ost. The concept of cost The concept of opportunity cost is important for economi			5	port-1.txt	0
16	Opportunity cost The concept of opportu			0	port-1.txt	0
17	f producing one pencil is one tomato. The opportunity cost of producing two pencils			4,950	port-1.txt	91
18	l tomato or pencil remains the same. The opportunity cost of production is constan			3,143	port-1.txt	57
19	l tomato or pencil remains the same. The opportunity cost of production is constan			3,866	port-1.txt	71
20	f producing one pencil is one tomato. The opportunity cost of producing two pencils			4,641	port-1.txt	85
21	to produce tomatoes and/or pencils. The opportunity cost of producing four pencils			2,622	port-1.txt	48
22	y own as a tax accountant. What is the opportunity cost of going out on my own			496	port-1.txt	8
23	efore, every action that we take has an opportunity cost. We could have chosen			367	port-1.txt	6
24	that were used to produce the one? The opportunity cost of producing one pencil i			4,940	port-1.txt	91
25	that were used to produce the one? The opportunity cost of producing one pencil i			4,631	port-1.txt	85
26	a growing tomatoes. This means that the opportunity cost of producing a product i			3,549	port-1.txt	65

Figure 3: Examples of concordances extracted from the Economics corpus using WordSmith Tools

1 or not to go shopping at the local mall, or not to do any number of other things. The opportunity cost of any action is the next best alternative forgone. For example, the opportunity cost of going to the beach will be different for the two of us. For me it is studying economics, for you it might be going to the mall, for someone else it might be having a sleep. It is very unlikely that each of us has the same next best alternative. Therefore, the opportunity cost

2 ies over time. I cannot become a lawyer and a medical doctor at the same time. If I choose to train as a lawyer I give up the possibility of training to be a medical doctor. The opportunity cost, or real cost, of training to be a lawyer is that I cannot become a medical doctor. Every day we are required to choose between various courses of action. Therefore, every action that we take has an opportunity cost. We could have chosen to do something else. Whatever

3 ach? It is that I cannot spend those hours studying economics. I must give up the opportunity of studying economics in order to go to the beach. Why is there an opportunity cost for our every action? Think about it for a moment. We cannot do more than one thing at a time. Time is a scarce resource. I cannot go to the mall and study economics simultaneously. I have to choose between the two activities. We only have so much time to spend doing

4 encils. We can see that if the economy is producing four pencils it is producing no tomatoes (bundle E). This means that all the resources are being used to produce pencils Opportunity cost and the production possibilities curve four tomatoes. The real cost of the two pencils is four tomatoes. To produce the second pencil two more tomatoes must be given up. The real cost of the second pencil is two tomatoes. and none is being used to produce tomatoes

5 se scarce resources to satisfy their unlimited wants. People cannot have everything that they would like to consume. They have to choose between different possibilities. .. Opportunity cost in production and consumption Production v. & *?>*. Let us say I have the choice between working in an office, and earning R5 000 a month, or going out on my own as a tax accountant. What is the opportunity cost of going out on my own? It is the next best

6 costs The production possibilities curve for the five bundles (A to f) shown in Table 3.2. We can use the production

Figure 4: Examples of expanded concordances extracted from the Economics corpus using WordSmith Tools

The contextual examples are saved as a Word document and then converted into tables for better analysis as shown in Table 1.

EXAMPLES IN CONTEXT	SENSE
<p>1. <p><s>The concept of opportunity cost is important for economists. <s>Whenever we choose to do something we also choose not to do something else. <s>For example, if I choose to go to the beach today it means that I have also chosen not to study economics, or not to go shopping at the local mall, or not to do any number of other things. (choice)</p>	<p>Basic meaning (cost in terms of choices)</p>
<p>2.<p><s>The opportunity cost of any action is <u>the next best alternative forgone</u>. <s>For example, the opportunity cost of going to the beach will be different for the two of us. <s>For me it is studying economics, for you it might be going to the mall, for someone else it might be having a sleep. <s>It is very unlikely that each of us has the same next best alternative. <s>Therefore, the opportunity cost for each of us is different.</p>	<p>Basic meaning (cost in terms of choices)</p>
<p>3. <p><s>The opportunity cost is also known as the real cost. <s>What is the real cost to me of going to the beach? <s>It is that I cannot spend those hours studying economics. <s>I must give up the opportunity of studying economics in order to go to the beach.</p>	<p>Additional meaning (real cost)</p>
<p>4. <p><s>Why is there an opportunity cost for our every action? <s>Think about it for a moment. <s>We cannot do more than one thing at a time. <s>Time is a scarce resource. <s>I cannot go to the mall and study economics simultaneously. <s>I have to choose between the two activities. <s>We only have so much time to spend doing various activities during any hour, day, month, year, or lifetime. <s>The quantity of time available to us is limited. <s>We cannot do everything that we would like to do.</p>	<p>Cost in terms of time</p>
<p>5. <p><s>Let us say I have the choice between working in an office, and earning R5 000 a month, or going out on my own as a tax accountant. <s>What is the opportunity cost of going out on my own? <s><u>It is the next best alternative</u>. <s>If I did not go out on my own I could be earning R5 000 a month. <s>Strictly speaking, the real cost is the output that I could buy with the R5 000.</p>	<p>Cost in terms of money</p>
<p>6. <p><s>What is the opportunity cost to me if I spend R100 on a CD? <s><u>It is the next best alternative</u>. <s>This may be a shirt. <s>Therefore the real cost to me of consuming the CD is that I cannot consume the shirt.</p>	<p>Cost in terms of production</p>
<p>7. <p><s>Societies choose the output that they wish to consume. <s>Whenever society decides to consume a particular good it decides not to consume the next best alternative. <s>All societies have to consider the opportunity costs of any production or consumption decisions that they take.</p>	<p>National context (Cost of production and consumption)</p>

Table 1: Concordances for *opportunity cost*

From these concordances, explanations can be developed by analysing the different contexts in which the term *opportunity cost* is used. Sometimes it is also possible to find a full definition of the term in the very concordances that are being analysed, concordances 1, 2 and 3, for example. From these concordances it can be observed that *opportunity cost* involves: (a) a choice between two alternatives, (b) a choice of the next best alternatives, and (c) the real cost. As shown above, an analysis of each of the concordances gives rise to different senses or meanings. The different contexts of the term *opportunity cost* provide the students with multiple exposures to the concept. Such multiple exposures to the term are essential for a student to understand its meaning or different senses. For students to understand the concepts, they need to be introduced to the concepts in their simplified form at the beginning, and then progress to learning more complex meanings of the concept. They also need to learn how to apply it in a range of contexts as shown in the examples above. With the aid of these examples, students are introduced to the basic meaning of the concept *opportunity cost* in the first concordance and then progress to its deeper meaning in the concordances that follow. In this way, definitions based on concordances are more elaborate and helpful to students than traditional definitions, especially from dictionaries not based on corpora. Providing students with mere definitions to memorize, results in a superficial understanding of the concept. According to O'Hara and Pritchard (2009: 11), "students must have both definitional and contextual information about words, as well as repeated exposures and opportunities to learn and review them". This supports the reception, production and cognitive functions of the glossaries.

3.6.3 Multilingual Concept Literacy Glossaries on Hypermedia (Vula)

The glossaries that are being developed at UCT are uploaded on Vula, the University Online Learning Environment developed by the Centre for Education Technology and powered by Sakai. This networked Online Learning Environment provides students with easy access to the glossaries and other online courses. It is observed that the use of the online environment exploits some of the developments utilised in the compilation of electronic dictionaries, a case of lexicographical development in the context of technological advances in the 21st century. The MEP Online Learning Environment on Vula is shown in Figure 5 followed by the Online Glossaries Site in Figure 6.

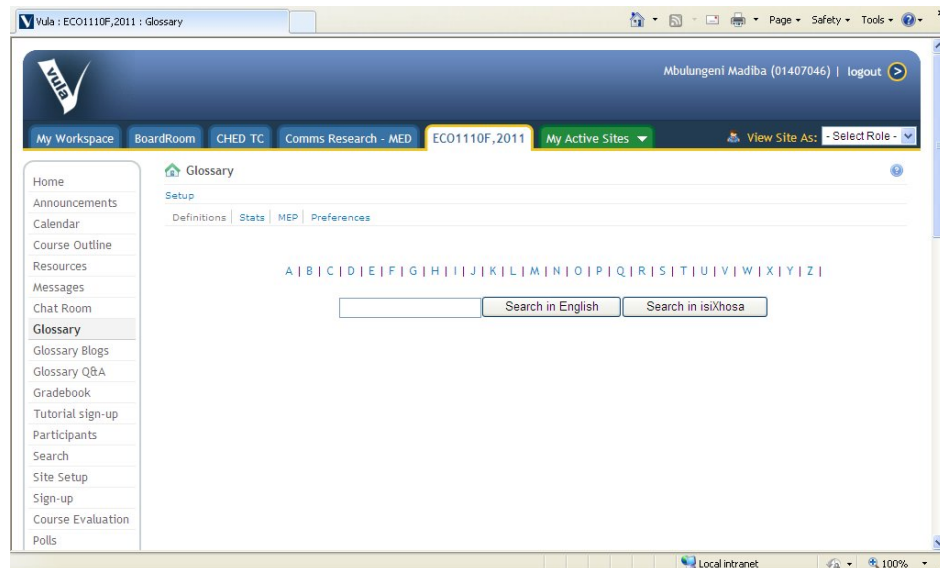


Figure 5: Economics Online Glossaries Hypermedia on Vula

From this screen shot, it may be observed that the Vula Hypermedia provides many search-route options, rendering the glossaries poly-accessible. Concepts can be accessed by keying the term into the search function or browsing the concepts in the selected languages. Language preference can be set from the start. In Figure 5, English and isiXhosa are shown as the preferred languages. However, any language can be selected as the preferred language. The search function is very effective in identifying conceptual relations as shown in the following examples of the concept *deficit*.

English

- Balance-of-payments deficit
- Budget deficit
- Capital account deficit
- Conventional deficit
- Cyclical deficit
- Deficit
- General government deficit
- International balance-of-payments deficit
- Payments deficit
- Primary deficit
- Trade deficit

Venda

- Ndinganyiso kha mbadelo ya Ṱhahlelelo
- Ṱhahlelelo ya mugaganyagwama
- Ṱhahlelelo kha akhauthu ya khephithaḷa
- Ṱhahlelelo ya mugaganyagwama
- Ṱhahlelelo yo vhangwaho nga risesheni
- Ṱhahlelelo
- Ṱhahlelelo ya muvhuso yo angaredzwaho
- Ṱhahlelelo ya zwiṰtundwa na zwivhambadze-lwannḍa
- Mbadelo dzo paḍaho
- Difisithi ya phuraimari
- Phaḍambambadzelaseli

If the user does not want to search a specific concept, browsing can be used in the traditional alphabetical glossary. In the following example, browsing is done on the letter A. This results in two wordlists, one for the source language and another for the target language.

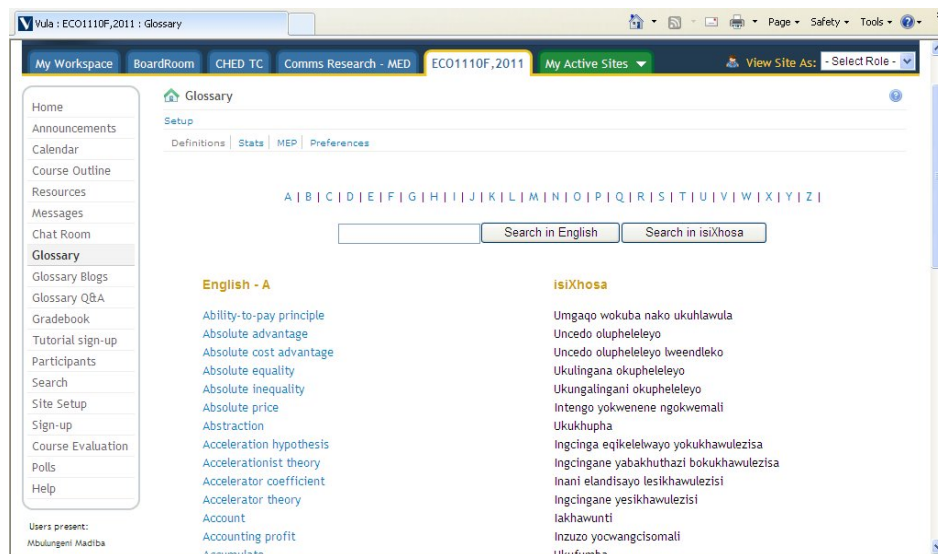


Figure 6: Economics Online Glossaries Hypermedia on Vula

Once the concept is identified, a further search can be made on its definition as shown in Figure 7.

The definition appears in both the source and the target languages. The list of other languages appears at the bottom of the screen shot. The glossary site also provides a space for the user to give feedback of the translation and the equivalent. The status of the translation equivalent is also shown as 'unassigned', 'draft', 'approved' and 'assigned'. This function is meant to facilitate the standardisation of the translated terms in the African languages based on popular usage and acceptability by the target users. While some of the terms were created by translators, the power to standardise the terms partly rests with the users, students and other professionals in the disciplines. Because terms and translation equivalents are not prescribed, synonyms and variants are provided, contrary to the stipulations of the GTT. The function of assigning status to an entry offers the users the best of both description and prescription, a practice which is called *proscription* in lexicography (Bergenholtz 2003).

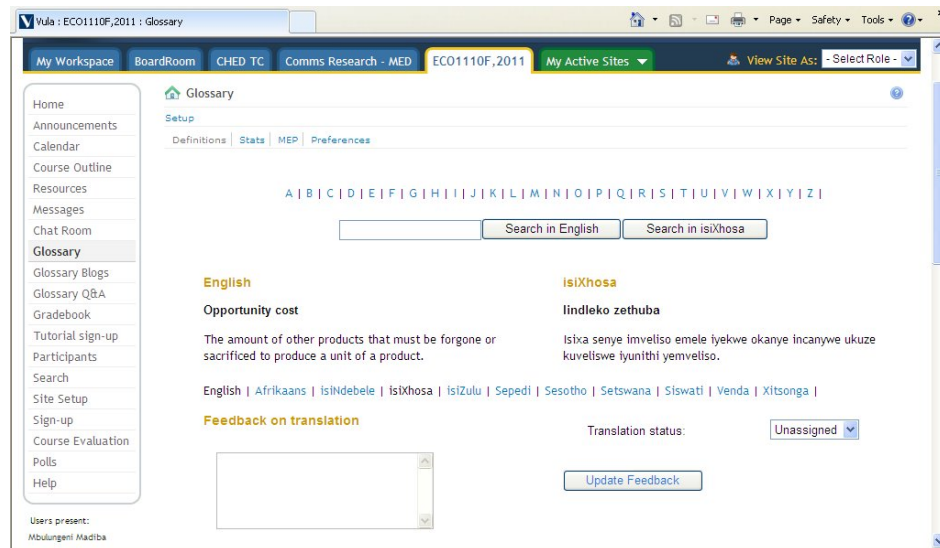


Figure 7: MEP Online Learning Environment on Vula

The definitions can further be linked with tutorials, pictures, graphics and podcasts. The glossary site also includes other functions such as modules, blogs, chat rooms and forums. Thus, unlike the traditional glossaries, the Vula Glossaries provide a more interactive environment. This makes the glossaries useful for users in reception, production and cognitive situations whereby punctual and extended searches are supported. This effectuates user-friendliness, a notion that has been discussed comprehensively in lexicography, especially with regard to the accessibility of lexicographical data.

4. Conclusion

It is still too early to assess the success of the UCT multilingual concept literacy glossaries. Further research will be conducted in this regard. However, their interactive nature and associated elements such as pictures, podcasts and blogs make the glossaries valuable resources for supporting non-native English speaking students learning concepts and their linguistic representations in their first languages. More critically, it is the user perspective which gives the glossaries more potential than existing glossaries and dictionaries. The user perspective is supported by the adoption of a corpus-based approach to the enterprise. Ahmad et al. (1992: 148) argue that the approach is "sound in as far as it is corpus-based and user-informed". Having identified the target users, their relevant characteristics and the situations in which their needs arise, the adopted corpus-based approach is characterised by a purposive selection of

texts and collaboration with both lecturers and students. All these ideas and principles have been adopted in modern lexicography. This therefore shows that the production of glossaries and other terminographical products may benefit from the advances made in lexicography.

References

- Ahmad, K., H. Fulford and M. Rogers.** 1992. The Elaboration of Special Language Terms: The Role of Contextual Examples, Representative Samples and Normative Requirements. Tommola, H. et al. (Eds.). 1992. *EURALEX '92 Proceedings. Papers Submitted to the 5th EURALEX International Congress, Tampere, Finland, 4-9 August 1992*: 139-149. Tampere: Department of Translation Studies, University of Tampere.
- Alberts, M.** 1999. Terminology in South Africa. *Lexikos* 9: 18-35.
- Alberts, M.** 2010. National Language and Terminology Policies — A South African Perspective. *Lexikos* 20: 599-620.
- Antia, B.** 2005. Lexicography versus Terminology: Some Practical Reasons for Distinction. Presentation at the *International Training Workshop on Basic Principles of Terminology Management*, organised by the Department of Arts and Culture, Republic of South Africa. Johannesburg, 26-30 September, 2005.
- Bergenholtz, H.** 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Bergenholtz, H. and S. Nielsen.** 2006. Subject-field Components as Integrated Parts of LSP Dictionaries. *Terminology* 12(2): 281-303.
- Bergenholtz, H. and S. Tarp.** 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes. Journal of Linguistics* 31: 171-196.
- Bergenholtz, H. and S. Tarp. (Eds.)** 1995. *Manual of Specialised Lexicography*. Amsterdam: John Benjamins.
- Bowker, L. and J. Pearson.** 2002. *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge.
- Cabré, M.T.** 2000. Elements for a Theory of Terminology: Towards an Alternative Paradigm. *Terminology* 6(1): 35-57.
- Free Online Dictionary, Thesaurus and Encyclopedia.** <http://www.thefreedictionary.com>. Accessed on 6 November 2009.
- Fuertes-Olivera, P.A and A. Arribas-Baño.** 2008. *Pedagogical Specialised Lexicography. The Representation of Meaning in English and Spanish Business Dictionaries*. Amsterdam: John Benjamins.
- Fuertes-Olivera, P.A. (Ed.)** 2010. *Specialised Dictionaries for Learners*. Berlin: De Gruyter.
- Hartmann, R.R.K. and G. James.** 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Lawson, A.** 2001. Collecting, Aligning and Analyzing Parallel Corpora. Ghadessy, M., A. Henry and R.L. Roseberry (Eds.). 2001. *Small Corpus Studies and ELT: Theory and Practice*: 279-310. Amsterdam: John Benjamins.
- Madiba, M.** 2004. Parallel Corpora as Tools for Developing the Indigenous Languages of South Africa with Special Reference to Venda. *Language Matters* 35(1): 133-147.
- Madiba, M.** 2010. Fast-tracking Concept Learning to English as an Additional Language (EAL) Students through Corpus-based Multilingual Glossaries. *Alternation* 17 (1): 225-248.

- McArthur, T.** 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablets to the Computer*. Cambridge: Cambridge University Press.
- Meyer, I. and K. Mackintosh.** 1996. The Corpus from a Terminographer's Viewpoint. *International Journal of Corpus Linguistics* 1(2): 257-285.
- Moropa, K.** 2004. A Parallel Corpus as a Terminology Resource for Xhosa: A Study of Strategies Used to Translate Financial Texts. *Language Matters* 35(1): 162-178.
- O'Hara, S. and R. Pritchard.** 2009. *Teaching Vocabulary with Hypermedia*. Boston: Pearson.
- POINTER Project Final Report.** 1996. Collated by the POINTER Workpackage 7 team (K.Ahmad, R. Bonthron, G. Engel, A. Fotopoulou, D. Fry, C. Galinski, J. Humbley, N. Kalfon, M. Rogers, C. Roulin, K. Schmalenbach and E. Tanke) under the co-ordination of D. Fry. <http://www.computing.surrey.ac.uk/ai/pointer/report/section1>. Accessed on 12 April 2011.
- Sager, J.A.** 1984. Terminology and the Technical Dictionary. Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*: 315-326. Tübingen: Max Niemeyer.
- Sager, J.A.** 1996. *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.
- Sauer, H.** 2008. Glosses, Glossaries and Dictionaries in the Medieval Period. A.P.Cowie (Ed.). 2009: 17-40. *The Oxford History of English Lexicography. Volume I: General-Purpose Dictionaries*. Oxford: Clarendon Press.
- Sinclair, J.M.** 2001. Preface. Ghadessy, M., A. Henry and R.L. Roseberry (Eds.). 2001. *Small Corpus Studies and ELT: Theory and Practice*: vii-xv. Amsterdam: John Benjamins.
- Tarp, S.** 2000. Theoretical Challenges to Practical Specialised Lexicography. *Lexikos* 10: 189-208.
- Tarp, S.** 2002. Basic Elements of Lexicographic Theory. Emejulu, J.D. (Ed.). 2002. *Éléments de Lexicographie Gabonaise. Tome II*: 7-20. New York: Jimacs-Hillman.
- Tarp, S.** 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.
- Tarp, S.** 2008. *Lexicography in the Borderland between Knowledge and Non-Knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Max Niemeyer.
- Temmerman, R.** 2000. *Towards New Ways of Terminology Description: The Sociocognitive Approach*. Amsterdam: John Benjamins.
- University of Cape Town.** 1999 (revised 2003). *Language Policy*. Cape Town: University of Cape Town.
- University of Cape Town.** 2003. Language Plan — Towards a Language Plan for the University of Cape Town: 2005-2010. Cape Town: University of Cape Town.
- Wiegand, H.E.** 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*: 13-30. Tübingen: Max Niemeyer.

A Critical Analysis of the Lemmatisation of Nouns and Verbs in isiZulu

D.J. Prinsloo, *Department of African Languages, University of Pretoria,
Pretoria, South-Africa (danie.prinsloo@up.ac.za)*

Abstract: This article is a critical evaluation of lemmatisation strategies for nouns and verbs in isiZulu with specific attention to the problem of stem identification. The presumed target users of dictionaries compiled according to these lemmatisation strategies are non-mother tongue learners of isiZulu. The advantages versus disadvantages of lemmatising verbal and nominal stems, verbal and nominal stems without suffixes, and nominal words will be considered mainly in terms of the entire paradigm containing the verbal root *-sebenz-* from an isiZulu corpus. The conclusion reached is that word lemmatisation is preferred over both stem lemmatisation and lemmatisation of stems without suffixes. It will be argued that the problem of stem identification can only be solved in electronic dictionaries and the electronic dictionary *isiZulu.net* will be analysed in this regard.

Keywords: CONJUNCTIVE ORTHOGRAPHY, DISJUNCTIVE ORTHOGRAPHY, ELECTRONIC DICTIONARIES, LEFT-EXPANDED ARTICLE STRUCTURE, LEMMATISATION, STEM IDENTIFICATION, STEM LEMMATISATION, USER PERSPECTIVE, WORD LEMMATISATION

Opsomming: 'n Kritiese evaluering van die lemmatiseringstrategieë vir naamwoorde en werkwoorde in isiZulu. Hierdie artikel is 'n kritiese evaluering van die lemmatiseringstrategieë vir naamwoorde en werkwoorde in isiZulu met spesifieke aandag aan die probleem van stamidentifikasie. Die veronderstelde teikengebruikers van woordeboeke wat volgens dié lemmatiseringstrategieë saamgestel is word omskryf as nie-moedertaal leerders van isiZulu. Die sterk- versus swakpunte van stamlemmatisering vir werkwoorde en naamwoorde, werkwoord- en naamwoordstamme sonder suffikse en woordlemmatisering vir naamwoorde sal oorweeg word hoofsaaklik in terme van die volledige paradigma 'n willekeurig gekose wortel, naamlik *-sebenz-* in 'n isiZulu korpus. Daar word tot die gevolgtrekking gekom dat woordlemmatisering bo stamlemmatisering verkies word. Daar sal aangevoer word dat die probleem van stamidentifikasie slegs in elektroniese woordeboeke opgelos kan word en die elektroniese woordeboek *isiZulu.net* sal vir dié doeleindes ontleed word.

Slutelwoorde: DISJUNKTIEWE SKRYFWYSE, ELEKTRONIESE WOORDEBOEKE, GEBRUIKERSPERSPEKTIEF, KONJUNKTIEWE SKRYFWYSE, LEMMATISERING, LINKS-UITGEBREIDE ARTIKELSTRUKTUUR, STAMIDENTIFISERING, STAMLEMMATISERING, WOORDLEMMATISERING

Introduction

The publication of the first dictionary for isiZulu using a word strategy, instead of the traditional device of stem lemmatisation, reopens the debate on stem versus word lemmatisation in African languages. In particular, the question is whether the problem of stem identification — which proved to be the major stumbling block for learners to find lemmas in isiZulu dictionaries — has been solved. To date, most publications on lemmatisation in the African languages contrast disjunctively written languages (e.g., Sepedi, Setswana and Sesotho) with those with a conjunctive orthography (e.g., isiZulu, Siswati and isiXhosa) in order to indicate the advantages and disadvantages of stem as opposed to word lemmatisation. The main argument has been that stem lemmatisation is an accepted, or even the best strategy for conjunctively written languages, but that word lemmatisation is a better option for disjunctively written languages. The principal reason for this is that stem lemmatisation introduces unnecessary problems for the user of a dictionary of a disjunctively written language, especially with regard to the identification of nominal stems. The stem tradition, nevertheless, supported by certain assumptions, such as being the more scientific option gained such momentum that a number of stem dictionaries were compiled for the Sotho languages as well. Word lemmatisation for conjunctively written languages was considered by Van Wyk (1995) and preliminary experiments were conducted at some of the National Lexicography Units in South Africa on the feasibility and possible advantages of word lemmatisation for conjunctively written languages. However, it was only in 2010 with the publication of the *Oxford Bilingual School Dictionary: Zulu and English (OZSD)* that the almost sacred stem tradition of lemmatisation for an Nguni language was broken using word lemmatisation for an isiZulu dictionary.

The focus of this article differs from earlier research in the sense that first, the issue of stem identification takes centre stage, and secondly, that the advantages and disadvantages of stem versus word lemmatisation are not described in terms of conjunctively versus disjunctively written languages but rather in terms of the advantages and shortcomings of these approaches for the conjunctively written Nguni languages, isiZulu being a case in point. Thirdly, although a selection of examples is offered, the analysis of examples focuses on a paradigm of approximately 2 500 occurrences of different words containing the root *-sebenz-* 'work' in the *Pretoria isiZulu Corpus (PZC)*.

Thus the main aim of this article is to critically evaluate lemmatisation strategies for nouns and verbs in isiZulu with specific attention to the problem of stem identification. The prime objective is to evaluate lemmatisation strategies rather than isiZulu dictionaries *per se*. For critical reviews of the two prominent isiZulu dictionaries *Isichazamazwi sesiZulu (ISZ)* and the *OZSD*, see Masubelele (2007) and Prinsloo (2010), respectively. It should be borne in mind, however, that the choice of lemmatisation strategy may depend on the type of dictionary being compiled and the probable linguistic ability of its intended users.

A consolidation of the most prominent views on stem versus word lemmatisation, which lie scattered over a number of publications, is also attempted. Finally, the success or potential of electronic dictionaries to solve stem identification problems that cannot be solved in printed dictionaries, irrespective of the lemmatisation strategy, is evaluated.

Word forms of *-sebenz-* in the PZC and a brief explanation of key terms and concepts

One of the objectives of this article is to study the success of the different lemmatisation strategies on an entire paradigm for a randomly selected word and not only, as has traditionally been done in similar discussions on lemmatisation, by quoting examples in a haphazard way as they fit the author's viewpoint. By taking an entire paradigm of real language use of a word and its derivations as input to the study, strengths and especially weaknesses in the different lemmatisation strategies, which may have been overlooked by an idiosyncratic selection of examples, come to the fore. The paradigm of the verb root *-sebenz-* has been selected on the basis of its high frequency as a verb (*-sebenza* plus verbal prefixes 4 907, *ukusebenza* 548 times in the PZC); its frequent occurrence with suffixes, e.g., *-sebenzisa* (3 373); and also for the high frequency of occurrences of nominal derivations (deverbatives) of *-sebenza*, e.g., *umsebenzi* (5 883), *emsebenzini* (1 456), *imisebenzi* (1 009), *isisebenzi* (81) and *abasebenzi* (174). See the appendix for a list of the forms occurring five times and more in the PZC. The PZC is a raw corpus of approximately six million tokens.

Detailed discussions of the morphological system of isiZulu can be found in grammar books, such as Doke (1945) and in the mini-grammars of dictionaries, such as the *Zulu-English Dictionary* (ZED) and the OZSD.

In their *Zulu-English Dictionary* (ZED), Doke and Vilakazi (1948: xxiv-xxv) define *stem* as "that part of a word depleted of all prefixal inflexions" and *root* as "the irreducible element of a word; the primitive radical form without prefix, suffix or other inflexion, and not admitting of analysis".

In *conjunctively written languages*, such as the *Nguni languages* isiZulu, isiNdebele, Siswati and isiXhosa, most word forms (tokens) contain verbal or nominal roots with affixes (prefixes or suffixes, or both) and they are written as one orthographic word. Consider the examples in (1):

(1)

sebenza 'to work' (verbal root *-sebenz-* + verbal ending *-a*)

sebenzisa 'make use' (verbal root *-sebenz-* + causative suffix *-is-* + verbal ending *-a*)

basebenza 'they work' (class prefix class 2 *ba-* + verbal root *-sebenz-* + verbal ending *-a*)

basebenzisa 'they use' (class prefix class 2 *ba-* + verbal root *-sebenz-*, + causative suffix *-is-* plus verbal ending *-a*)

isineke 'patience' (class prefix class 7 *isi-* + nominal stem *-neke*)

indlovukazi 'queen' (class prefix class 9 *in-* + nominal stem *-dlovu* + augmentative/feminine suffix *-kazi*)

nendlovukazi 'with/and the queen' (conjunctive *na-* + class prefix class 9 *in-* + nominal stem *-dlovu* + augmentative/feminine suffix *-kazi*)

As a prerequisite to subsequent discussion, a brief but more corpus-based analysis is given with the focus on the chosen paradigm of *-sebenz-*. The nominal forms of *-sebenz-* occur with prefixes and suffixes and the combination of such affixes, e.g. *emsebenzini* (1 456) (*e+umsebenz(i)+ini*) 'at work' <*umsebenzi* 'work', *nomsebenzi* (309) (*na+umsebenzi*) 'and the work', *ngomsebenzi* (253) (*nga+umsebenzi*) 'with work', *imisebenzana* (21) (*imisebenzi+ana*) 'little tasks' and *imisebenzikazi* (1) (*imisebenzi+kazi*) 'big jobs'. Affixes and combinations of affixes occurring with the verbal forms of *-sebenza* include the infinitive form *ukusebenza* (548) 'to work, working', subject concords, e.g., *usebenza* (299) 'he/she works', causative suffix, *-sebenzisa* (219) (*-sebenza+isa*) 'use, utilize', infinitive and future *ukusebenzelana* (8) 'working for each other' *ukuyosebenza* (49) (*uku+yo+sebenza*) 'to go and work', combination of concords and auxiliary *wawusebenza* (10) 'it was working'.

The conjunctive way of writing consequently results in very long words; the average word length of isiZulu words (tokens) in the PZC is 6.93 characters, cf. (2a). In *disjunctively written languages*, such as Sepedi, Setswana, Sesotho, Tshivenda and Xitsonga, nouns, verbs, concords, etc. are written as separate orthographic words, e.g., as in (2b). By contrast, the average word length for Sepedi words in the *Pretoria Sepedi Corpus* (PSC) is a mere 3.88 characters.

(2)

- a. Angifuni ukusebenza (isiZulu) 2 linguistic words, 2 orthographic words
 b. Ga ke rate go šoma (Sepedi) 2 linguistic words, 5 orthographic words
 'I do not want to work'

A popular definition of *lemmatisation* is the selection of a canonical form to represent a specific paradigm. A clear though simplified example is that *walk* is chosen as lemma to represent the paradigm *walk, walks, walked, walking*.

Stem lemmatisation refers to the selection of the verbal stem *-sebenza* from verbal forms, such as *ukusebenza* 'to work, working', *usebenza* 'he/she works', *wawusebenza* 'it was working', for instance, as the canonical form for lemmatisation purposes. For nominal forms, the nominal stem *-sebenzi* is selected for *umsebenzi* 'work, worker', *emsebenzini* 'at work', *nomsebenzi* 'and the work', *ngomsebenzi* 'with work', etc. In terms of stem lemmatisation a distinction is drawn between stem lemmatisation and left-expanded stem lemmatisation as opposed to word lemmatisation.

Word lemmatisation would have the lemma *umsebenzi* for the paradigm *umsebenzi, emsebenzini, nomsebenzi, and ngomsebenzi*.

The user perspective

As echoed in many publications, e.g., Hartmann (1989), Gouws and Prinsloo (2005a) and (2005b), contemporary lexicography is dominated by a user-driven approach. Consequently, all arguments in this article depart from the users' needs, and/or their reference skills and ability to find lemmas in isiZulu dictionaries. The target users in mind for this discussion on lemmatisation strategies in isiZulu dictionaries are learners of isiZulu with text production needs as well as the need for text reception of the prescribed books for isiZulu readers.

It may be stated at the outset that the inability of users to identify nominal and verbal stems can impede successful word searches or even result in the total failure to look up words in isiZulu dictionaries successfully. The situation is aggravated by the reality in Africa that users generally lack a dictionary culture and dictionary using skills (cf. Gouws and Prinsloo 2005a: 42).

The stem versus the word tradition in lemmatisation

Bennett (1986) as quoted by De Schryver (2010: 163) rightfully points to the complexity of nouns and verbs in African languages and asserts that stem identification can be problematic.

There has been debate as to the proper arrangement of the Bantu lexicon, and the question is far from settled. The inflection of nominals and verbals by means of prefixes, and the complex and productive derivational system, both characteristic of Bantu languages, pose difficulties [...] If items are alphabetized by prefix [...] a verb will be listed far from its nominal derivations, however transparent these may be. [...] A competing school arranges the lexicon by stem or root; this usefully groups related items, and saves on cross-referencing. Unfortunately, in such a system the user must be able to identify the stem, which given the sometimes complex morphophonemics of Bantu languages may not be easy. Bennett (1986: 3-4)

Van Wyk (1995) puts the issue of stem versus word lemmatisation in perspective in relation to disjunctively versus conjunctively written languages. Van Wyk (1995: 82) notes that two lexical traditions exist in the African languages in South Africa, i.e., the *word* tradition and the *stem* tradition:

According to the word tradition, lemmas are based on complete written words, and there is a one-to-one correspondence between written words and lemmas. According to the stem tradition, lemmas are based on the stems of written words without their prefixes.

Subsequent publications dealing with problematic aspects of lemmatisation in African languages, such as Prinsloo (1994), Prinsloo and Gouws (1996), Prinsloo and De Schryver (1999) and Prinsloo (2009) have departed from Van Wyk's pioneering analysis. The fundamental issues raised in Van Wyk's 1995 study

are not discussed in any detail here — only aspects that are relevant for this article are briefly outlined as a basis for the discussion that follows.

Van Wyk (1995) begins by dismissing the claim that the stem tradition is in any way superior to the word tradition. He states:

[...] many lexicographers have come to the erroneous conclusion that only the stem tradition is linguistically justified. Ziervogel [...], for example, claims that it is scientifically sound, and Ziervogel and Mokgokong [...] state categorically that it is the only scientific method. (Van Wyk 1995: 84).

Then, he refutes the claim that stem lemmatisation is more economical, in terms of dictionary space, than word lemmatisation. Thirdly, he highlights the misconception that verbal affixation and nominal affixation are equally productive and therefore necessitate stem lemmatisation for nouns. Finally, he highlights the problems in respect of stem lemmatisation, especially in the case of some class 9 nouns where neither the lexicographer nor the user can identify the stem. For example, stem identification is very problematic in cases such as *intaba* 'mountain', *intombi* 'girl', *inkosi* 'king' and *inkabi* 'ox'. The uninitiated user would not know that the stem form of *intaba* is *-ntaba*, but for *intombi* it is *-thombi*, for *inkosi*, *-khosi* and for *inkabi*, *-nkabi* in order to look these terms up. Mpungose (1998: 65) agrees by saying that the process of lemmatising nouns in classes 9 and 10 is problematic and he refers to the traditional method as lemmatising the "lexical noun by etymological noun-stem". Mtuze (1992: 17), in reference to nominals of Class 9 and Class 10 in isiXhosa, bluntly states:

You never knew how these nominals were lemmatised [...] In some cases, you had to struggle trying to look up words such as *ingulube* (the wild pig) as the entry could either be under *g* or under *n*.

There is, fortunately, no dispute regarding stem versus word lemmatisation in the case of verbs. Followers of both traditions agree that verbs should be lemmatised on their stems. Sources, such as Prinsloo (2009) debating the issue, consequently focus on nominal stem identification as the problematic area. However, it should be emphasized that the fact that both traditions agree on stem lemmatisation for verbs does not alleviate the problem of stem identification. It is argued here that stem identification for verbs in Nguni languages is as problematic for verbs as it is for nouns — the identification of *-sebenza* from the numerous verbal forms, or *-sebenzi* from the nominal forms in the paradigm of *-sebenz-* is proof thereof. From the 31 orthographic forms in the appendix occurring more than 100 times in the PCZ, 15 are verbs and 16 are nouns. The challenge to identify the stem is exactly the same for nouns and verbs.

The advantages and disadvantages of lemmatising stems

This strategy entails the lemmatisation of nominal and verbal stems with or without suffixes. For verbs, all suffixes can be removed rendering the verbal

root, e.g., *-sebenz-* plus verbal ending *-a*, *-sebenza* as lemma and also for words such as *ukusebenza* (548), *usebenza* (299) and *esebenza* (290) from the appendix. Verb stems can also be lemmatised with their suffixes, e.g., *-sebenzisa* for the paradigm *ukusebenzisa* (244), *-sebenzisa* (219) and *ngokusebenzisa* (199). For nouns, the result of affixal stripping will render the nominal stem *-sebenzi* as lemma for *ngomsebenzi* (253), *izisebenzi* (242), *ngumsebenzi* (179). Consider verbal stem without suffix(es), verbal stem with suffix(es), and noun stem lemmatisation in ZED in (3).

(3) ZED

sebenza (3.2.9) v. [> perf. *-sebenzile*; pass. *setshenzwa*; neut. *sebenzeka*; ap. *sebenzela*; rec. *sebenzana*; caus. *sebenzisa*; int. *sebenzisa*; dim. *sebesebenza*; *insebenzo*; *isisebenzi*; *umsebenzi*; *umsebenzi*.]
 1. Work, labour; work at. *ukusebenza emasimini* (to work in the fields). *Usebenza imithi* (He is a druggist). *Usebenza izimoto* (He is a motor-mechanic). *Woke usebenze* (You will have a tough job).
 2. Administer medical treatment to. *ukusebenza umuntu ogulayo* (to treat a sick person).
-sebenzi (*isisebenzi*, 2.2-4.4.3.9, *izisebenzi*) n. [< *sebenza*.]
 Servant, employee; workman.
-sebenzi (*umsebenzi*, 2.2-4.4.3.9, *abasebenzi*) n. [< *sebenza*.]
 Workman, labourer.

-sebenzi (*umsebenzi*, 2.2-4.4.3-8.9, *imisebenzi*) n. [< *sebenza*.]
 Work, job, occupation. *Ngifuna umsebenzi* (I want work).
sebenzisa (3.3.2.9) v. [caus. <sebenza. > perf. *-sebenzisile*; pass. *setshenziswa*; neut. *sebenziseka*; ap. *sebenzisa*; rec. *sebenzisa*.]
 1. Cause to work, help to work, work with.
 2. Use, utilize. *ukusebenzisa umshanelo* (to use a broom).
-sebenzo (*insebenzo*, 2.4.3.9, *izinsebenzo*) n. [< *sebenza*.]
 Wages, payment for work. *Bathola insebenzo yabo ngesonto* (They get weekly wages).

An advantage of stem lemmatisation is that it is the undisputed option for the lemmatisation of verbs, not only for conjunctively written languages but also for the disjunctively written ones. Van Wyk (1995: 85) states clearly that "except for the notational device of a hyphen [...] the entry for 'see' will [...] be found as **bona** in word dictionaries and as **-bona** in stem dictionaries".

For nouns, stem dictionaries normally provide the stem and the possible nominal prefixes in brackets, e.g., *-sebenzi* (um-, aba-, imi-) and the user can rightfully conclude that the forms are *umsebenzi*, *abasebenzi* and *imisebenzi*. The lexicographer must of course make sure that the possible combinations suggested by this notation, i.e., *umsebenzi*, *abasebenzi* and *imisebenzi* are correct. An example where it is not the case is the lemmatisation of *inkosi/amakhosi* 'king(s)', in *Woordeboek Afrikaans-Zoeloe, Zoeloe-Afrikaans* (WAZ) in (4).

(4) WAZ

-khosi, (in-, ama-), b; 1. koning, regent, hoofman

The user's conclusion is: **inkhosi* 'king', *amakhosi* 'kings', of which the plural form is correct but the singular form is incorrect — it must be *inkosi*. This is a very serious mistake, since the dictionary should never guide the user to such incorrect conclusions.

Van Wyk (1995), however, has fundamental problems with stem lemmatisation for nouns. In his view, the first disadvantage stems from the misconception that nominal and verbal stems are equally productive in combining with prefixes. Van Wyk (1995: 87) quite correctly shows that verbs stems may, for

instance, productively combine with all the subject concords, object concords, negative morphemes and modal morphemes, $18 \times 19 \times 6 \times 2$ which comes to 4 104 possible combinations. Noun stems can only be used with a small number of class prefixes. In the case of nominalizations of *-sebenz-*, nouns occur in classes 1, 2, 3, 4, 7 and 8 as in (5a). The starred forms in (5a) indicate ungrammatical combinations in terms of the class prefix paradigm for isiZulu. For other nouns, the number of possible combinations can be even less, as indicated in (5b).

(5)

- a. Class 1 and 3: *umsebenzi* 'worker, work', Class 2: *abasebenzi* 'workers', Class 4: *imisebenzi* 'work, jobs', Class 5: **i(li)sebenzi*, Class 6: **amasebenzi*, Class 7: *isisebenzi*, 'employee' Class 8 *izisebenzi* 'employees', Class 9: **insebenzi*, Class 10 **izinsebenzi*, Class 11: **u(lu)sebenzi*, Class 14: **ubusebenzi*
- b.
- ntu (umu-, aba-, isi-, ubu-, u(lu)-): *umuntu* 'a human', *abantu* 'humans', *isintu* 'Bantu culture', *ubuntu* 'humaneness', *untu* 'common people'
 - khosi (in- (inkosi), ama-): *inkosi* 'a king', *amakhosi* 'kings'
 - khaya (i(li)-, ama-, um-): *ikhaya* 'home', *amakhaya* 'homes', *umkhaya* 'members of the family'
 - tho (isi-, izi-, in-, izin- (u(lu)-): *isitho* 'limb', *izitho* 'limbs', *into* 'thing', *izinto* 'things', *utho* 'something'
 - daba (in-, izin-, u(lu)-): *indaba* 'story', *izindaba* 'stories', *udaba* 'a serious affair'

Thus Van Wyk (1995) concludes that there is no linguistic justification for treating nouns and verbs in the same way in terms of stem lemmatisation. Van Wyk's criticism is valid if the view is restricted to the consideration of concords in terms of verb stems and class prefixes in relation to noun stems. If, however, the complex orthographic forms of nouns and verbs are considered, e.g., as for *-sebenz-* in the appendix, then noun stems and verbs stems are on a par in terms of productive combination with affixes, such as the conjunctives *na*, in *nomsebenzi* (*na+umsebenzi*) 'and the work', *nga* in *ngomsebenzi* (*nga+umsebenzi*) 'with work' and the possessive *sa* in *somsebenzi* (*sa+umsebenzi*) 'of the work' or with the combinations of affixes. The distinction between verbal stem identification and nominal stem identification therefore effectively falls away for the target users. The question could therefore be asked: if the user has to deal with affixation on such a massive scale anyway in his/her effort to find the lemmas for nouns and verbs in the dictionary, why not depart from the stem in all instances for nouns and verbs?

Identifying the stem remains the underlying challenge for the discussion on the following two lemmatisation strategies: stems lemmatised with their prefixes, and word lemmatisation.

An issue of special relevance for a critical analysis of stem and word lemmatisation in the Nguni languages is the lemmatisation of infinitives. In all of

the lemmatisation strategies for isiZulu, i.e., stem, left-expanded and even word lemmatisation, verbs are by default lemmatised as stems. Linguists agree that the infinitive has characteristics of both nouns and verbs. Infinitives such as *ukuhamba* 'to walk, a/the walking', *ukusebenza* 'work, a/the working' and *ukukhuluma* 'speak, a/the speaking' are nouns (of class 15) and verbs at the same time. In traditional grammars the infinitive is therefore positioned and formally described within the two major categories of *verb* and *noun*. Consider (6a) in contrast to (6b) where the verbal versus nominal meanings of the infinitive are foregrounded.

(6)

- a. Angithandi ukuhamba ngezinyawo 'I do not like *to walk* on foot'
- b. Ukuhamba kuyakhathaza '*Walking* is exhausting'

Unlike the other noun classes, the stem of the infinitive noun is not a nominal, but a verbal stem and, unlike verbs, infinitives contain a noun class prefix (class prefix of Class 15). In dictionaries following a stem lemmatisation strategy, such as the ZED, all infinitives are lemmatised under their stem forms, e.g., the lemmas *-hamba*, *-sebenza* and *-khuluma*. No effort towards the lemmatisation of the nominal forms *ukuhamba*, *ukusebenza* or *ukukhuluma* is made in the ZED and no effort towards treating nominal meanings in the articles of *-hamba*, *-sebenza* and *-khuluma* has been made. Ironically, the advocates of stem lemmatisation are forced to lemmatise full words with *uku-*, e.g., in the case of the ZED for *ukuthi* (conjunctive) 'so that', *ukufa* (interjection) 'how magnificent!', *ukuba* (conjunctive) 'that (after verbs of knowing, etc.)', *ukuphela* 'only that', etc. These words belong to different parts of speech and the lemmatisation strategy could be justified. The problem lies with the fact that users are conditioned to ignore infinitive prefixes in the process of determining the lemma, i.e., not to consider the *uku-* and therefore they will look up *-thi*, *-fa*, *-ba* and *-phela* and indeed find such lemmas with treatment in the ZED without any cross-reference to the lemmatised full forms *ukuthi*, *ukufa*, *ukuba* and *ukuphela*.

Ukuthi, as a conjunctive or connective, is the most frequently used word in isiZulu and the lack of a cross-reference could simply mean that the users will not find the meaning of *ukuthi* 'so that, in order that' representing 90% of its use. IsiZulu dictionaries, such as the *Compact Zulu Dictionary* (CZD) and the *English and Zulu Dictionary* (EZD) do not handle this issue satisfactorily by either not lemmatising the conjunctive or not giving a cross-reference.

(7)

<p>a. CZD -thi (v) (ukuthi), to say.</p>	<p>b. EZD -thi defect. V. 1. say. 2. mean, intend; ... ukuthi conj. 1. (foll. by indic.) that. 2. (foll. by subjunct.) so that, in order that.</p>
--	---

The strengths and shortcomings of stems lemmatised with their prefixes

This is the lemmatisation strategy employed by dictionaries, such as the ISZ, the *Concise SiSwati Dictionary* (CSD), the *Dictionary of the Tebele & Shuna languages* (DTS) (the latter as quoted by Gauton (forthcoming)), and even for a disjunctively written language, Sesotho, in the *Southern Sotho-English Dictionary* (SSED) where stems are lemmatised but the full form of the word is given. In the ISZ, for verbs, it entails lemmatising the stem and presenting it in boldface but adding the infinitive prefix *uku-* in italics, e.g., *ukusebenza* will be lemmatised as *ukusebenza* in the alphabetical stretch S of the dictionary. Likewise, the noun *umsebenzi* will be lemmatised as *umsebenzi* under S by giving its full form with the stem in bold and the class prefix in italics as in (8).

(8)

íkúsebéntza (sz) [-an-; -el-; -ek-; -is-; -w-] - 1. ukwenza okuthile ngenhloso yokuthola umkhiqizo noma umphumela othile. *Ubaba usebenza eGoli.* 2. ukwelapha. *Inyanga yabe isebenza isiguli.* 3. isimo sokufakazela ukuphila kokuthile. *Ungalibali ukufonela uZitha, phela seluyasebenza ucingo lwakhe.*

isísebenzi (bz) - umuntu oqashiwe owenza umsebenzi othile; um(u)sebenzi; um(u)qashwa.

ím(ú)sebenzi (bz) - 1. okwenziwa ngenhloso yokuthola umkhiqizo othile noma umphumela othile. *Umsebenzi usihlasele kulezi zinsuku.* 2. idili elenziwa ekhaya njengokubuyisa osashona.

ím(ú)sebenzi (bz) - oqashelwe ukwenza umsebenzi othile; isisebenzi; um(u)qashwa.

íkúsebéntzisa (sz) [-an-; -el-; -ek-; -is-; -w-] - ukwenza okuthile kube kukhona okukusizayo ukwenza lokho okwenzayo; ukusebenza ngokuthile. *Ukusebenzisa isikele ukugunda intshebe yakhe yikhona okufanelekile.*

Adding the prefix has certain advantages, e.g., the reassuring factor (in cases of a 1-1 match), e.g., the user wants to look up *isidaka* and looks for *-daka* and finds all the different full nominal forms including *isidaka*. So (s)he knows that the process of information retrieval has been successful. The ISZ goes even further by implicitly giving morphological information about the prefix. Examples, such as *ú(lu)sinsi* 'hair growing low on the forehead' *í(lí)fasíkoti** 'apron' *í(lí)bhoklóló*° 'a brave, confident male', contain additional information, i.e., that the full form of the prefixes are, respectively, *ulu-* and *ili-*. Presenting the full word also enables the indication of tone. The symbols 't', '*' and 'o' following the lemma indicate that no plural form exists for *ú(lu)sinsi*; that *í(lí)fasíkoti* is a loan word; and that *í(lí)bhoklóló* is a neologism. Apart from giving the infinitive class prefix with verbal stems, verbal suffixes are given in brackets following the stem, i.e., not as separate lemmas, e.g., *úkúlobola* ... [-an-, -el-, -ek-, -is-, -w-]. Indicating a number of frequently used verbal suffixes in this way does enhance the comment on form but does not contribute to the comment on semantics. For the user, it means that stems containing these suffixes have to be

looked up under the basic stem, and the meanings conveyed by these suffixes then have to be added on. It also gives insufficient guidance in cases where sound changes occur as a result of affixation. As Masubelele (2007: 460) rightfully remarks, "variants of words which are the result of phonological processes, such as the passive construction have been omitted, e.g., *úkúlobola* which changes to *úkúlotsholwa* in the passive". The same holds true for the passive form of *-sebenza* where the inexperienced user is unlikely to link the passive form *-setshenzwa* with *-sebenza+w*. It would be better to lemmatise and treat derivations containing these suffixes as in (3).

Gouws and Prinsloo (2005b: 29) refer to the lemmatisation strategy where stems are lemmatised, but the full form of the words are given as left-expanded article structures.

[...] a left-expanded procedure [...] can [...] accommodate the prefixal element in a slot preceding the stem. This phasing out of the prefixal element to the article-initial position does not, however, change the status of the lemma sign as guiding element of the article because the lemmatization is still done according to an initial-alphabetical ordering in which the stem is the alphabetical point of reference.

Consider the following examples from CSD in (9).

(9) a. CSD

- (kú)-hamba v. walk, go, travel, move,
↓
leave. (For going to ... cf. -ya).
- (kú)-hambela v. t. 1. travel for or on
behalf of. 2. visit. 2. press a claim.
- (kú)-hambelana v. 1. go with each other,
travel together. 2. be on good terms
- si-hambi / ti- n. visitor, tourist stranger.
- um-hambi / bá- n. traveller.
- (kú)-hambisa v. t. 1. drive, cause or
help to move, send off. 2. bid fare-
well to; send greetings to. 3. purge.
- (kú)-hambisana v. associate, accompany
- lu-hambo / ti- n. journey

b.

↓
<i>kuhamba</i> v. walk, go ...
<i>kuhambela</i> v.t. ...
<i>kuhambelana</i> v. ...
<i>sihambi</i> n. ...
<i>umhambi</i> n. ...
<i>kuhambisa</i> v.t. ...
<i>kuhambisana</i> v. ...
<i>luhambo</i> n. ...
↓

The verb stems *-hamba*, *-hambela*, *-hambelana*, *-hambisa* and *-hambisana* are lemmatised with the infinitive class prefix *ku-*. Full nouns, i.e., *sihambi*, *umhambi* and *luhambo* are given, but they are lemmatised on their stem forms *-hambi* and *-hambo*. As in the case of the verbs, the alphabetization is done on H, the first letter of the stem as indicated by the arrows in (9a). Gouws and Prinsloo (2005b) suggest that the first letters of the stems should be vertically aligned to visually strengthen the alphabetical alignment on H in (9b). In the ISZ, vertical alignment on the first letter of the stem has also not been done as in (8), but indentation of the amount of space equal to approximately three characters and the contrast between italics for prefixes and boldface for lemmas provide for a user-friendly layout and alleviate the need for vertical alignment.

Left-expanded article structures can in principle be extended to go beyond

left-expansion of class prefixes to other types of prefixes and prefixal combinations, such as conjunctives and concords. The lexicographer could for example decide to lemmatise words that occur with a high frequency in the corpus, such as *nokusebenza* (121) 'and to work', *ngokusebenza* (56) 'by working', *wayesebenza* (57) 'he/she/it was working' *ukuyosebenza* (49) 'to go and work' in (10).

(10)

nokusebenza
ngokusebenza
wayesebenza
ukuyosebenza

On the one hand, the lemmatisation strategy in the ISZ and the CSD shows characteristics of stem lemmatisation in the sense that alphabetical ordering runs on the first letter of the stem, thus ignoring the different nominal and verbal prefixes. On the other hand, it resembles a word dictionary since full words, i.e., the full infinitive form of, e.g., the verb, *kuhamba* 'to go' and the full form of the noun *sihambi* 'visitor/tourist' are lemmatised. An even closer resemblance to word dictionaries is found in the ISZ's layout, e.g., by not putting the infinite prefix in brackets or separating the prefix from the stem by means of a hyphen as was done in the CSD but using the normal orthography. Masubelele (2007: 459) quotes the following paradigm: *í(lí)daka* 'dry cattle dung', *ísídaka* 'black soil', *ú(lú)daka* 'mud', *úkú)daka* 'to become drunk' and *úm(u)daka* 'heavy, brown bracelet, bestowed as royal honour'. Advocates of the word tradition for lemmatisation will be quick to point out that no gain in terms of space saving is achieved in such cases. Giving the prefix has certain advantages. First, the reassuring factor should not be underestimated. It is of special value in the case of those class 9 nouns quoted by Van Wyk (1995: 90) where lexicographer and user have difficulty in identifying the stem form. Secondly, utilizing the opportunity to give additional morphological and tonological information is a positive aspect, provided that the user is familiar with the tonal markers. The convention used to indicate the full form of the prefix, however, carries the risk of misinterpretation that, e.g., both *usinsi* and **ulusinsi*; *ifasikoti* and **ilifasikoti*; and *ibhoklolo* and **ilibhoklolo* are grammatical, because this convention normally suggests the part in brackets as being an alternative. Thirdly, the symbols '+' for indicating that the noun does not have a plural form or '‡' in the case of certain plurals not having a singular form is a positive. This convention has, however, to be weighed against the convention '(x/y)' as, for example, used in the OZSD where both singular and plural class numbers are given and related and the appropriate class to which the word belongs is indicated by the boldface as in (14).

Masubelele (2007: 459) regards the fact that no plural indication is given in the ISZ as a problem, "since only a singular noun prefix is given with each stem, this might be problematic, especially to users who are not mother-tongue

speakers, because they might not know what the plural form of the specific word is". The question, however, is whether the user who wants to find the meaning of *umsebenzi* 'worker' will be interested to know what the plural form is? The problem is rather the amount of knowledge required from the user to look up singular or plural forms in dictionaries employing a left-expanded strategy and whether this strategy contributes in any way to resolving the problem of stem identification. The answer is *no*, although this lemmatisation strategy provides for user-friendly elements, such as additional information and reassurance, stem identification still has to be done.

Returning to the issue of infinitives: in dictionaries following a left-expanded lemmatisation strategy, the lemmas will (for example) be *ukuhamba*, *ukusebenza* and *ukukhuluma*, respectively. In the case of *ukufa* 'die, death; how magnificent!' and *ukuthi* 'say, a/the saying; so that' all three semantic distinctions in each case are accounted for and accommodated together in, say, two *subsequent* lemmas as in (11) and (12).

(11) <i>ukufa</i> ... inf 1. to die ... 2. death .. <i>ukufa</i> ... interjection ... how magnificent!	(12) <i>ukuthi</i> ... conjunctive ... <i>ukuthi</i> ... inf 1. to say ... 2. saying ..
--	---

This layout represents an improvement to stem lemmatisation as argued in terms of (7) above, because nominal, verbal and other meanings are presented together.

Advantages and disadvantages of word lemmatisation

The title of De Schryver's (2010) text: *Revolutionizing Bantu Lexicography — A Zulu Case Study*, suggests that word lemmatisation has fundamentally transformed the lexicography of the African languages. Word lemmatisation for nouns, in word dictionaries where both singular and plural forms are lemmatised means that the full singular and plural forms will be lemmatised with alphabetical sorting on the first letter of the word, e.g., *umsebenzi* 'work, worker', *abasebenzi* 'workers' and *isisebenzi* 'employee', can be looked up directly under *u*, *a* and *i*, respectively.

(13) OZSD

umsebenzi *** *noun* 3/4 (pl. **imisebenzi**)
 C-SEBENZA ■ work; job • Kunzima ukuthola
umsebenzi kule mihla. *It is very difficult to get a job these days.*
 ◇ **umsebenzi wokuqondisisa**
 ■ comprehension test
abasebenzi ** *pl. noun* 7/2 See sg.
 ISISEBENZI

isisebenzi *noun* 7/2, 7/8 (pl. **abasebenzi, izisebenzi**) C-SEBENZA ■ worker; servant
 • Okhulumela **abasebenzi**, uMnu. Khulekani Mzimela. *The one who speaks on behalf of the workers is Mr Khulekani Mzimela.*

The OZSD goes beyond the lemmatisation of the basic nouns *umsebenzi*, *abasebenzi*, *imisebenzi*, *isisebenzi* and *izisebenzi*, cf. (5a) above, and offers articles for derived forms, such as *ekusebenzeni* 'in the working', *ekusebenziseni* 'in using', *ekusetshenzisweni*, *emisebenzini* 'at work' and *emsebenzini* 'at work'.

Word lemmatisation also solves the difficulties mentioned by Mtuze (1992), Van Wyk (1995) and Mpungose (1998); and for those words where it is difficult to identify the stem, e.g., *intaba*, *intombi*, *inkosi* and *inkabi*. In a dictionary where full nouns are lemmatised, the problem is avoided by the lemmatisation of these forms exactly as they are. So, if the user is given the nominal form, access to the lemma is straightforward and easy.

A typical argument against the lemmatisation of the full forms of nouns as echoed by Van Wyk (1995: 95) is that the alphabetical stretches, especially U, A, I into which nouns fall, will be very large, because nouns in classes 1, 3, 11 and 14 begin with *u-*, classes 2 and 6 with *a-*, classes 4, 5, 7, 8, 9 and 10 with *i-*. Van Wyk's estimate for isiZulu is U: 18%, I:20% and A:5%. In the OZSD, *i-* takes up 62 pages, representing 23.5%, i.e., almost a quarter of the dictionary, *u-* 40 (15.2%) pages and *a-* 14 pages (5.3). However, users are unlikely to find this at all disturbing as can be judged by looking at a typical example taken from the alphabetical stretch *isi-* in (14). In the *Collins COBUILD English Dictionary* (COBUILD), the alphabetical stretch CON- is almost 30 pages long and to the best of our knowledge no complaints have been voiced in this regard.

(14) OZSD

isiZulu–isiNgesi 113	isisebenzi → isitayela
<p>rallying cry "The Power is Ours" means that the power must be with the people.</p> <p>isisebenzi noun 7/2, 7/8 (pl. abasebenzi, izisebenzi) C -SEBENZA ■ worker; servant • Okhulumela abasebenzi, uMnu. Khulekani Mzimela. <i>The one who speaks on behalf of the workers is Mr Khulekani Mzimela.</i></p> <p>isisekelo noun 7/8 (pl. izisekelo) C -SEKELA ■ support; foundation; basis • Uphenyoluthole ukuthi izinsolo eMhlali azinaso isisekelo. <i>The investigation found that the complaints in Mhlali are without foundation.</i></p> <p>◊ gotha imbokodo nesisekelo ■ destroy completely • UShaka wayeseqothe imbokodwe nesisekelo kuzo zonke izizwe ayakhelene nazo. <i>Shaka completely destroyed all the surrounding tribes.</i></p> <p>◊ umthetho-sisekelo ■ constitution</p> <p>isishayamthetho noun 7/8 (pl. izishayamthetho) C -SHAYA + UMTHETHO ■ legislature • Isishayamthetho sesifundazwe sikhethelwa iminyaka emihlanu. <i>A provincial legislature is elected for a term of five years.</i></p> <p>isishiyagalolunye noun 7/- ■ nine</p>	<p>on a programme to lose weight, fibre can prevent you becoming hungry quickly.</p> <p>isisinga noun 7/8 (pl. izisinga) ■ rope (used for tying a calf, goat, ...)</p> <p>◊ inkonyane isethuka isisinga ■ not yet used to • Inkonyane yafike yethuka isisinga kuNompumelelo nxa efika esikoleni esikhulu saKwaMagwaza ngakho wafike waqaphelisisa konke okwenzekayo wakubhala encwadini yakhe. <i>Nompumelelo was not yet used to it when he came to the high school in Magwaza; that's why he first took careful note of everything that happened and wrote it down in his book.</i></p> <p>isisombululo noun 7/8 (pl. izisombululo) ■ solution • Unqamule uhambo lwakhe lwaseNdiya ngesonto ukuze ahlngane nezikhulu zalo mnyango emizamweni yokuthola isisombululo kule nkinga. <i>She shortened her visit to India by a week, in order to meet with the senior officials of the department in an effort to find a solution to the problem.</i></p> <p>isisu ** noun 7/8 (pl. izisu) 1 ■ stomach; belly; abdomen • Waze wafika emanzini,</p>

A second argument against the lemmatisation of the full forms of nouns pertains to the lemmatisation of plural forms of nouns, first, in terms of the additional space in the dictionary taken up by these lemmas and, secondly, that it results in overuse of the mediostructure (cross-referencing system), because all such lemmas function as cross-references to the singular forms. It cannot be denied that lemmatising plural forms takes up a great deal of additional space. However, in terms of the reassuring aspect mentioned above as well as the amount of information carried by these skeleton dictionary articles as in (15), their inclusion in the macrostructure could be justified. First, the user is reassured that (s)he is dealing with the correct lemma; secondly, information on the frequency of use is indicated (by means of e.g. *, ** and ***); thirdly, noun class information is provided; and finally a cross-reference is given to the singular form where full treatment is offered.

(15) OZSD

imiphakathi * <i>pl. noun 3/4</i> See sg. UMPHAKATHI	iziphuzo <i>pl. noun 7/8</i> See sg. ISIPHUZO
imiphefumulo <i>pl. noun 3/4</i> See sg. UMPHEFUMULO	iziqalo <i>pl. noun 7/8</i> See sg. ISIQALO
imiphumela ** <i>pl. noun 3/4</i> See sg. UMPHUMELA	iziqephu <i>pl. noun 7/8</i> See sg. ISIQEPHU
imiqhudelwano <i>pl. noun 3/4</i> See sg. UMQHUDELWANO	iziqu * <i>pl. noun 7/8</i> See sg. ISISU
imiqondo * <i>pl. noun 3/4</i> See sg. UMQHONDO	iziqubulo <i>pl. noun 7/8</i> See sg. ISIQUBULO
imisakazo <i>pl. noun 3/4</i> See sg. UMSAKAZO	izisebenzi <i>pl. noun 7/8</i> See sg. ISISEBENZI
...	izisho <i>pl. noun 7/8</i> See sg. ISISHO
imisebenzi *** <i>pl. noun 3/4</i> See sg. UMSEBENZI	izisu <i>pl. noun 7/8</i> See sg. ISISU
	izitatimende <i>pl. noun 7/8</i> See sg. ISITATIMENDE
	iziteshi <i>pl. noun 7/8</i> See sg. ISITESHI
	izitha * <i>pl. noun 7/8</i> See sg. ISITHA

Consider now the presumed or likely dictionary needs of a learner of isiZulu who wants to use the dictionary for text production. First, a typical situation in class is considered where the learner is instructed to find the meaning of a number of isiZulu words, say nouns, or to write an essay on *abasebenzi* 'workers'. In the latter case, the users simply take the dictionary, look for the lemma *abasebenzi* and the worst that can happen is of (s)he having to follow up a cross-reference to the singular form *umsebenzi* where appropriate treatment is offered. No problematic stem identification is required as in the case of stem or left-expanded lemmatisation discussed in the previous paragraphs. However, a substantial part of learners' needs is to find the meanings of words used in their prescribed books, especially isiZulu literary works, such as novels, poetry and prose.

This means they are from the outset confronted by the full/complex orthographic forms which can be more than 2 000, e.g., in the case of *-sebenza*. In

order to find the meaning of *nomsebenzi* (309), *ngomsebenzi* (253), *emsebenzini* (1 456), *somsebenzi* (112), etc. (s)he has to *identify the noun*. This has to be done principally by stripping off affixes. Even more problematic is where (s)he has to *add* characters to the word in order to reconstruct the full noun in order to look it up, e.g., *umsebenzi* for *msebenzi* (470). Adding or stripping affixes in order to find the *word form* for the word search is as challenging to the user as stem identification. This unfortunately means that the problem of stem identification is simply replaced by the challenge of identifying word forms. Advocates of the stem tradition could argue that, if identification of the lemma entails the selection of a section of the complex orthographic word anyway, why not then also cut the noun prefixes, which brings one back to stem lemmatisation? It is not possible to lemmatise the entire paradigms of all isiZulu words in printed dictionaries. Prinsloo (2010) tries to make a case for selection on the basis of frequency in this regard by saying that the lexicographer should ensure that the frequently used forms are included. The OZSD indeed lemmatises quite a number of frequently used derivations of *-sebenz-*, i.e., *abasebenzi*, *ekusebenzeni*, *ekusebenziseni*, *ekusetshenzisweni*, *emisebenzini*, *emsebenzini*, *imisebenzi*, *isisebenzi*, *izisebenzi*, *-sebenzela*, *-sebenzisa*, *-sebenzisana*, *-setshenziswa*, *-setshenzwa*, *ukusebenza*, *ukusetshenziswa*, *umsebenzi*. This is useful, but for the learner reading an isiZulu novel, the low frequency words also need to be decoded for him/her to understand the specific utterance.

Returning to the infinitive, dictionaries following word lemmatisation will, by default, also at least have the lemmas *-hamba*, *-sebenza* and *-khuluma*, honouring the non-disputed stem lemmatisation approach for verbs, but will lemmatise infinitives as nouns according to the default word lemmatisation strategy for nouns, i.e., on the first letter of full forms. The OZSD accordingly lemmatises the full forms of a number of frequently used infinitive nouns, such as *ukudla* 'food', *ukuhamba* 'departure' and *ukukhuluma* 'a/the talking' in the alphabetical stretch *uku-* and treats them appropriately for their nominal meanings. These infinitive *nouns*, however, stand in contrast with the infinitive verbs *ukudla* 'to eat', *ukuhamba* 'to walk/go' and *ukukhuluma* 'to speak' in isiZulu. *Ukudla*, *ukuhamba* and *ukukhuluma* have, therefore, also been lemmatised in the OZSD under their stem forms *-dla*, *-hamba* and *-khuluma* as well with applicable treatment for their verbal meanings. However, as argued above, dictionary users become used to looking up infinitive verbs under their stem forms. When looking up *ukudla*, *ukuhamba* and *ukukhuluma*, the user is unlikely to consider the possibility that (s)he should also check under *uku-* for the possible existence of an infinitive noun with the same stem. As in the case of lemmatising stems, a cross-reference in this case is imperative from the articles of the verb stems to the full nouns in such cases. In many instances the nominal and verbal meanings are closely related, e.g., *ukukhuluma*: 'to talk, a/the talking', *ukuhamba* 'to travel, a/the travelling', *ukusebenza* 'to work; a/the working', but in cases, such as the infinitive noun *ukujula* 'depth', the infinitive verb stem, i.e., *-jula*, means 'consider carefully'. A cross-reference from *-jula* to *ukujula* is imperative to

avoid misguiding the user. Inserting such cross-references would of course require additional space in the dictionary. In addition, users should be alerted in the user's guide to the dictionary to check for possible nominal forms when looking up infinitives under their verbal stems and vice versa.

There should be no doubt that word lemmatisation contributes substantially to reducing the problems that stem lemmatisation causes to users. However, the problem of stem identification or word identification is still not solved, and probably never will be solved in printed dictionaries and that moves the focus to electronic dictionaries.

Electronic dictionaries for isiZulu — the final frontiers?

In the early nineties, the electronic era was met with great enthusiasm and expectations expressed in relation to electronic dictionaries and their enormous potential to supersede printed or paper dictionaries in imaginative ways. As the title *Lexicographers' Dreams in the Electronic-Dictionary Age* of De Schryver (2003) suggests, early publications on EDs were dreams about the potential of the new medium and the expected revolution it would bring along, such as antiquating the paper dictionary in a decade or two. These publications list dozens of advantages of EDs, such as accessibility, user-friendliness and especially the availability of space and processing speed. Many of these issues are discussed in detail by Dodd (1989), Bolinger (1990), Atkins (1996), Nesi (1999), Geeraerts (2000), Harley (2000) and Prinsloo (2001), to name but a few. Meijs (1990) even predicted the end of the paper dictionary by 2000. Prinsloo (2005) believes that the potential of electronic dictionaries lies in the utilization of what he calls true electronic features, such as pop-up access, bringing together of related items, new routes to the data, less dependency on alphabetical order, fuzzy spelling, intelligent extrapolation of characters keyed in and audible pronunciation. For the purposes of this article, the question is what the status of currently available isiZulu dictionaries is in terms of lemmatisation and solving the issue of stem or word identification that dominate the discussion in the previous sections of this article.

In principle, catering for all of the approximately 2 500 occurrences of *-sebenz-* in the PZC is not a problem in electronic dictionaries, given the almost unlimited available space and the speed of information retrieval, cf. Prinsloo (2001) and De Schryver (2003). The question, however, is whether this goal has been achieved.

A number of electronic dictionaries and word lists are available for isiZulu, such as the *Webster's Online Dictionary*, *Freelang.net* and the *Dicts.info*. However, the most sophisticated online dictionary is the *isiZulu.net*. The major stongpoints of this dictionary are that it is extensive; there is no need for stem search; and it automatically gives a morphological analysis of the stem plus affixes. The *isiZulu.net* offers some promising features in solving the most problematic cases discussed in terms of Van Wyk (1995)

above where stem identification is problematic. To illustrate: it offers two access routes to *impilo*, i.e., *impilo* and *mpilo*, and both *intombi* and *ntombi* for *intombi*. Plural forms of these nouns can also be directly looked up by typing their full forms, *izimpilo* and *izintombi*, respectively. In addition, this dictionary is useful in cases where the learner finds it difficult to isolate stems/words.

From the examples given in (16), it is clear that the stem identification problem has at last been resolved. The inexperienced learner can simply type in the word or part of it and is (re)routed to the appropriate lemma. Moreover, the quality of the treatment is good.

(16)

The image shows two side-by-side screenshots of the isiZulu.net dictionary interface. Both screenshots have a search bar containing the word 'umsebenzi' and a search button. The left screenshot shows the search results for 'umsebenzi', categorized under 'Direct Matches'. It lists two main entries: 'umsebenzi/abasebenzi' (n, 1/2 (-sebenzi) [umse'be:nzi] [abase'be:nzi]) with the translation 'worker; labourer', and 'umsebenzi/imisebenzi' (n, 3/4 (-sebenzi) [umse'be:nzi] [imise'be:nzi]) with translations 'duty; job; task; work', 'employment', 'business; trade', and 'service'. The right screenshot shows the search results for 'ngomsebenzi', categorized under 'Potential Decompositions (verbatim)'. It lists three entries: 'umsebenzi/abasebenzi' (n, 1/2 (-sebenzi) [umse'be:nzi] [abase'be:nzi]) with the translation 'worker; labourer', 'umsebenzi/imisebenzi' (n, 3/4 (-sebenzi) [umse'be:nzi] [imise'be:nzi]) with translations 'duty; job; task; work', 'employment', 'business; trade', and 'service', and 'ngomsebenzi' [ngomse'be:nzi] ← nga + umsebenzi with the translation 'about/by/at/on/in/per/as per (a/the) worker'. Below this, another entry for 'ngomsebenzi' [ngomse'be:nzi] ← nga + umsebenzi is shown with the translation 'about/by/at/on/in/per/as per (a/the) duty'.

The question, however, is how comprehensive the *isiZulu.net* is in terms of coverage of entire paradigms of words, such as the paradigm for *-sebenz-* as given in the appendix? Formulated differently, can all orthographic forms of nouns and verbs in isiZulu be looked up in the *isiZulu.net*? To answer these questions, the *isiZulu.net* was subjected to a number of random tests in terms of the paradigm for *-sebenz-*, as well as to random selections from a number of published isiZulu dictionaries.

For the first test, the top 10 frequencies were taken, i.e., *umsebenzi* (5 583), *emsebenzini* (1 456), *imisebenzi* (1 009), *ukusebenza* (548), *msebenzi* (470), *lomsebenzi* (446), *nomsebenzi* (309), *usebenza* (299), *esebenza* (290), *ngomsebenzi* (253). All of these were found in the *isiZulu.net*. See, for example, the treatment of *umsebenzi* and *ngomsebenzi* in (16).

For the inexperienced user, the automatic guidance from *msebenzi* to *umsebenzi*, *abasebenzi* and *imisebenzi* is excellent because, as mentioned above, no addition of characters to the word is required to look it up.

(17)



In the case of the successful automatic retrieval of *umsebenzi* from *lomsebenzi* (lo+umsebenzi), *nomsebenzi* (na+umsebenzi) and *ngomsebenzi* (nga+umsebenzi), the results are equally satisfying, because the search was successful and morpho-phonological processes are reversed by the dictionary and presented to the user in a clear and user-friendly way.

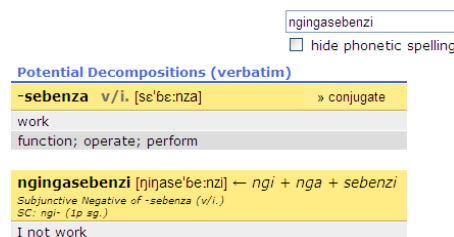
For the second test, ten words were selected from the paradigm for *-sebenz-* that occur five times in the PZC as given in Table 1 and even though these words occur with a low frequency, six were found in the *isiZulu.net*.

Table 1: A random selection of derivations of *-sebenz-* occurring 5 times in the PZC and their presence or absence in *isiZulu.net*

ayisebenzele	YES		singasebenzi	YES
ngingasebenzi	YES		uzosebenzani	NO
ngisebenzile	YES		sisebenzile	YES
nisasebenza	YES		wayesasebenza	NO
njengokwemisebenzi	NO		wayesesebenza	NO

Example (18) reflects the quality of treatment for *ngingasebenzi* in the *isiZulu.net*.

(18)



First, decomposition of *ngingasebenzi* to *-sebenza* gives a useful morphological breakdown into stem with prefixes. The user learns that (s)he is dealing with a derivation of *-sebenza*. Secondly, a translation equivalent of the full word *ngingasebenzi*, 'I not work' (even though the latter is a very direct translation) is given. So, for this example one could say that the problem of stem/word identification has been resolved and the user finds sufficient comment on form as well as comment on semantics of the full word.

For the third test, a study of the first lemma on every 25th page of the WAZ (nine lemmas) was done in terms of its inclusion or omission from the *isiZulu.net* and its presence in the PZC. It was found that five of the nine lemmas occurred in the PZC. Only three were lemmatised and treated in the *isiZulu.net*. A similar selection of the first lemma on every 50th page of the ZED and its inclusion or omission from the *isiZulu.net* and the PZC revealed that from the 19 lemmas in question seven occurred in the PZC and three in the *isiZulu.net*. From these three tests, it is clear that the *isiZulu.net* electronic dictionary performs well on the more frequently used words but substantial enlargement will be required to cover less frequently used words as well.

The least electronic dictionaries could do is to link paradigms, such as those in the appendix to the stem/word, i.e., *-sebenza* in this case.

Conclusion

The weakest option for lemmatising nouns and verbs in isiZulu is to lemmatise verbal stems without suffixes and in the case of nouns, noun stems without their prefixes and without the augmentative and diminutive nominal suffixes. This lemmatisation strategy is not user-friendly; stem identification is a major obstacle; a vast amount of knowledge of morphophonetics is presupposed; and the user is often in doubt whether (s)he has successfully retrieved information. Even if the users do manage to identify the stem and to look it up, all the additional information conveyed by the affixes have to be 'added back on' and the user will not know for sure whether (s)he has come to the right conclusion. Lemmatising verb stems represents a slight improvement. At least the meanings of the suffixes need not be artificially added on as in the case of lemmatising stems without their suffixes.

Lemmatising stems with their prefixes merely added on (left-expanded) is a better option, because the user has the advantage of seeing the full form of infinitive verbs and the full forms of nouns with additional information, such as tonal indication. This strategy is more user-friendly, but stem identification remains problematic and a substantial amount of knowledge of morphophonetics is still presupposed.

Word lemmatisation applicable to nouns is by far the better strategy, because nouns can be looked up under the first letter. For given non-derived nominal forms, the problem of stem identification is solved for all nouns. This strategy is especially beneficial for those nouns where stem identification is

problematic. The strategy is user-friendly and no knowledge of the grammar is presupposed. However, for nominal and verbal derivations, especially those where nominal and verbal stems occur with multiple prefixes, the problem of stem/word identification remains unsolved.

The problem of word/stem identification which is present in all of the lemmatisation strategies employed for isiZulu can only be solved in electronic dictionaries. Most electronic dictionaries are mere translated word lists and are not of much use to the target users especially for their productive needs. A clear exception is the *isiZulu.net* online dictionary, where the problem of stem/word identification has been solved for most of the frequently used words in isiZulu, but more comprehensive electronic isiZulu dictionaries are required to alleviate the need for stem/word identification for less frequently used words as well.

References

A. Dictionaries and corpora

- COBUILD = Sinclair, J. (Ed.). 1995². *Collins COBUILD English Dictionary*. London: HarperCollins.
- CSD = Rycroft, D.K. 1982. *Concise SiSwati Dictionary*. Second impression. Pretoria: J.L. van Schaik.
- CZD = Dent, G.R. and C.L.S. Nyembezi. 1961. *Compact Zulu Dictionary*. Pietermaritzburg: Shuter and Shooter.
- DTS = Elliott, W.A. 1897. *Dictionary of the Tebele & Shuna Languages. With Illustrative Sentences and Some Grammatical Notes*. Frome and London: Butler & Tanner.
- EZD = Doke, C.M., D. McK. Malcolm and J.M.A. Sikakana. 1958. *English-Zulu Dictionary*. Johannesburg: Witwatersrand University Press.
- ISZ = Mbatha, M.O. (Ed.). 2006. *Isichazamazwi sesiZulu*. Pietermaritzburg: New Dawn Publishers.
- OZSD = De Schryver, G.-M. (Ed.). 2010. *Oxford Bilingual School Dictionary: Zulu and English*. First Edition. Cape Town: Oxford University Press Southern Africa.
- SSED = Mabile, A. and H. Dieterlen. 1988. *Southern Sotho-English Dictionary*. Revised by R.A. Paroz. Morija: Morija Sesotho Book Depot.
- WAZ = Dekker, A.M. and J.H. Ries. 1958. *Woordeboek Afrikaans-Zoeloe Zoeloe-Afrikaans/Isichazamazwi Isibhunu-IsiZulu IsiZulu-Isibhunu*. Johannesburg: Afrikaanse Pers Boekhandel.
- ZED = Doke, C.M. and B.W. Vilakazi. 1948. *Zulu-English Dictionary*. First Edition. Johannesburg: Witwatersrand University Press.
- Dicts.info*: <http://www.dicts.info/>
- Freelang dictionaries*: <http://www.freelang.net/>
- isiZulu.net*: <http://isizulu.net>
- PSC: Pretoria Sepedi Corpus. University of Pretoria
- PZC: Pretoria Zulu Corpus. University of Pretoria
- Webster's Online Dictionary*: <http://www.websters-online-dictionary.org/>

B. Other literature

- Atkins, B.T. Sue. 1996. *Bilingual Dictionaries: Past, Present and Future*. Gellerstam, Martin, Jerker Järborg, Sven-Göran Malmgren, Kerstin Norén, Lena Rogström and Catarina Røjder Pap-

- mehl (Eds.). 1996. *Euralex '96 Proceedings I-II, Papers Submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*: 515-546. Gothenburg: Department of Swedish, Göteborg University.
- Bennett, P.R.** 1986. Grammar in the Lexicon, Two Bantu Cases. *Journal of African Languages and Linguistics* 8(1): 1-30.
- Bolinger, D.** 1990. Review of the *Oxford Advanced Learner's Dictionary of Current English*. *International Journal of Lexicography* 3(2): 133-45.
- De Schryver, G.-M.** 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.
- De Schryver, G.-M.** 2010. Revolutionizing Bantu Lexicography — A Zulu Case Study. *Lexikos* 20: 161-201.
- Dodd, W. Steven.** 1989. Lexicocomputing and the Dictionary of the Future. *Lexicographers and their Works*. James, Gregory (Ed.). 1989. *Lexicographers and Their Works*: 83-93. Exeter Linguistic Studies 14. Exeter: University of Exeter Press.
- Doke, C.M.** 1945. *Text-Book of Zulu Grammar*. London: Longmans, Green and Co. Ltd.
- Gauton, Rachelle.** Forthcoming. Lexicography of the Nguni Languages. *Handbooks of Linguistics and Communication Science, Volume Dictionaries: An International Encyclopedia of Lexicography*. HSK Vol. 5.4. Dictionaries: Supplementary Volume. Berlin: Mouton de Gruyter.
- Geeraerts, D.** 2000. Adding Electronic Value: The Electronic Version of the *Grote Van Dale*. Heid, Ulrich et al. (Eds.). *Proceedings of the Ninth EURALEX International Congress. EURALEX 2000. Stuttgart, Germany, August 8th–12th, 2000*: 75-84. Stuttgart: Stuttgart University.
- Gouws, R.H. and D.J. Prinsloo.** 2005a. *Principles and Practice of South African Lexicography*. Stellenbosch: African Sun Media.
- Gouws, R.H. and D.J. Prinsloo.** 2005b. Left-expanded Article Structures in Bantu with Special Reference to IsiZulu and Sepedi. *International Journal of Lexicography* 18(1): 25-46.
- Harley, Andrew.** 2000. Software Demonstration: Cambridge Dictionaries Online. Heid, Ulrich et al. (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress. EURALEX 2000. Stuttgart, Germany, August 8th–12th, 2000*: 85-88. Stuttgart: Stuttgart University.
- Hartmann, R.R.K.** 1989. Sociology of the Dictionary User: Hypotheses and Empirical Studies. Hausmann, F.J. et al. (Eds.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*: 102-111. Berlin: De Gruyter.
- Masubelele, Rose.** 2007. [Review of] Mbatha, M.O. (Ed.). 2006. *Isichazamazwi sesiZulu*. Pietermaritzburg: New Dawn Publishers. *Lexikos* 17. 458-461.
- Meijs, W.J.** 1990. Morphology and Word-Formation in a Machine-Readable Dictionary: Problems and Possibilities. *Folia Linguistica: Acta Societatis Linguisticae Europaeae* 24(1-2): 45-71.
- Mpungose, M.H.** 1998. Analysis of the Word-Initial Segment with Reference to Lemmatizing Zulu Nasal Nouns. *Lexikos* 8(AFRILEX series 8:1998): 65-87.
- Mtuzze, P.T.** 1992. A Critical Survey of Xhosa Lexicography 1772–1989. *Lexikos* 2 (AFRILEX series 2:1992): 165-177.
- Nesi, Hilary.** 1999. A User's Guide to Electronic Dictionaries for Language Learners. *International Journal of Lexicography* 12(1): 55-66.
- Prinsloo, D.J.** 1994. Lemmatization of Verbs in Northern Sotho. *South African Journal of African Languages* 14(2): 93-102.

- Prinsloo, D.J.** 2001. The Compilation of Electronic Dictionaries for the African Languages. *Lexikos* 11: 139-159.
- Prinsloo, D.J.** 2005. Electronic Dictionaries Viewed from South Africa. *Hermes, Journal of Linguistics* 34: 11-35.
- Prinsloo, D.J.** 2009. Current Lexicography Practice in Bantu with Specific Reference to the *Oxford Northern Sotho School Dictionary*. *International Journal of Lexicography* 22(2): 151-178.
- Prinsloo, D.J.** 2010. [Review of] G.-M. de Schryver (Ed.). 2010. *Oxford Bilingual School Dictionary: Zulu and English*. First Edition. Cape Town: Oxford University Press Southern Africa. *Lexikos* 20 (AFRILEX series 20:2010). 760-766.
- Prinsloo, D.J. and Gilles-Maurice de Schryver.** 1999. The Lemmatization of Nouns in African Languages with Special Reference to Sepedi and Cilubà. *South African Journal of African Languages* 19(4): 258-275.
- Prinsloo, D.J. and Rufus H. Gouws.** 1996. Formulating a New Dictionary Convention for the Lemmatization of Verbs in Northern Sotho. *South African Journal of African Languages* 16(3): 100-107.
- Van Wyk, E.B.** 1995. Linguistic Assumptions and Lexicographical Traditions in the African Languages. *Lexikos* 5: 82-96.

Appendix A: Words containing *-sebenz-* in the PZC with their frequencies indicated in brackets

UMSEBENZI (5883), EMSEBENZINI (1456), IMISEBENZI (1009), UKUSEBENZA (548), MSEBENZI (470), LOMSEBENZI (446), NOMSEBENZI (309), USEBENZA (299), ESEBENZA (290), NGOMSEBENZI (253), UKUSEBENZISA (244), IZISEBENZI (242), SEBENZISA (219), NGOKUSEBENZISA (199), NGUMSEBENZI (179), ABASEBENZI (174), NEMISEBENZI (173), USEBENZISE (167), SEBENZA (155), EMISEBENZINI (153), USEBENZISA (143), ABASEBENZA (134), KOMSEBENZI (132), ASEBENZE (131), WASEBENZA (127), OSEBENZA (125), NGISEBENZA (124), NOKUSEBENZA (121), SOMSEBENZI (112), NGEMISEBENZI (112), USEBENZE (101), YOMSEBENZI (100), UYASEBENZA (97), ISEBENZA (96), ASEBENZA (96), BASEBENZA (92), SISEBENZA (90), ISEBENZISA (90), AYESEBENZA (88), KULOMSEBENZI (87), KUNGUMSEBENZI (86), ISISEBENZI (81), SISEBENZISE (74), ASEBENZISE (72), SISEBENZISA (69), LOWOMSEBENZI (68), ZOMSEBENZI (63), KOSEBENZA (63), KUSEBENZA (62), ZEMISEBENZI (58), NASEMSEBENZINI (58), WAYESEBENZA (57), NGOKUSEBENZA (56), BASEBENZE (55), YOKUSEBENZA (53), MISEBENZI (53), ISEBENZISE (53), EZISEBENZA (53), ZISEBENZISA (51), LEMISEBENZI (51), BESEBENZA (51), ZISEBENZA (50), SISEBENZE (50), UKUYOSEBENZA (49), KWEMISEBENZI (49), ESEBENZISA (49), ENGASEBENZI (49), ZOKUSEBENZA (45), YOKUSEBENZISA (45), NGISEBENZE (45), NEZISEBENZI (44), BASEBENZA (44), BASEBENZISE (42), USEBENZILE (41), NGIYASEBENZA (40), ISEBENZE (37), OKUSEBENZA (36), AKASEBENZI (36), YEMISEBENZI (35), UKUSEBENZELA (35), ABASEBENZAYO (35), ABABESEBENZA (35), ESISEBENZA (34), OSEBENZISA (33), NINGASEBENZI (32), EKUSEBENZENI (32), YIMISEBENZI (31), OMSEBENZI (31), NOKUSEBENZISA (31), IYASEBENZA (31), AYOSEBENZA (31), UZOSEBENZA (30), NABASEBENZI (30), LOKUSEBENZISA (30), USEBENZELA (29), LISEBENZA (29), ASEBENZISA (29), USEBENZAPHI (28), LOKUSEBENZA (28), KUBASEBENZI (28), ENGENAMSEBENZI (27), BESEBENZISA (27), WOMSEBENZI (26), KUSEBENZE (26), ZABASEBENZI (25), OWAYESEBENZA (25), KWAKUNGUMSEBENZI (25), KWABASEBENZI (25), EZISEBENZISA (25), BOMSEBENZI (25), LISEBENZE (24), EMISEBENZI (24), BABESEBENZA (24), NGALOMSEBENZI (23), KOKUSEBENZA (23), BAYASEBENZA (23), UYOSEBENZA (22), NALOMSEBENZI (22), EKUSEBENZISENI (22), ANGISEBENZI (22), AKUNAMSEBENZI (22), ZISEBENZISE (21), NGABASEBENZI (21), IMISEBENZANA (21), EZINGENAMSEBENZI (21), ESESEBENZA (21), YASEBENZA (20), UZISEBENZISE (20), ONGASEBENZI (20), NGASEBENZA (20), ANGINAMSEBENZI (20), USEMSEBENZINI (19), EZISEBENZINI (19), ENGISEBENZA (19), ANGASEBENZI (19), ZISEBENZE (18), UKULUSEBENZISA (18), OSEBENZAYO (18), ESASEBENZA (18), ABASEBENZISA (18), WOKUSEBENZA (17), USASEBENZA (17), UNGASEBENZI (17), SOKUSEBENZA (17), SIYASEBENZA (17), AYISEBENZI (17), WUMSEBENZI (16), USEBENZANI (16), UMSEBENZANA (16), OSEBENZELA (16), NGISEBENZELA (16), ZIYASEBENZA (15), UKUZISEBENZISA (15), UKUSISEBENZISA (15), SINGASEBENZISA (15), NISEBENZE (15), NGISEBENZISE (15), ESASISEBENZA (15), BAYOSEBENZELA (15), ZEZESEBENZI (14), YABASEBENZI (14), WEZISEBENZI (14), UNOMSEBENZI (14), UNGASEBENZISA (14), NISEBENZA (14), NGIYOSEBENZA (14), ISEBENZI (14), EZISEBENZELA (14), EYAYISEBENZA (14), ESEBENZE (14), ANGASEBENZISA (14), ABANGASEBENZI (14), YAYISEBENZA (13), WAYOSEBENZA (13), UWASEBENZISE (13), UKUZOSEBENZA (13), UKUYISEBENZISA (13), NJJENGOMSEBENZI (13), NGANGISEBENZA (13), KWAKUSEBENZA (13), KUWUMSEBENZI (13), KUNOMSEBENZI (13), ESEMSEBENZINI (13), ESEBENZELA (13), EMSEBENZI (13), AZOSEBENZA (13), UNGASEBENZISI (12), SIYOSEBENZA (12), SEKUSEBENZA (12), OLUSEBENZAYO (12), NGIZOSEBENZA (12), NGISASEBENZA (12), INSEBENZO (12), INOMSEBENZI (12), EZAZISEBENZA (12), BENGISEBENZA (12), BAYOSEBENZA (12), ZOKUSEBENZISA (11), WUKUSEBENZISA (11), WASEBENZISA (11), WABASEBENZI (11), USEYASEBENZA (11), UBEBESEBENZA (11), SIZOSEBENZA (11), SASEBENZA (11), ABASEBENZELA (11), YIZISEBENZI (10),

WAWUSEBENZA (10), YEZISEBENZI (10), WEMISEBENZI (10), UZISEBENZISA (10), UYISEBENZISE (10), UKUWASEBENZISA (10), SENGISEBENZA (10), SEBENZANI (10), OKUNGUMSEBENZI (10), NGISEBENZISA (10), NASEMISEBENZINI (10), NANGOMSEBENZI (10), KUMSEBENZI (10), INGUMSEBENZI (10), ELISEBENZA (10), BESEBENZELA (10), AYASEBENZA (10), ZISEBENZI (9), WOKUSEBENZISA (9), USEBENZELE (9), UKUNGASEBENZI (9), SEZISEBENZI (9), SEBESEBENZA (9), SEBENZELA (9), NANGEMISEBENZI (9), KWASEBENZA (9), KUSEBENZISA (9), KULOWOMSEBENZI (9), ISISEBENZA (9), INGASEBENZISA (9), ESIZISEBENZISAYO (9), ESISEBENZISA (9), ESISEBENZAYO (9), ENGASEBENZI (9), BENGASEBENZI (9), BASEBENZELE (9), BABESEBENZISA (9), AWUSEBENZI (9), ASISEBENZELI (9), ASEBENZELE (9), ASEBENZELA (9), YOKUSEBENZELA (8), WAYESEBENZISA (8), WAYENGASEBENZI (8), USESEBENZE (8), UKUSEBENZELANA (8), UKULISEBENZISA (8), UBUSEBENZA (8), OZOWASEBENZISA (8), OYISEBENZISAYO (8), NIZOSEBENZA (8), NGOKOMSEBENZI (8), NGISEMSEBENZINI (8), NESISEBENZI (8), NAMEBENZI (8), LUSEBENZA (8), KAYISEBENZI (8), IZOSEBENZA (8), IYOSEBENZA (8), EZISEBENZAYO (8), ESEBENZAYO (8), ELISEBENZISA (8), BEMISEBENZI (8), AYISEBENZISE (8), AWASEBENZISE (8), ALISEBENZISE (8), YISEBENZI (7), YASEMSEBENZINI (7), UYISEBENZISA (7), UYAKUSEBENZA (7), USUSEBENZA (7), USESEBENZA (7), UNGASEBENZA (7), ULISEBENZISE (7), UKUZISEBENZELA (7), UKUSEBENZISANA (7), SISEBENZELA (7), SILISEBENZISE (7), SEBENZE (7), OZOSEBENZA (7), OKOMSEBENZI (7), NGINAMSEBENZI (7), NEZABASEBENZI (7), LEYOMISEBENZI (7), KWEZISEBENZI (7), KOKUSEBENZISA (7), INGASEBENZI (7), ESEBENZISE (7), ENGANGISEBENZA (7), BOKUSEBENZISA (7), BAZOSEBENZA (7), BASEBENZELA (7), AZISEBENZISE (7), ASISEBENZI (7), ANGAZISEBENZISA (7), AKASEBENZI (7), ZASEBENZA (6), YALOMSEBENZI (6), UZISEBENZELA (6), USUSEBENZE (6), UNGAZISEBENZISA (6), SINGASISEBENZISA (6), SINGASEBENZA (6), SIKUSEBENZISE (6), NISEBENZILE (6), NGIZOSEBENZISA (6), NGINEMISEBENZI (6), NGAEMSEBENZINI (6), MAKASEBENZE (6), LIZOSEBENZA (6), LIYASEBENZA (6), LISEBENZISA (6), LASEBENZA (6), ILUSEBENZISE (6), EYOSEBENZA (6), ESIWASEBENZISA (6), ESEBENZE (6), ESEBENZEKAYO (6), ENGASENAMSEBENZI (6), BOKUSEBENZA (6), BESEMSEBENZINI (6), BASEBENZISANE (6), AYISASEBENZI (6), ANGENAMSEBENZI (6), ANGASEBENZA (6), ALUSEBENZISE (6), ZIWUMSEBENZI (5), ZILISEBENZISA (5), ZENSEBENZO (5), WAYESEBENZA (5), WAYESASEBENZA (5), UZOSEBENZANI (5), UWASEBENZISA (5), USISEBENZISE (5), USISEBENZISA (5), SIYISEBENZISA (5), SISISEBENZISE (5), SISEBENZILE (5), SINGASEBENZI (5), SESISEBENZA (5), SEMSEBENZINI (5), SEMISEBENZI (5), SEKUNGUMSEBENZI (5), SASISEBENZA (5), OYISEBENZI (5), OWASEBENZA (5), OKUSEBENZISA (5), NJENGOKWEMISEBENZI (5), NJENGOKUSEBENZISA (5), NJENGEMISEBENZI (5), NISASEBENZA (5), NGOKWEMISEBENZI (5), NGISEBENZILE (5), NGINGASEBENZI (5), NGAKUSEBENZISA (5), NABASEBENZISA (5), NABASEBENZA (5), LISEBENZILE (5), LINGASEBENZA (5), LASEMSEBENZINI (5), LABASEBENZI (5), KUSEBENZISEKE (5), ISEBENZELE (5), EZOSEBENZA (5), EZISEBENZI (5), ESIZISEBENZISA (5), ENGUMSEBENZI (5), ENGASEBENZA (5), BESASEBENZA (5), BASASEBENZA (5), AYISEBENZELE (5), AWUSEBENZISE (5), ASEBENZELANA (5), ASEBENZAYO (5), ABAZOSEBENZA (5)

Management and Internal Standardization of Chemistry Terminology: A Northern Sotho Case Study*

Elsabé Taljard, *Department of African Languages,
University of Pretoria, Pretoria,
Republic of South-Africa (elsabe.taljard@up.ac.za)*

and

Mahlodi Jean Nchabeleng, *Department of African Languages,
University of Pretoria, Pretoria,
Republic of South-Africa*

Abstract: One of the many implications of the process of language democratization which started post-1994 in South Africa is the empowerment of the previously marginalized South African Bantu languages to become languages of higher functions, i.e. languages of learning and teaching, and also of scientific discourse. This in turn implies the development, consolidation and especially standardization of terminology for each of these languages, and the compilation of LSP dictionaries. This article describes the terminological processing of a technical source text prior to translation, which formed part of the compilation of a *Quadrilingual Explanatory Dictionary of Chemistry*. It reports on the model of terminology management that was utilized and explores strategies for the internal standardization of terms in the absence of readily available, standardized chemistry terminology.

Keywords: TERMINOLOGY MANAGEMENT, TERMINOLOGY STANDARDIZATION, NORTHERN SOTHO CHEMISTRY TERMINOLOGY, USERS' PREFERENCES, TERM EXTRACTION, TERM EQUIVALENCE, TECHNICAL TRANSLATION

Opsomming: Bestuur en interne standaardisering van chemieterminologie: 'n Noord-Sotho gevallestudie. Een van die talle implikasies van die proses van taaldemokratiesering wat na 1994 in Suid-Afrika plaasgevind het, is die bemagtiging van die voorheen benadeelde Suid-Afrikaanse Bantoetale om ook tale van hoër funksies te word, dit wil sê tale van onderlig en leer, en ook tale van wetenskaplike diskoers. Dit impliseer die ontwikkeling, konsolidasie en veral standaardisering van terminologie vir elkeen van hierdie tale, asook die saamstel van vakwoordeboeke. Hierdie artikel beskryf die terminologiese prosessering van 'n tegniese teks voor die vertaling daarvan. Die vertaling vorm deel van die samestelling van 'n *Viertalige Verklarende Chemiewoordeboek*. Die artikel lewer verslag oor die model van terminologiebestuur wat gebruik is

* This article is a revised version of a paper presented at the Fifteenth International Conference of the African Association for Lexicography which took place at the University of Botswana, Gaborone, Botswana, 19–21 July 2010.

en verken strategieë vir die interne standaardisering van terme in die afwesigheid van algemeen-beskikbare, gestandaardiseerde chemieterme.

Sleutelwoorde: TERMINOLOGIEBESTUUR, STANDAARDISERING VAN TERMINOLOGIE, NOORD-SOTHO CHEMIETERME, GEBRUIKERSVOORKEURE, TERMONTTREKKING, TERMEKWIVALENSIE, TEGNIESE VERTALING

1. Contextualization

One of the many implications of the process of language democratization which started post-1994 in South Africa is the empowerment of the previously marginalized South African Bantu languages to become languages of higher functions, i.e. languages of learning and teaching, and also of scientific discourse. This in turn implies the development, consolidation and especially standardization of terminology for each of these languages, and the compilation of LSP dictionaries. Rising to this challenge, the *Suid-Afrikaanse Akademie vir Wetenskap en Kuns* (SAAWK) initiated a project that had as one of its aims the compilation of a *Quadrilingual Explanatory Dictionary of Chemistry*. A detailed account of issues pertaining to the planning of this dictionary is given in Taljard and Gauton (2000) and will not be repeated here; suffice it to mention briefly that the intended target users of this dictionary are senior secondary school learners and undergraduate chemistry students. Subject field experts, in cooperation with a team of terminographers, compiled an English lemma list of 500 high frequency chemistry terms to be included in the envisaged dictionary. Terminological definitions were compiled in English for each of these terms, followed by the provision of Afrikaans term equivalents and the translation of the definitions into Afrikaans. The SAAWK then approached the Department of African Languages at the University of Pretoria to assist with the translation of the terms and their definitions into Northern Sotho (as representative of the Sotho languages) and Zulu (as representative of the Nguni languages). This project was subsequently incorporated into the M.A. (course work) programme of the Department of African Languages. Each student participating in the project received 50 terms and their definitions in English (and Afrikaans), which constituted the source text that had to be translated into the two languages mentioned above. Two teams consisting of ten participants for each language were therefore envisaged to participate in the project. Participants had a dual role to fulfil, i.e. that of both translator and terminologist. The nature of the project, particularly the physical circumstances and terminological context within the participants found themselves, posed a number of terminological challenges which had to be addressed in order to produce a terminologically sound final product. The aim of this article is to briefly describe the model of terminology management that was utilized for this project, and secondly, to report on the terminological processing of one of the ten source texts prior to translation, focussing on the internal standardization of terms in the

absence of readily available, standardized chemistry terminology. The second point specifically investigates the involvement of target users in preliminary standardization procedures.

2. Terminology management model

For this particular project, participants were required to translate a technical text which constitutes a random extract from the domain of chemistry. Since the students participating in this project are by no means full-time translators and/or terminologists, the translated text represents an instance of once-off text production, and therefore calls for what Wright and Wright (1997: 147) call *ad hoc* terminology management. The default model for terminology management advocated in most theoretical treatises is that of systematic terminology management, but as Wright and Wright (op. cit.) point out, this model does not make provision for the limitations that exist in the conventional translation workplace. Consequently, translation oriented terminology management is set apart from other terminological activities. Terminologists working within a systematic model have access to subject field experts and usually have the time to collect material, select terminology and organize it according to logical concept systems. In contrast, *ad hoc* terminology management calls on the translator-terminologist to create and manage his/her own terminology resources. Furthermore, whereas systematic terminology management is subject-field driven, *ad hoc* terminology management is text-driven. Translators/terminologists are often confronted with source texts which constitute random extracts from a specific domain, and their lack of expertise in the subject field makes it difficult, if not impossible, to reconstruct the logical concept structure of the domain in question. Wright and Wright (1997: 148) list several disadvantages experienced by translators/terminologists, and these are particularly applicable to the participants in the chemistry dictionary project. In the first instance, they are not subject field experts, and as a result have no knowledge of the concept system within which they are working. Secondly, they have no or limited access to subject field experts and lastly, available research materials on chemistry in especially the two target languages, i.e. Northern Sotho and Zulu are almost non-existent. In order to decide on an appropriate approach to even small-scale terminology management, the conditions that prevail in the individual working environment have to be considered. Every student who participated in the chemistry dictionary project worked alone, without any interaction with other members of the translation team. The majority of participants have only a very basic level of computer literacy with limited or no access to internet facilities. Translation is done without the benefit of sophisticated electronic terminology management systems or translation memory systems. Despite these constraints, it is only reasonable to expect a minimum level of documentation to support the translation work at hand. For this particular project, and taking the non-ideal situation of the translators/terminologists into

consideration, it was decided that a bilingual glossary of source terms with their equivalents in the target languages was the minimum requirement, even though Wright and Wright (1997: 151) indicate that such a list does not really meet the minimum requirement for terminological documentation. They do concede however, that terminologist-translators need to determine for themselves what the basic minimum terminological entry must look like in their specific working environment. During the initial planning of the project it was decided that the project leader would take responsibility for collecting and consolidating glossaries submitted by all participants and making a final bilingual (English–Northern Sotho for the Northern Sotho speaking participants and English–Zulu for the Zulu speaking participants) term list available to all. As pointed out by Tufiş (2004), it is common knowledge that terminological consistency over a large collection of thematic documents is hard to attain, and having such an internally standardized list would ensure terminological consistency not only in the work of each individual participant, but also within the project as a whole. It is furthermore our intention to submit the respective glossaries to the National Language Body for Northern Sotho and the National Language Body for Zulu, sub-structures of PanSALB which are responsible for standardization and consequent dissemination of terminology. In cases where multiple term equivalents exist, all equivalents will be retained, but terms which have shown themselves during the course of the investigation to be preferred will be listed first. It would then be the task of the standardization body to make a choice from amongst competing equivalents. In this way, the project can make a positive contribution to terminology development in Northern Sotho and Zulu respectively. The Northern Sotho glossary appears as Addendum A.

It should be clear from the foregoing discussion that some form of rudimentary terminology management is essential in any project dealing with technical translation, even if it is on a small scale and on an *ad hoc* basis.

The discussion below focuses first on the procedural steps which are followed in order to compile a bilingual (English–Northern-Sotho) glossary of chemistry terms. Secondly, the issue of internal standardization is addressed, with particular reference to the potential role that target users can play in this regard. The data generated by one of the participants (who is also the co-author of this article) form the basis for the discussion. Sections of this article are furthermore based on her unpublished mini-dissertation, of which full particulars appear in the bibliography.

3. Terminological processing of the source text

Translating technical texts into a lesser resourced language such as Northern Sotho requires proper and sometimes innovative terminological processing of the source text prior to the actual translation.

3.1 Terminology extraction

The first step in the translation of the source text is the (semi-automatic and manual) extraction of terminology from the source text. To this end, an electronic special purpose corpus was compiled, consisting of the 50 terms and their definitions. This is necessary in order to semi-automatically extract all terms from the source text, making use of WordSmith Tools' KeyWord function. Simply put, the KeyWord function compares the frequency with which an item occurs in the special purpose corpus with its frequency in a larger, general reference corpus, and isolates KeyWords with a significantly higher or lower frequency of occurrence. (For a detailed description of the procedure for the identification of KeyWords, i.e. potential term candidates, see Taljard et al. 2007: 160). Not all KeyWords thrown up by the KeyWord search are however necessarily terms, and manual perusal of the list is necessary to eliminate non-terms from the term candidate list.

For the purpose of this study, the existing English definitions of the 50 chemistry terms which is the source text that is to be translated into Northern Sotho, automatically constitute the special purpose corpus, whereas the *University of Pretoria English Internet Corpus* (PEIC) is used as the general or reference corpus. The special purpose corpus consists of 1 225 tokens and 158 types, whereas the reference corpus has approximately 12,5 million tokens and 118 193 types. Running the KeyWord search on our special corpus resulted in 81 term candidates being thrown up, of which 72 turned out to be terms. This procedure can be carried out not only for single word terms, but also for multiword terms, and an additional 18 two word terms were extracted semi-automatically. However, semi-automatic term extraction does not succeed in extracting all terms from a source text. According to Taljard and De Schryver (2002), semi-automatic term extraction accounts for approximately 60% of terms in a running text. Therefore, computational extraction needs to be complemented by manual term excerption. Term conscious reading of the source text resulted in a further 40 single and 19 two word terms being identified, thus giving a total of 149 terms isolated from the source text. Compare Table 1 in this regard:

Table 1: Results of term extraction

	Single word terms	Multiword terms
Terms extracted semi-automatically	72	18
Terms excerpted manually	40	19
Subtotal	112	37
TOTAL	149	

3.2 Sourcing of term equivalents

The next step is the sourcing of translation equivalents for all terms occurring in the source text. This would result in the bilingual term list or glossary, to be made available to all participants in the project. Availability of such a list will ensure terminological consistency in the project at large, making sure that all participating translators make use of the same translation equivalents for terms occurring in the source texts. As will be pointed out below, this ideal is not always easily attainable.

The first preference when sourcing term equivalents for the 149 source terms should be perusal of standardized sources. These would normally include dictionaries, preferably LSP dictionaries dealing with the subject field at hand, and official term lists. Due to the lack of LSP dictionaries for Northern Sotho, participants in the project had no option but to consult existing LGP dictionaries, and the only available official terminology list, the Terminology and Orthography of 1988. Consultation of these sources produced equivalents for only 57 of the 149 source terms. This means that only 38.2% of source terms can be covered by terms from standardized Northern Sotho sources. Trawling through the available standardized sources revealed a further problem — for 36 of the 57 source terms, multiple equivalents were found; for some terms as many as four equivalents were found. As can be seen from Table 2 below, the multiplicity of TEs is to be found on various levels: in some cases variation is on the lexical level (cf. TEs for 'decomposition', 'particle' and 'separation'), in other cases on the orthographical level, i.e. different spellings of the same term (cf. *nekethifi* vs. *neketifi*) and in yet other cases, the variation concerns the term formation strategy (indigenous term vs. transliteration, cf. *sedilana* vs *esiti*).

Table 2: Multiple term equivalents for source terms

Source term	TE1	TE2	TE3	TE4
acid	<i>sedilana</i>	<i>esiti</i>		
decomposition	<i>polo</i>	<i>kamologano</i>	<i>tlharamollo</i>	
particle	<i>karolwana</i>	<i>seripana</i>	<i>lerathana</i>	<i>sekgwana</i>
separation	<i>karogano</i>	<i>kgaogano</i>	<i>tlogelano</i>	
negative	<i>ganetšago</i>	<i>latolago</i>	<i>nekethifi</i>	<i>neketifi</i>

It therefore seems that even so-called standardized sources suffer from proliferation of terms, a symptom of inadequate implementation of standardization procedures.

Seeing that only little more than a third of the source terms could be covered by consulting standardized sources, the translator/terminologist was compelled to also consult non-standardized sources. For the purpose of this project, non-standardized sources consisted mainly of informal term lists com-

piled by individuals working in the field of chemistry, who are also speakers of Northern Sotho. With regard to the use of non-standardized sources, it needs to be acknowledged that these sources need to be treated with the necessary circumspection. However, in this particular instance, these sources provided well-formed and appropriate TEs for many of the source terms, revealing an exceptional engagement with both linguistic and conceptual issues. By utilizing these sources, equivalents for a further 22 source terms could be provided, leaving 70 source terms without term equivalents. For these source terms, equivalents had to be coined by making use of the appropriate term formation strategies for Northern Sotho. Figure 1 summarizes the different sources from which term equivalents were harvested.

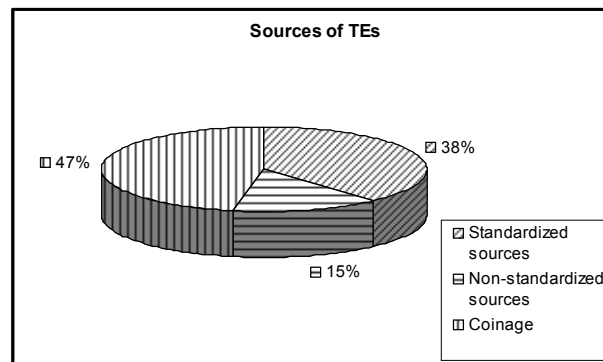


Figure 1: Sources of translation equivalents

Ideally, coining of terms should be done in collaboration with subject field experts who are also speakers of Northern Sotho, but from a practical point of view this is not always possible. Term translation strategies that are available to the terminologist include the following:

- **Semantic transfer**, specifically semantic specialization, a process during which a word from the language for general purpose (LGP) attains the status of a language for special purposes (LSP) term by acquiring an additional, more technical meaning. Compare the following examples in this regard:
 - Source term 1: **acid**
 - Translation equivalent: *sedilana*
 - LGP meaning of TE: something sour, acid
 - LSP meaning of TE: chemical substance that contains hydrogen and, when dissolved in water, has a Ph level of below 7, and a sour taste.

- Source term 2: **particle**
 - Translation equivalent: *lerathana*
 - LGP meaning of TE: crumb
 - LSP meaning of TE: smallest building block of an organism
- **Paraphrasing**, i.e. explanation or description of a concept by making use of a phrase or even a sentence, cf.:
 - Source term 1: **carbonated water**
 - Translation equivalent: *meetse ao a tšhetšwego khapone*, i.e. water to which carbon has been added
 - Source term 2: **empirical formula**
 - Translation equivalent: *fomula ye e netefaditšwego*, i.e. formula which has been confirmed / proven
- **Compounding**, i.e. formation of new terms by joining existing words or lexical items. There seems to be a natural chronological link between paraphrasing and compounding, whereby paraphrases tend to become compounds, probably because of frequency of use. Compare the following examples in this regard:
 - Source term 1: **atomic**
 - Translation equivalent: *seka-athomo* 'atom-like, like an atom'
 - Source term 2: **time unit**
 - Translation equivalent: *motšonako* < *motšo wa nako* 'unit of time'
 - Source term 3: **trivial name**
 - Translation equivalent: *leinatlwaelo* < *leina la tlwaelo* 'name of habit'
- **Borrowing**, which includes the use of loan words, where the term and its meaning is retained intact with no adaptation of the morphological structure of the word. In this particular case, symbols representing chemical elements and units of measurement are, according to international practice, retained as is. Examples are the following:
 - Source term 1: **Debye** (unit of measurement named after the Dutch physicist P.J.W. Debye)
 - Translation equivalent: *Debye*
 - Source term 2: **H₂O** (symbol)
 - Translation equivalent: **H₂O** (seka)

Borrowing also includes the use of adoptives or transliterations, in which case the adopted word is completely adapted — morphologically and phonologically — to the structure of the borrowing language. Compare the following examples:

- Source term 1: **dimer**
 - Translation equivalent: *taemara*
- Source term 2: **magnetite**
 - Translation equivalent: *makenetaete*
- Source term 3: **titration**
 - Translation equivalent: *taetereišene*

After having completed the coining of equivalents for those source terms for which no equivalents could be found, the first deliverable of the initial terminological processing of the source text was available, i.e. a bilingual term list, containing all source terms isolated from the source text, followed by their Northern Sotho equivalents. However, the problem of multiple term equivalents for 43 of the 149 source terms still persisted. Ideally, such a list of source terms and their multiple equivalents should be submitted to an official standardization body for formal standardization, a process during which a preferred term from amongst multiple translation equivalents is identified, again in consultation with subject field experts. In practice this is rarely feasible, due to the time pressure under which translators normally operate. Furthermore, the standardization process of terminology in SA seems rather flawed, one of the main problems being the dissemination and general accessibility of standardized terms. As a result, translators use their own discretion in deciding on appropriate term equivalents for source terms. This practice does however not solve the issue of the multiplicity of term equivalents, and may even contribute to the unnecessary proliferation of terms.

4. Internal standardization of the bilingual term list

For this particular project, it was decided to use the preferences of the target users of the terminology as a guideline for internal standardization. Consequently, a small case study was conducted in three secondary schools in the Limpopo province where chemistry forms part of the curriculum. The aim of this case study was to determine the feasibility of involving target users in the standardization process, even if it is only a preliminary and internal standardization. A questionnaire (Addendum B) consisting of four sections was administered to 30 grade 12 learners and three science educators. It was assumed that, being in the final year of schooling, these learners would already have internalized the basic chemistry concepts and that it would be appropriate to administer the questionnaire to them. All of them have indicated that their mother tongue is Northern Sotho. The first section of the questionnaire concerns the attitude of learners towards the use of Northern Sotho as a language of instruction of especially chemistry. The second section concentrates on establishing the users' preference in the case of multiple TEs, using conceptual appropriateness as guiding principle. In the third section, users' preference

with regard to the use of transliterations versus indigenous terms is investigated, and the fourth section examines users' preferences pertaining to the phonological adaptation and resultant spelling of transliterations. The results collected from the questionnaires were used to internally standardize the bilingual term list to be used in the translation of the chemistry texts, thus ensuring terminological consistency in the translated text. The list would then also serve as the minimum level of terminological documentation as required by Wright and Wright (1997) for this particular project. Insight into target users' preferences can also have a wider impact, in that it may provide some guidelines for future terminology development.

Space constraints do not allow a detailed analysis of the results obtained from all four sections of the questionnaire; the results of section 1 will therefore be dealt with very briefly. The two anchor questions put to learners in this section of the questionnaire were the following:

- (1) Is it easy to learn chemistry in English?
- (2) Do you think that teaching subjects such as chemistry in the mother tongue will have a positive influence on the matric pass rate?

Learners' responses to these questions present an interesting contradiction: 80% of learners indicated that it was easy studying chemistry in English, which is not their mother tongue. On the other hand, they do seem to sense that learning a difficult subject in a language other than their mother tongue may have a negative impact on their successful mastering of the subject: all 30 responded that they believe that teaching subjects such as chemistry in the mother tongue will have a positive influence on the matric pass rate.

In the second section of the questionnaire, learners were presented with 11 source terms for which multiple TEs had been harvested in order to identify the preferred term. These 11 source terms were selected in such a way that the TEs for any particular source term represented the same term translation strategy. Respondents were provided with all TEs for a specific source term, and asked to select the one they prefer. They were also provided with the definition of each term to ensure that they select the term which is conceptually the closest match to the source term.

The terms which were preferred by the majority of respondents were then regarded as being internally standardized for the purpose of the project. As can be seen from Table 3 below, in some cases preferences were very clear — 80% of respondents for example preferred the term *tlemagano* as equivalent for 'bond', 70% preferred *mafolofolo* as equivalent for 'energy'. In other cases, preferences were not so clear-cut. Preferences for the equivalent of the source term 'dispersion' were as follows: *tšitlano* (43%), *phatlalatšo* (30%), and *phatlalalo* (24%). Nevertheless, it was possible to identify a preferred term for all the source terms. Compare Table 3 for an analysis of the results of this section of the questionnaire:

Table 3: Learners' preferences with regard to multiple TEs

	ST and TEs	School 1	School 2	School 3	Total	%
ST BOND						
TE1	<i>tlemagano</i>	7	10	7	24	80
TE2	<i>kamanyo</i>	-	-	1	1	3
TE3	<i>setlamo</i>	-	-	1	1	3
TE4	<i>pofa</i>	-	-	-	-	-
ST DECOMPOSITION						
TE1	<i>pola</i>	1	6	8	15	50
TE2	<i>tlharamollo</i>	7	2	2	11	37
TE3	<i>kamolagano</i>	2	2	-	4	13
ST DEHYDRATION						
TE1	<i>komo</i>	8	5	4	17	57
TE2	<i>tšhwabo</i>	-	1	1	2	6
TE3	<i>meetsefatšollo</i>	2	4	5	11	37
ST DENSITY						
TE1	<i>pitlagano</i>	1	1	2	4	13
TE2	<i>kitlano</i>	1	4	1	6	20
TE3	<i>teteano</i>	1	1	3	5	17
TE4	<i>kitlagano</i>	6	2	2	10	33
TE5	<i>pitlagantšho</i>	1	2	2	5	17
ST DIFFUSION						
TE1	<i>phatlalatšo</i>	9	6	7	22	73
TE2	<i>kgašano</i>	1	4	3	8	27
ST DISPERSION						
TE1	<i>phatlalatšo</i>	1	3	5	9	30
TE2	<i>tšitlano</i>	3	6	4	13	43
TE3	<i>phatlalalo</i>	6	1	1	8	24
ST DISSOCIATION						
TE1	<i>tlogelano</i>	4	2	3	9	30
TE2	<i>kgaogano</i>	3	6	4	13	43
TE3	<i>tšhwalalano</i>	3	2	3	8	27
ST ENERGY						
TE1	<i>mafolofolo</i>	10	17	4	21	70
TE2	<i>mooko</i>	-	-	-	-	-
TE3	<i>maatla</i>	-	3	6	9	30
ST PARTICLE						
TE1	<i>sekgwana</i>	-	-	1	1	3
TE2	<i>lerathana</i>	6	1	6	13	43
TE3	<i>seripana</i>	1	5	2	8	27
TE4	<i>tsekana</i>	-	1	1	2	7
TE5	<i>karolwana</i>	3	3	-	6	20

ST REACTION						
TE1	<i>phetogo</i>	7	6	6	19	63
TE2	<i>kgohlagano</i>	3	4	4	11	37
ST SEPARATION						
TE1	<i>karogano</i>	2	3	1	6	20
TE2	<i>kgaoogano</i>	1	3	4	8	27
TE3	<i>tlogelano</i>	7	4	5	16	53

The purpose of the third section of the questionnaire was to establish whether the target users have a specific preference for, or resistance against the use of transliterations to form term equivalents. The use of transliterations is a much debated issue amongst academics, but as far as we could ascertain, no investigation has thus far been made into the preferences of target users. In this section of the questionnaire, respondents were presented with 12 source terms, each of which has two term equivalents, one being a transliteration, the other a so-called indigenous term. Results for this section indicated that 51% of the preferred equivalents were transliterations, the rest (49%) being indigenous terms. The results obtained from the educators present an interesting contrast to those of the learners: an analysis of the educators' preferences indicated that only 28% of their preferred equivalents are transliterations, the rest being indigenous terms. This could be ascribed to the fact that educators feel that they have to promote the use of 'pure' language, thus discouraging the use of transliterations. Compare Table 4 for an analysis of learners' preferences.

Table 4: Indigenous words versus transliterations

TERM			School 1	School 2	School 3	Total #	Total %
acid	IND	<i>sedilana</i>	3	4	3	10	33
	TRL	<i>esiti</i>	7	6	7	20	66
cell	IND	<i>lelahle</i>	5	4	5	14	47
	TRL	<i>sele</i>	5	6	5	16	53
compound	IND	<i>tlhakanišhetšo</i>	7	5	4	16	53
	TRL	<i>khomphaonte</i>	3	5	6	14	47
copper	IND	<i>mpshiri</i>	8	8	7	23	77
	TRL	<i>koporo</i>	2	2	3	7	23
dimer	IND	<i>phokotšaetša</i>	5	4	5	14	47
	TRL	<i>timara</i>	5	6	5	16	53
element	IND	<i>setho</i>	5	6	4	15	50
	TRL	<i>elemente</i>	5	4	6	15	50
formula	IND	<i>kaelo</i>	3	2	1	6	20
	TRL	<i>fomula</i>	7	8	9	24	80
gas	IND	<i>moya</i>	6	7	2	15	50
	TRL	<i>gase</i>	4	3	8	15	50

layer	IND	<i>tlhatlagano</i>	7	6	8	21	70
	TRL	<i>llaga</i>	3	4	2	9	30
melt	IND	<i>nyaoga</i>	3	4	2	9	30
	TRL	<i>meleta</i>	7	6	8	21	70
mercury	IND	<i>tshipimeetse</i>	5	6	1	12	40
	TRL	<i>mekhuri</i>	5	4	9	18	60
negative	IND	<i>ganetšago</i>	6	5	7	18	60
	TRL	<i>nekethifi</i>	4	5	3	12	40
TOTAL	IND					173	49
	TRL					187	51

Although we acknowledge the fact that the preferences of target users, who can at most be regarded as lay people to perhaps semi-experts, cannot be the final criterion in the selection of a standardized term from multiple equivalents, it surely needs to be taken into consideration that target users seem to have no serious objection to the use of transliterations. Furthermore, since they are probably already familiar with the concepts denoted by these terms, the non-transparency of the transliterated TEs, which is often used as an argument against the use of these forms, is no longer a stumbling block to the conceptual understanding of these terms.

One of the problems with regard to the use of transliterations in Northern Sotho is the phonological adaptation and resultant spelling of these items. The preferred syllable structure in Northern Sotho is a CV-structure, which implies that whenever a word is borrowed from English or Afrikaans, its phonological structure and consequently its spelling needs to be adapted to conform to this requirement. However, this rule is not applied consistently, resulting in multiple equivalents which differ on orthographical level. In the last section of the questionnaire, respondents were requested to choose between one of two variants, the one displaying an adapted syllable structure and spelling corresponding to the syllable structure, the other equivalent being the non-adapted variant. A second aspect that complicates the spelling of transliterations is the indication of aspiration, specifically with regard to the three voiceless plosives [p], [t] and [k]. Speakers often differ with regard to the pronunciation of these sounds, which consequently leads to differences in spelling, cf *molekhule* vs *molekule*. Four items were included in the list of terms where respondents had to choose between a version where aspiration was indicated and one where it was not. The official spelling rules of Northern Sotho provide no guideline with regard to these two issues.

Table 5: Learners' preferences with regard to spelling of transliterations

TERM		School 1	School 2	School 3	Total #	Total %
atom	<i>athomo</i>	9	9	8	26	87
	<i>atomo</i>	1	1	2	4	13

compound	<i>khompaonte</i>	-	1	2	3	10
	<i>khomphaonte</i>	10	9	8	27	90
electrode	<i>eleketerote</i>	10	8	3	21	70
	<i>elektrote</i>	-	2	7	9	30
electron	<i>eleketerone</i>	8	7	6	21	70
	<i>elektrone</i>	2	3	4	9	30
hydrogen	<i>haeterotšene</i>	6	5	6	17	57
	<i>haetrotšene</i>	4	5	4	13	43
isotope	<i>aesothopo</i>	6	7	6	19	63
	<i>isothopo</i>	4	3	4	11	37
molecule	<i>molekhule</i>	9	10	9	28	93
	<i>molekule</i>	1	-	1	2	7
positive	<i>phosethifi</i>	6	7	6	19	63
	<i>phosetifi</i>	4	3	4	11	37
titration	<i>taetereišene</i>	6	6	5	17	57
	<i>taetreišene</i>	4	4	5	13	43
TOTAL	ADAPTED					63
	NON-ADAPTED					27

Analysis of target users' preferences revealed a clear bias towards those forms where the orthographical representation reflects the adapted phonological structure. With regard to the indication of aspiration, in all four examples preference was for those forms where the aspiration was reflected in the spelling, cf. *khomphaonte*, *molekhule*, *phosethifi* and *athomo*. It must however again be emphasized that these preferences should by no means be interpreted as definitive principles — they are merely indications of the preferences of a very restricted sample of Northern Sotho speakers.

5. Conclusion

A statement that is often heard with regard to the African languages of South Africa is that there is a lack of technical terminology in these languages. This statement is however only partially true and represents a very simplified view of a complex matter.

First, the fact that TEs need to be coined for almost 50% (46.9%) of terms isolated from the source text is especially worrisome and indeed confirms the need for a concerted effort of proper terminology development. This need is further substantiated by the fact that less than 40% (38.2%) of source terms can be provided with equivalents from standardized sources. The possibility to recover equivalents for 14.7% of source terms from non-standardized sources adds another dimension to the picture — it implies that the main challenge is not so much a lack of terminology, but rather a lack of standardized termi-




nology. In view of the seeming inability of official standardization bodies to properly manage terminology development, alternative measures of standardization need to be considered. One such an alternative is to use target users' inputs as a guideline for preliminary and project specific standardization. Involving target users in the development of terminology will furthermore encourage them to take ownership thereof. This would also make potential users more inclined to actually use the terminology, since they would feel themselves to have been involved in the creation thereof. However, this can never be more than an interim measure. It can never function as a substitute for a fully functional central standardization body.



References

- Nchabeleng, M.J.** 2011. *Terminological Issues in the Translation of Chemistry Texts from English to Northern Sotho*. Unpublished M.A. Mini-Dissertation. Pretoria: University of Pretoria.
- Taljard, E. and G.-M. de Schryver.** 2002. Semi-Automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho. *Lexikos* 12: 44-74.
- Taljard, E. and R. Gauton.** 2001. Supplying Syntactic Information in a Quadrilingual Explanatory Dictionary of Chemistry (English, Afrikaans, isiZulu, Sepedi): A Preliminary Investigation. *Lexikos* 10: 191-208.
- Taljard, E., R. Gauton and L.A. Gauton.** 2007. Issues in the Planning and Design of a Bilingual (English–Northern Sotho) Explanatory Dictionary for Industrial Electronics. *Lexikos* 17: 152-169.
- Tufis, D.** 2004. Term Translations in Parallel Corpora: Discovery and Consistency Check. *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, Portugal*: 1981-1984.
- Wright, S.E. and G. Budin (Eds.).** 1997. *Handbook of Terminology Management*. Volume 1. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Wright, S.E. and L.D. Wright.** 1997. Terminology Management for Technical Translation. Wright, S.E. and G. Budin (Eds.). 1997: 145-159.

Addendum A: Bilingual term list*Key:*

Where statistical information could be retrieved from the questionnaires with regard to multiple equivalents, the preferred equivalent is listed first and marked with ^P, with other equivalents following in descending order. If no statistical information is available, multiple equivalents are listed in alphabetical order.

-  = equivalent sourced from standardized sources
 = equivalent sourced from non-standardized sources
 = coined equivalent

	Source term	TE1	TE2	TE3	TE4
	accelerate	akgofiša			
	acid	esiti ^P	sedilana		
	actinides	diakethenaete			
	anion	anayone			
	anode	anote			
	aqueous	[sa] meetse			
	atom	athomo ^P	atomo		
	atomic	[sa] go ba le athomo	seka-athomo		
	atomic number	athomopalo			
	bond	tlemagano ^P	kamanyo	setlamo	pofa
	carbon	khapone			
	carbon monoxide	monokosaete wa khapone			
	carbonated water	meetse ao a tšhetšwego khapone			
	cathode	khatote			
	cation	khateayone			
	cell	sele ^P	lelahle		
	charged particles	dikgawana tše maatlafaditšwego			
	chemical	[sa] khemise	khemikhale		
	chemical compound	tihakantšhetšo ya sekhemise			
	chemical reaction	kgohlagano-khemikhale			
	Cl (symbol)	Cl (seka)			
	coke	khokhu			
	colloidal	thulano			
	compound	tihakantšhetšo ^P	khomphaonte		
	concentration	motswakoti			
	conductor	sesepediši			
	control rods	melangwanataolo			
	copper sulphate	salafeiti ya koporo			
	current	moela			
	cyclotron	sesesedi			

	debye	<i>tepaye</i>			
	decompose	<i>bola</i>			
	decomposition	<i>polo^p</i>	<i>tlharamollo</i>	<i>kamologano</i>	
	decomposition reaction	<i>phetogo ya polo</i>			
	dehydration	<i>komo^p</i>	<i>meetsefatšollo</i>	<i>tšhwabo</i>	
	density	<i>kitlagano^p</i>	<i>kitlano</i>	<i>pitlagantšho</i>	<i>tetemo</i>
	device	<i>sešomišwa</i>			
	diatomic	<i>seathomopedi</i>			
	diffusion	<i>phatlalatšo^p</i>	<i>kgašano</i>		
	dilute	<i>hlaphola</i>			
	dimer	<i>taemara</i>			
	dipole	<i>ntlhapedi</i>			
	dipole moment	<i>lebaka la ntlhapedi</i>			
	diprotic	<i>phorothone-pedi</i>	<i>taephorothiki</i>		
	dispersion	<i>tšitlano^p</i>	<i>phatlalatšo</i>	<i>phatlalalo</i>	
	dissociation	<i>kgaogano^p</i>	<i>tlogelano</i>	<i>tswalalano</i>	
	dissolution	<i>motološo</i>			
	dissolve	<i>nyaoša (tr), nyaoga (itr)</i>	<i>tološa (tr), tologa (itr)</i>		
	dissolved substance	<i>selo se se tološitšwego</i>	<i>setološwa</i>		
	distillate	<i>hlwekiša</i>	<i>sekiša</i>		
	donating	<i>[sa] go aba</i>	<i>[sa] go fana</i>		
	E (symbol)	<i>E (seka)</i>			
	electric	<i>[sa] mohlagase</i>			
	electric circuit	<i>lepatlelo la mohlagase</i>			
	electric current	<i>moela wa mohlagase</i>			
	electrochemical	<i>mohlagasekhemikhale</i>			
	electrochemical cell	<i>lelahle la mohlagasekhemikhale</i>			
	electrochemical reaction	<i>phetogo ya mohlagasekhemikhale</i>			
	electrode	<i>eleketerote^p</i>	<i>elektrote</i>		
	electrolysis	<i>eleketerolaesese</i>			
	electrolyte	<i>eleketerolaete</i>			
	electrolytic	<i>eleketerolaetiki</i>			
	electrolytic cell	<i>sele ya eleketerolaete</i>			
	electromotive	<i>mohlagase-sepedi</i>			
	electromotive force	<i>maatla a mohlagase-sepedi</i>			
	electron	<i>eleketerone^p</i>	<i>elektrone</i>		
	electron configuration	<i>peakanyo ya dieleketerone</i>			
	electron density	<i>pitlagano ya dieleketerone</i>			
	electronegative	<i>dieleketerone tše di ganetšago</i>			
	electroplating	<i>pharo ka mohlagase</i>			
	element	<i>elemente^p</i>	<i>setho</i>		

✂	emission	<i>go ntšha</i>	<i>tokollo</i>		
✂	empirical formula	<i>fomula ye e netefadiitšwego</i>			
✂	emulsify	<i>phaphamala</i>			
✂	end point	<i>ntlhaphetšo</i>			
📄	endothermic	<i>entothemiki</i>			
📖	energy	<i>mafolofolo^P</i>	<i>maatla</i>	<i>mooko</i>	
✂	enthalpy	<i>entalipi</i>			
✂	enthalpy change	<i>phetogo ya entalipi</i>			
✂	equivalent point	<i>ntlhatekatekano</i>			
✂	ethoxy ethane	<i>etheini ya ethokisi</i>			
✂	ethylene	<i>ethelini</i>			
📖	formula	<i>fomula^P</i>	<i>kaelo</i>		
📖	gas	<i>gase^P</i>	<i>moya</i>		
✂	gas exchange	<i>neeletšano ya gase</i>			
✂	H (symbol)	<i>H (seka)</i>			
✂	HCl (symbol)	<i>HCl (seka)</i>			
✂	homogeneous	<i>[se] swanago</i>			
✂	homogeneous mixture	<i>motswakotshwano</i>			
📄	hydration	<i>meetsefatšo</i>			
📖	hydrogen	<i>haeterotšene^P</i>	<i>haetrotšene</i>		
📄	hydrogen chloride	<i>tloloraete ya haeterotšene</i>			
📖	immersed	<i>inetše</i>			
📄	ion	<i>ayone</i>			
📖	iron	<i>tshipi</i>			
📄	isotope	<i>aesothopo^P</i>	<i>isothopo</i>		
✂	lanthanide	<i>lanthenaete</i>			
📄	layer	<i>tlhatlagano^P</i>	<i>leyara</i>	<i>llaga</i>	
📖	liquid	<i>seela</i>			
📖	magnetic field	<i>karolokogedi</i>	<i>lepatlelokogedi</i>		
✂	magnetite	<i>makenetaete</i>			
✂	melt	<i>meleta^P</i>	<i>nyaoga</i>		
📖	mercury	<i>mekhuri^P</i>	<i>tshipimeetse</i>		
✂	metallic element	<i>elemente ya metale</i>			
📖	molecule	<i>molekhule^P</i>	<i>molekule</i>		
✂	monoxide	<i>monokosaete</i>			
📖	motion	<i>tšhikinyo</i>	<i>tšhišinyego</i>		
📖	movement	<i>tshepelo</i>	<i>tšhikinyego</i>	<i>tšhišinyego</i>	
📖	negative	<i>[se] ganetšago</i>	<i>[se] latolago</i>	<i>nekethifi</i>	
✂	nuclear reactors	<i>difetoši tša nutliliya</i>			
✂	orbital	<i>[se] dukulogago</i>			
✂	organic compound	<i>tlhakantšhetšo ya tše di phelago</i>			
✂	overall	<i>kakaretšo</i>			
📄	oxidation	<i>okositšenefatšo</i>			
📖	particle	<i>lerathana^P</i>	<i>seripana</i>	<i>karolwana</i>	<i>tsekana</i>
📖	periodic	<i>[sa] lebaka</i>	<i>nako</i>	<i>paka</i>	<i>sebaka</i>
✂	periodic table	<i>papetla ya pheriotiki</i>			

	point	<i>khutlo</i>	<i>ntlha</i>	
	positive	<i>phosithifi^p</i>	<i>Phositifi</i>	
	process	<i>tshepetšo</i>		
	propagated	<i>[sa] go gašanywa</i>		
	property	<i>seeng</i>		
	proton	<i>porotone</i>		
	radio waves	<i>maphotho a radio</i>		
	radioactive	<i>radio-ekethifi</i>		
	reaction	<i>phetogo^p</i>	<i>kgohlagano</i>	
	reduce	<i>fokotša</i>		
	region	<i>selete</i>	<i>tikologo</i>	
	resultant mixture	<i>motswakopheletšo</i>		
	separation	<i>tlogelano^p</i>	<i>kgaogano</i>	<i>karogano</i>
	solute	<i>setologi</i>		
	solution	<i>setološwa</i>		
	solvent	<i>setološa</i>	<i>setološi</i>	
	spin	<i>dikologa (itr), dikološa (tr)</i>		
	stable	<i>[se sa] fetogego</i>		
	strength	<i>maatla</i>		
	substance	<i>nto</i>	<i>selo</i>	
	sulphate	<i>salafeiti</i>		
	symbol	<i>seka</i>	<i>sešupo</i>	
	synonym	<i>lehlalošetšagotee</i>	<i>sinonime</i>	
	time unit	<i>motšonako</i>		
	titration	<i>taetereišene^p</i>	<i>taetreišene</i>	
	trivial name	<i>leinatlwaelo</i>		
	unit	<i>botee</i>	<i>motšo</i>	
	via	<i>ka tsejana</i>		
	volume	<i>bolumo</i>		
	water	<i>meetse</i>		
	x-ray	<i>eksrei</i>		
	zero	<i>lefela</i>		

Addendum B: Questionnaire**QUESTIONNAIRE FOR LEARNERS****SECTION A**

The purpose of this section is to find out how well you are acquainted with chemistry as a learning area. Tick only the applicable answer.

1. Is the subject (learning area) chemistry taught at your school?
Yes No
2. If Yes, from which grade is it taught?
8 9 10 11 12
3. Is it compulsory to learn chemistry in English in your school?
Yes No
4. Is it easy to learn chemistry in English?
Yes No
5. If No, in which language would you prefer to be taught?
Northern Sotho Zulu Tsonga Afrikaans
Swati Ndebele Venda Xhosa
Tswana Southern Sotho
6. Is the language chosen in 5. your mother tongue?
Yes No
7. If Yes, what do you think are the advantages of learning a subject in your mother tongue?

8. Do you think that teaching subjects such as chemistry in the mother tongue will have a positive influence on the matric pass rate?
Yes No Don't know

SECTION B

In this section a definition of a basic chemistry concept will be provided together with possible terms to represent that concept. Please tick the term which you think is the most suitable.

A. Bond: A thing that joins two things together.

Term 1: *tlemagano*

Term 2: *kamanyo*
Term 3: *setlamo*
Term 4: *pofo*

B. Decomposition: A reaction involving the chemical separation of a given compound into two or more simple compounds or substances e.g. $2\text{H}_2\text{O} \rightarrow 2\text{H}_2 + \text{O}_2$

Term 1: *polo*
Term 2: *tlharamollo*
Term 3: *kamologano*

C. Dehydration: The removal of water from a substance e.g. $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (Hydrated copper sulphate) $\rightarrow \text{CuSO}_4 + 5\text{H}_2\text{O}$ (dehydrated copper sulphate + water)

Term 1: *komo*
Term 2: *tšhwabo*
Term 3: *meetsefatšollo*

D. Density: Mass per unit of volume e.g. the density of mercury is $13,5\text{g.cm}^3$

Term 1: *pitlagano*
Term 2: *kitlano*
Term 3: *teteano*
Term 4: *kitlagano*
Term 5: *pitlagantšho*

E. Dispersion: To go in different directions or to scatter.

Term 1: *phatlalatšo*
Term 2: *tšitlano*
Term 3: *phatlalalo*

F. Dissociation: The separation of compounds or atoms, e.g. the dissociation of acetic acid in water to form H^+ ions and acetate ions.

Term 1: *tlogelano*
Term 2: *kgaozano*
Term 3: *tšhwalalano*

G. Energy: The potential to do work or to transfer heat.

Term 1: *mafolofo*
Term 2: *mooko*
Term 3: *maatla*

H. Diffusion: The process during which a substance moves from a higher to a lower concentration.

Term 1: *phatlalatšo*
Term 2: *kgášano*

I. Particle: A very small bit or piece of a substance.

Term 1: *sekgwana*
 Term 2: *karolwana*
 Term 3: *lerathana*
 Term 4: *seripana*
 Term 5: *tsekana*

J. Separation: Stop being combined, to remove elements from each other.

Term 1: *karogano*
 Term 2: *kgagogano*
 Term 3: *tlogelano*

K. Reaction: Chemical change produced by two or more substances acting upon each other.

Term 1: *phetogo*
 Term 2: *kgohlagano*

SECTION C

In this section a term will be given together with two possible translation equivalents (TE). One TE is a coined term, the other is a transliteration. Choose the TE that you prefer. Tick with \checkmark next to the chosen TE.

Term	TE1	TE2
1. acid	<i>sedilana</i>	<i>esiti</i>
2. cell	<i>lelahle</i>	<i>sele</i>
3. compound	<i>tlhakantšhetšo</i>	<i>khomphaonte</i>
4. dimer	<i>phokotššetša</i>	<i>timara</i>
5. element	<i>setho</i>	<i>elemente</i>
6. formula	<i>kaelo</i>	<i>fomula</i>
7. gas	<i>moya</i>	<i>gase</i>
8. negative	<i>ganetšago</i>	<i>nekethifi</i>
9. mercury	<i>tshipimeetse</i>	<i>mekhuri</i>
10. melt	<i>nyaoga</i>	<i>meleta</i>
11. layer	<i>tlhatlagano</i>	<i>llaga</i>
12. copper	<i>mpshiri</i>	<i>koporo</i>

SECTION D

In this section, a term will be given together with two possible term equivalents. The two term equivalents (TEs) are the same, but spelled differently. Choose the TE that you prefer. Tick with \checkmark next to the chosen TE.

Term	TE1	TE2
compound	<i>khompaonte</i>	<i>khomphaonte</i>
electrode	<i>eleketerote</i>	<i>elektrote</i>
electron	<i>eleketerone</i>	<i>elektrone</i>
hydrogen	<i>haeterotšene</i>	<i>haetrotšene</i>
isotope	<i>aesothopo</i>	<i>isothopo</i>
molecule	<i>molekhule</i>	<i>molekule</i>
positive	<i>phosethifi</i>	<i>phosetifi</i>
titration	<i>taetereišene</i>	<i>taetreišene</i>
atom	<i>athomo</i>	<i>atomo</i>

Pedagogical Lexicography: Towards a New and Strict Typology Corresponding to the Present State-of-the-Art

Sven Tarp, *Department of Afrikaans and Dutch, University of Stellenbosch, South Africa, and Centre for Lexicography, Aarhus School of Business and Social Sciences, University of Aarhus, Denmark (st@asb.dk)*

Abstract: A frequent and well-known problem within lexicography is the use of various terms to denominate the same phenomenon as well as the use of the same term to denominate various, completely different phenomena. Such a non-systematic terminology may lead to confusion in the discipline and hamper its theoretical and practical development. The problem is especially severe within so-called pedagogical lexicography. A short panoramic review shows that especially the terms "pedagogical lexicography/dictionaries", "didactic lexicography/dictionaries", "school dictionaries" and "learners' dictionaries" are used with a lot of different meanings that vary from author to author, from country to country, from culture to culture. Although publishing houses could hardly be expected to use a strict terminology for their products, this should nevertheless be expected from theoretical lexicography. In order to overcome the present confusion, it is therefore urgent to establish a typology that can be used as reference by scholars dealing theoretically with the subfield of pedagogical lexicography. The article will first show the amazing variety of meanings addressed to the various terms in the theoretical literature. It will then approach the problem along two different lines: 1) establishing a clear definition of the terms "pedagogical", "didactic", "school" and "learner" in a lexicographical perspective, and 2) referring to the existing practice where the terms are frequently used in a much broader sense than in the theoretical literature. Based upon these considerations, a new and strict typology that corresponds to the present practice of pedagogical lexicography will be presented.

Keywords: PEDAGOGICAL LEXICOGRAPHY, PEDAGOGICAL DICTIONARIES, LEARNERS' LEXICOGRAPHY, LEARNERS' DICTIONARIES, SCHOOL DICTIONARIES, CHILDREN'S DICTIONARIES, DESK DICTIONARIES, COLLEGE DICTIONARIES, DICTIONARIES FOR FOREIGN-LANGUAGE LEARNERS, DICTIONARIES FOR MOTHER-TONGUE LEARNERS, DICTIONARIES FOR LEARNERS OF SCIENTIFIC DISCIPLINES

Opsomming: Pedagogiese leksikografie: Op weg na 'n nuwe en streng tipologie wat ooreenstem met die huidige gebruik daarvan. 'n Algemene en bekende probleem binne die leksikografie is die gebruik van verskeie terme om na 'n enkele verskynsel te verwys, sowel as die gebruik van 'n enkele term om na verskeie, heeltemal verskillende verskynsels te verwys. So 'n niesistematiese terminologie kan tot verwarring binne die vakgebied lei en kan die teoretiese en praktiese ontwikkeling daarvan belemmer. Die probleem is besonder ernstig

binne die sogenaamde pedagogiese leksikografie. 'n Kort, wye oorsig toon dat veral die terme "pedagogiese leksikografie/woordeboeke", "didaktiese leksikografie/woordeboeke", "skoolwoordeboeke" en "aanleerderswoordeboeke" gebruik word in baie verskillende betekenis wat van outeur tot outeur, land tot land, en kultuur tot kultuur verskil. Alhoewel daar allermins van uitgewers verwag kan word om 'n streng terminologie ten opsigte van hulle produkte te gebruik, behoort dit tog van die teoretiese leksikografie verwag te kan word. Om die huidige verwarring te oorkom, is dit dus dringend noodsaaklik om 'n tipologie te vestig wat gebruik kan word deur vakkundiges wat teoreties omgaan met die subveld pedagogiese leksikografie. Hierdie artikel sal eers tens die groot verskeidenheid betekenis aantoon wat aan die verskillende terme toegeskryf word wat in die teoretiese literatuur voorkom. Daarna sal die probleem van twee verskillende kante benader word: 1) die vestiging van 'n duidelike definisie vir die terme "pedagogies", "didakties", "skool-" en "aanleerder-" vanuit 'n leksikografiese perspektief, en 2) die verwysing na die bestaande praktyk waar die terme dikwels in 'n baie breër sin as in die teoretiese literatuur gebruik word. Gebaseer op hierdie oorwegings word 'n nuwe en streng tipologie wat ooreenstem met die huidige gebruik in die pedagogiese leksikografie voorgestel.

Sleutelwoorde: PEDAGOGIESE LEKSIKOGRAFIE, PEDAGOGIESE WOORDEBOEKE, AANLEERDERSLEKSIKOGRAFIE, AANLEERDERSWOORDEBOEKE, SKOOLWOORDEBOEKE, KINDERWOORDEBOEKE, HANDWOORDEBOEKE, WOORDEBOEKE VIR VREEMDETAAL-AANLEERDERS, WOORDEBOEKE VIR MOEDERTAALLEERDERS, WOORDEBOEKE VIR LEERDERS VAN WETENSKAPLIKE DISSIPLINES

1. Introduction

A frequent and well-known problem within lexicography is the use of various terms to denominate one and the same phenomenon as well as the use of one and the same term to denominate various, completely different phenomena. Such a non-systematic terminology may lead to confusion in the discipline and hamper its theoretical and practical development. The problem is especially severe within so-called pedagogical lexicography as shown by Welker (2008) who provides the — so far — most comprehensive panorama of this important subfield of lexicography with references to almost 800 titles. A small anecdote will illustrate the problem. In a book which I published in 2008 and where I, among other things, criticized this inconsequent practice, I defined a *learner's dictionary* as follows (Tarp 2008: 125):

A learner's dictionary is a dictionary whose genuine purpose is to satisfy the lexicographically relevant information needs that learners may have in a range of situations in connection with the foreign-language learning process.

Of course, one could discuss some minor things in this definition, for instance whether it should be called foreign language, second language, non-native language, additional language, etc. But apart from that, it cannot be denied that the above definition corresponds to what is generally understood by *learner's dictionary* among English-speaking lexicographers. It was therefore rather sur-

prising to read an article by Giacomini and Rovere (2009) who first quoted a contribution by Cowie (1996) entitled *The 'Dizionario Scolastico': a Learner's Dictionary for Native Speakers* and then commented (p. 21-22):

The question is whether the chosen title is an example of the inconsequence in the use of the concept 'learner's dictionary' criticized by Tarp (2008: 125), or whether his definition is too narrow.

My first reaction was that Giacomini and Rovere's criticism was completely unjustified, first of all because I had defined the term in such a way that nobody could misunderstand it even if they disagreed, and secondly because the definition corresponded to general practice with Cowie as a possible exception. However, after giving it a lot of thought, I reached the conclusion that Giacomini and Rovere may have pointed at a problem that is worth considering. In lexicography, as in any other academic discipline, it is not only necessary to work with well-defined terms; it is also necessary that the definitions reflect the practical and theoretical state-of-the-art of a discipline in constant development. There should also be a logical linguistic relation between the term used and its conceptual content.

In this respect, the lack of a consistent and systematic terminology within pedagogical lexicography is especially deplorable. A short panoramic review shows that especially the terms "pedagogical lexicography/dictionaries", "didactic lexicography/dictionaries", "school dictionaries" and "learners' dictionaries" are used with a lot of different meanings that vary from author to author, from country to country, from culture to culture. Wiegand, for instance, has edited two books on pedagogical lexicography "involving German" (Wiegand 1998 and 2002), which only deal with dictionaries for foreign-language learners, whereas other authors also include first-language learners in the concept. Welker himself insists that the terms "school dictionary" and "learners' dictionaries" should be reserved for dictionaries designed for first-language and second-language learners, respectively, whereas other authors such as Giacomini and Rovere (2009) include dictionaries for first-language learners in the concept of a learners' dictionary; in some countries like South Africa, school dictionaries — frequently bilingual — are produced both for first- and second-language learners; etc.

Although publishing houses may hardly be expected to use a strict terminology for their products, this should nevertheless be expected from theoretical lexicography. In order to overcome the present confusion, it is therefore urgent to establish a typology that can be used as reference by scholars dealing theoretically with the subfield of pedagogical lexicography.

The following reflections aim at establishing such a typology. They will be based upon a study of existing practice as well as the corresponding theoretical literature, especially Herbert Andreas Welker's book *Panorama Geral da Lexicografia Pedagógica* (General Survey of Pedagogical Lexicography) which contains a collection of — frequently contradictory and even opposed — opinions and

ideas expressed by a large number of scholars that permits the reader to get a quick overview of the confusion in terms of terminology within pedagogical lexicography (cf. Tarp 2010). This terminological confusion concerns basic concepts such as:

- (1) Pedagogical lexicography/dictionaries
- (2) Didactic lexicography/dictionaries
- (3) Learner's lexicography/dictionaries
- (4) School lexicography/dictionaries
- (5) Children's lexicography/dictionaries
- (6) College lexicography/dictionaries

In the following, these concepts will be discussed and redefined in the light of existing practice with a view to establishing a new and strict typology that corresponds to the present state-of-the-art of pedagogical lexicography. All quotations from non-English texts have been translated into English by the author.

2. Pedagogical dictionaries versus didactic dictionaries

For many researchers these two terms are synonymous, but some, e.g. Wiegand (1998, 2002), seem to prefer the term *pedagogical dictionaries* while others, e.g. Hernández (1998), prefer *didactic dictionaries*. However, although in the current daily language the terms *pedagogical* and *didactic* are frequently mixed up and used as synonyms, in the scientific literature it is nevertheless, as Welker (2008) rightly points out, necessary to use the two terms in a strict and correct way. With reference to existing pedagogical literature, Welker himself (p. 21) defines pedagogy as the "theory and science of teaching and education" whereas didactics is the "totality of methods and techniques used in education". In this sense a pedagogical dictionary is a *type of dictionary* (used in teaching and education), whereas the adjective *didactic* refers to the quality of the dictionary in terms of learner-friendliness. Welker (2008: 22) concludes:

... *didactics* should only be used in order to qualify the *way* in which the information is provided: more or less didactic, more or less clear, more or less adapted to the users' skills. In this respect, even pedagogical dictionaries may vary in their didactic quality.

So, what is *pedagogical lexicography*? What is a *pedagogical dictionary*? First of all, it is important to state that the two terms are hypernyms. But to what? Welker (2008: 18) provides the following definition:

Pedagogical lexicography includes all dictionaries conceived for learners of either a foreign language or the mother tongue.

This definition seems to be accepted by many scholars and also corresponds to the one given in the *Dictionary of Lexicography* (Hartmann and James 1998: 107):

pedagogical dictionary

A reference work specifically designed for the practical didactic needs of teachers and learners of a language. The distinction usually made between a dictionary for native speakers (school dictionary) and one for non-native learners (learner's dictionary) is not helpful.

as well as to the one furnished by the *Nordisk Leksikografisk Ordbok* (Bergenholtz et al. 1995: 214):

pedagogical dictionary

A dictionary especially designed for certain user groups as a tool to assist language learning.

However, when one looks at existing lexicographical practice it seems that all three definitions — i.e. the definitions provided by Welker, Hartmann and James, and Bergenholtz et al. — are somewhat problematic. Among the thousands of relevant dictionaries, in South Africa alone one may find, among others, the following "atypical" pedagogical dictionaries:

- Illustrated Dictionary of Natural Sciences and Technology Today (Maskew Miller Longman 2006)
- Multilingual Science Dictionary for South African Schools (Longman 2007)
- Sasol Science & Technology Resource (New African Education Publishers 1998)

None of the above dictionaries seems to have been conceived mainly as a language-learning tool. For instance, in the presentation provided on the back cover of the first of the three dictionaries, one may read that,

It enhances conceptual understanding of key concepts and will help the learners succeed in Natural Sciences and Technology.

The main objective of this and other similar *pedagogical dictionaries* is, according to the authors or the publishing house, to assist the learning of natural sciences, technology or whichever scientific or practical discipline, i.e. not specifically the learning of neither a native (first) or a foreign (second and third) language. In this sense, i.e. in correspondence with existing lexicographical practice, the terms *pedagogical dictionary* and *pedagogical lexicography* should be redefined as follows:

Definition:

A *pedagogical dictionary* is a dictionary especially designed to assist learners of languages (whether a native or a foreign language) and of scientific and practical disciplines.

Definition:

Pedagogical lexicography is the branch of lexicography dealing with the theory and practice of pedagogical dictionaries.

3. School dictionaries versus learner's dictionaries

For many researchers (e.g. Lagane 1990), school dictionaries are dictionaries for learners of the mother tongue (first-language learners), whereas learner's dictionaries are produced for learners of a foreign (second, third, etc.) language. For other researchers, the two terms refer to dictionaries designed for both types of learners. As already indicated above, in their *Dictionary of Lexicography* Hartmann and James (1998: 107) are critical of the traditional distinction between school dictionaries and learner's dictionaries:

The distinction usually made between a dictionary for native speakers (school dictionary) and one for non-native learners (learner's dictionary) is not helpful.

Welker (2008: 19) disagrees specifically with this statement by Hartmann and James and writes:

On the contrary, I think that this distinction is very useful and necessary.

There are several reasons why it is both useful and necessary to distinguish between dictionaries for first-language learners and dictionaries for learners of a second, third, fourth, etc. language. The main reason is that these two types of learners do not have the same lexicographical needs — especially at a beginner's level — and these needs have to do with their proficiency level in their first (native) and second (foreign) language, respectively:

- First-language learners, without mastering their own language completely (they are learners), are supposed to be fluent in this language in terms of basic vocabulary and grammar when they start using dictionaries, whereas second-language learners are just beginning to learn the second language when they first consult dictionaries in this language.
- First-language learners are generally still in the process of alphabetization when they start using dictionaries, whereas second-language learners, as a rule, are already alphabetized (although there are important exceptions, especially in multi-language countries like South Africa).
- First-language learners are developing their conceptual thinking and knowledge of the world together with their language skills, whereas second-language learners have already developed their conceptual thinking and knowledge beyond their second-language skills.

So, how should a learner's dictionary be defined? Should it only be understood as a dictionary for learners of a second language? Should it include dictionaries

for first-language learners? And should it also include learners of a specific LSP?

The broadest definitions found include all three types of language learners. Welker (2008), for instance, defines a learner's dictionary as a dictionary for foreign-language learners (as did Tarp 2008). Cowie (1991) discusses a "Learner's Dictionary for Native Speakers". In the same vein, Giacomini and Rovere (2009) define a learner's dictionary as a dictionary for learners of either a native or a foreign language, i.e. what Welker (2008) defines as a "pedagogical dictionary". Fuertes-Olivera and Arribas-Baño (2008) and Fuertes-Olivera (2010) include dictionaries for LSP learners, e.g. specialized translation students, in the concept of a learner's dictionary. But are these definitions broad enough?

The answer once more has to come from practical lexicography. Although its title does not indicate it, the famous South African *Sasol Science & Technology Resource* (Hartmann-Petersen et al. 2001) is, in fact, a specialized dictionary as can be seen from the following article which is formulated and structured according to traditional lexicographical patterns:

layering

A method of artificial propagation in which a shoot from a plant stem is pegged down and covered with soil. Adventitious roots grow from nodes on the shoot, and when they are established the young plant can be detached from the parent plant. *See also* marcotting.

The subtitle, printed on the cover of this dictionary, is *A book resource for learners, students and trainees in science and technology*. In this respect, it cannot be considered anything else than a *dictionary for learners of science and technology*. This concept of a learner corresponds to the one generally used within education where the term *learner* is not only used to refer to language learners, but also to learners of any scientific or practical discipline. In this sense, and being consequent with the general use of the term, a learner's dictionary may also be conceived for learners of, say, accounting, law, or even mechanics.

However, it is important to distinguish between learners in the broad sense of the word (life-long learning) and those who are following a study or teaching program. If the term *learner's dictionary* were to include dictionaries also for life-long learners then it would inevitably include almost all dictionaries as it could be claimed that anybody who consults a dictionary does so in order to "learn" something. The term *learner's dictionary* should therefore be reserved for dictionaries conceived for learners in the narrow sense of the word, i.e. those following a study, teaching or education program. In this respect, a *learner's dictionary* is synonymous to a *pedagogical dictionary*, with the only difference that it is looked at from the point of view of the learner and not from the point of view of the "educator", and it could be redefined as follows:

Definition:

A learner's dictionary is a dictionary especially designed to assist learners of languages (whether a native or a foreign language) and of scientific and practical

disciplines. A learner's dictionary is synonymous to a pedagogical dictionary but focuses on the learner and not the educator.

This definition of a learner's dictionary, however, leaves the discussion of school dictionaries and their definition unsolved. As mentioned above, there are various approaches and definitions among scholars and this terminological confusion is also reflected in the three dictionaries of lexicography consulted for this article. For example, in his *Diccionario de lexicografía práctica*, Martínez de Sousa (1995: 143) provides the following double definition under the lemma *diccionario escolar* (school dictionary):

school dictionary. 1. A language dictionary especially conceived to be used by school children in first-grade education. 2. A language dictionary defining terms used in certain text books in school.

The second part of this definition is interesting because Martínez de Sousa seems to include what is normally called glossaries in his definition. Although almost never discussed in the theoretical literature, these glossaries are highly useful and oft-used lexicographical works and it is, hence, important to consider them as part of the concept of school dictionaries as it has also been emphasized by Gouws and Tarp (2010). The second part of the above definition, however, raises the question why it is only dictionaries conceived for first-grade education that should be considered school dictionaries. Surprisingly, this definition almost totally contradicts the one found in *Nordisk Leksikografisk Ordbok* (Bergenholtz et al. 1995: 214) where there is a reference from the lemma *skoleordbok* (school dictionary), which contains no definition, to the lemma *pedagogisk ordbok* (pedagogical dictionary) under which the authors, among other things, write that school dictionaries are conceived for:

pupils in the secondary school and part of the upper secondary school.

Finally, the third of the three consulted dictionaries, the *Dictionary of Lexicography* (Hartmann and James 1998: 122), provides a definition that does not distinguish between various grades of school children:

school dictionary

A dictionary written for school-children, common features of which are a controlled defining vocabulary, a clear design and the incorporation of illustrations. The boundaries between the school dictionary on the one hand, and children's dictionaries, college dictionaries and desk dictionaries on the other, are not clearly demarcated.

Although this definition to a certain degree seems to mix the purpose of a school dictionary and its features, the most important statement seems to be that a school dictionary is "a dictionary written for school-children". In this respect, it is necessary to distinguish between two fundamental types of school dictionaries:

- (1) A school dictionary that is an abbreviated or reduced version of a general-language dictionary.
- (2) A school dictionary that is designed from the outset as a lexicographical tool specifically conceived for school children.

A *genuine school dictionary* is a dictionary of the second type, i.e. a dictionary designed to be used by school children and adapted to their mental, linguistic, cultural, and encyclopedic development. In fact, the Spanish lexicographer Hernández (1989: 52) seems to exclude most of the first type of dictionaries, which he calls "deplorable", from the world of school dictionaries:

When a dictionary is the result of a purely quantitative reduction, it can in no way be considered a work conceived for the user.

As to the distinction between dictionaries for school children of various grades and the different (and opposing) definitions provided by Martínez de Sousa (1995) and Bergenholtz et al. 1995), respectively, this apparent contradiction may be explained by different national traditions. In a global perspective it therefore seems that the definition furnished by Hartmann and James (1998), i.e. a dictionary written for school-children without taking into account their respective grades, is not only the one that corresponds to lexicographical practice in general, but also the only one with a logic linguistic relation between the term and its conceptual content.

However, it is one thing to determine the target user group of school dictionaries in general, it is another to determine their respective content. As mentioned above, in the theoretical literature several different definitions are to be found in terms of this content. The broadest of these definitions — i.e. the one including all types of dictionaries designed to be used in school whatever the specific content and discipline — corresponds to the lexicographic practice in South Africa which can be seen from the following dictionaries published within the last few years:

- Bilingual school dictionary. Afrikaans–Engels/English–Afrikaans (Pharos 2007)
- Learner's Dictionary for Schools. Afrikaans–Engels/English–Afrikaans (Pharos 2010)
- Multilingual Science Dictionary for South African Schools (Longman 2007)

It has already been stated that no scientific theory can be built upon the commercial titles given by the publishing houses in order to stimulate sales. However, it goes without saying that the three dictionaries listed above — and many more could be added to the list — are conceived and published in order to be used by school children in their various learning activities, i.e. the learn-

ing of languages (whether a native or a foreign language) and scientific disciplines. The concept of a school dictionary should include this common lexicographical practice. In this respect, the term *school dictionary* may be considered a hyponym to the terms *learner's dictionary* or *pedagogical dictionary* and may be redefined in the following way:

Definition:

A school dictionary is a learner's dictionary (or pedagogical dictionary) especially designed to assist school children in learning languages (whether a native or a foreign language) and scientific and practical disciplines.

4. Children's dictionaries versus school dictionaries

Just as it was the case with school dictionaries there are also various definitions of children's dictionaries to be found in the theoretical literature. In his *Diccionario de lexicografía práctica*, Martínez de Sousa (1995: 158) provides the following definition under the lemma *diccionario infantil* (children's dictionary):

children's dictionary

A dictionary especially conceived to initiate the children in the use of this type of works.

This short definition is somewhat surprising because it only focuses on the didactic aspect in terms of initiation in dictionary use and not on the usefulness of the dictionary itself when consulted by its target user group of children who most probably do not consult it in order to be acquainted with dictionary use but to learn and understand new words, etc. There is no further help in the *Nordisk Leksikografisk Ordbok* (Bergenholtz et al. 1995: 73-74) which in the relatively long article under the lemma *barneordbok* (children's dictionary) only lists some features of the type of dictionary without defining it. The *Dictionary of Lexicography* (Hartmann and James 1998: 20), on the other hand, provides the following article:

children's dictionary

A dictionary aimed at children. While the transition between the dictionary for younger children and the school dictionary is fluid, the former is less bound by the conventions of the traditional, fully-fledged general dictionary than the latter. It is based on a limited basic vocabulary and uses pictorial illustrations and 'stories' — often humorous — rather than formal definitions, to explain the meaning of the (predominantly concrete) words.

The small definition *A dictionary aimed at children* is almost self-explanatory and even seems troublesome to the authors as it can be seen in the next sentence where they have problems establishing a dividing line between this type of dictionary and the school dictionary. In fact, almost all scholars discussing children's dictionaries and school dictionaries agree that it is difficult to establish such a dividing line between these two types of dictionaries. For instance,

Hausmann (1990: 1365), which probably represents the so far most interesting contribution on children's dictionaries, notes that there is no clear dividing line between the two dictionary types and that their titles frequently do not reveal their real content, an opinion also shared by (Bergenholtz et al. 1995: 214) who write:

It can be difficult to establish a boundary between most elementary school dictionaries (beginners' dictionaries) and children's dictionaries.

There seem to be three problems in this apparently difficult discussion. First, it is evident from looking at the *features* — and not at the *purpose* or *functions* — of the two dictionary types that there is a fluid transition from one to the other as stated above by Hartmann and James. Nothing else could be expected. Second, a scientific typology should never be built upon the *titles* provided by the publishing houses but should be on the *purpose* and *functions* of the respective dictionaries. And finally, there seems to be a problem with the *logical linguistic relation* between the terms used and their conceptual content. In most countries, children start school between the age of 5 and 7, are considered children at least up to the age of 12 or 14, and continue school up to the age of 15 or 16. This means that they, for a long period, are school children. Consequently, if a school dictionary is defined as a dictionary conceived to be used by the pupils in school, most school dictionaries are at the same time "children's dictionaries".

The real distinction to be made is between preschool dictionaries and school dictionaries with both being graded and subdivided into various categories according to the pupils' mental and linguistic development and their growing knowledge about the world. And the same holds true when a distinction has to be made between school dictionaries, on the one hand, and college, desk, and university dictionaries or whatever they are called in the different traditions, on the other hand. The content and features of a dictionary are, of course, extremely important and essential for lexicographical works but if they are used as a basis for a typology this will only lead to terminological confusion because of the evident fluid transition from one type to another. Upon this basis, the term *pre-school dictionary* as a hyponym to the terms *learner's dictionary* and *pedagogical dictionary*, may be defined in the following way:

Definition:

A pre-school dictionary is a learner's dictionary (or pedagogical dictionary) especially designed to assist pre-school children in learning their native (first) language (or, hypothetically, a foreign language and a scientific or practical discipline).

5. Towards a new and strict typology

Based upon the above reflections, it is possible to outline a new and strict typology corresponding to the present state-of-the-art of pedagogical dictionaries. This typology, of course, could be approached in different ways. If one focuses on *the age of the target user group* (its place in the education system), the following typology could be suggested:

Typology A for pedagogical dictionaries (learner's dictionaries)

- (a) Preschool dictionary
- (b) School dictionary
- (c) College dictionary

If the focus instead is on *the type of learning* that the dictionaries are supposed to assist then the following typology could be suggested:

Typology B for pedagogical dictionaries (learner's dictionaries)

- (a) Dictionary for native-language learners
- (b) Dictionary for non-native-langue learners
- (c) Dictionary for learners of scientific disciplines, etc.

If the two above criteria (age and type of learning) are combined, then the following two typologies could be suggested (the parenthesis indicates that the corresponding type of dictionary will probably not be produced for children at pre-school level):

Typology A for pedagogical dictionaries (learner's dictionaries)

- (1) Dictionary for native-language learners
 - (a) Preschool dictionary
 - (b) School dictionary
 - (c) College dictionary
- (2) Dictionary for non-native-langue learners
 - (a) (Preschool dictionary)
 - (b) School dictionary
 - (c) College dictionary
- (3) Dictionary for learners of scientific disciplines, etc.
 - (a) (Preschool dictionary)
 - (b) School dictionary
 - (c) College dictionary

Typology B for pedagogical dictionaries (learner's dictionaries)

- (1) Preschool dictionary
 - (a) Dictionary for native-language learners
 - (b) (Dictionary for non-native-langue learners)
 - (c) (Dictionary for learners of scientific disciplines, etc.)

- (2) School dictionary
 - (a) Dictionary for native-language learners
 - (b) Dictionary for non-native-language learners
 - (c) Dictionary for learners of scientific disciplines, etc.

- (3) College dictionary
 - (a) Dictionary for native-language learners
 - (b) Dictionary for non-native-language learners
 - (c) Dictionary for learners of scientific disciplines, etc.

6. Four basic considerations

The conclusions in this article may seem strange after so many years of terminological uncertainty and confusion. But they are the only logical ones to be drawn if the definitions of concepts should reflect the-state-of-the-art of lexicography and if a logical linguistic relation between the terms used and their content should be established.

The study of the theoretical literature shows an amazing variety of meanings addressed to the various terms used. Furthermore, the study of existing practice indicates that pedagogical, learners' and school dictionaries are produced, not only to assist the learning of a language (whether mother tongue, second language or a specific LSP), but also to assist the learning of science and various disciplines such as technology, mathematics, etc. It is time for this practice to be reflected in the theoretical literature and in a new and strict typology. For the future, it is especially important to take into account the following four basic considerations:

- (1) Commercial publishing houses cannot be expected to use scientifically correct terms for their products, especially when some terms have already taken root.
- (2) Theoretical lexicography should always work with well-defined terms and concepts.
- (3) The definitions should reflect the state-of-the-art of both practical and theoretical lexicography.
- (4) There should be a logical linguistic relation between the term used and its conceptual content.

Literature

Basel, Barbara and Louise Banks (Eds.). 2006. *Illustrated Dictionary of Natural Sciences and Technology Today*. Cape Town: Maskew Miller Longman.

- Bergenholtz, Henning, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson and Bo Svensén.** 1997. *Nordisk leksikografisk ordbok*. Oslo: Universitetsforlaget.
- Bosman, D.B., I.W. van der Merwe and A.S.V. Barnes.** 2007. *Bilingual School Dictionary. Afrikaans–Engels/English–Afrikaans*. Cape Town: Pharos.
- Cowie, Anthony P.** 1996. The 'Dizionario Scolastico': A Learner's Dictionary for Native Speakers. *International Journal of Lexicography* 9(2): 118-131.
- Du Plessis, Madaleine.** 2010. *Learner's Dictionary for Schools. Afrikaans–Engels/English–Afrikaans*. Cape Town: Pharos.
- Fuertes-Olivera, Pedro A. and Ascención Arribas-Bano.** 2008. *Pedagogical Specialised Lexicography*. Amsterdam/Philadelphia: John Benjamins.
- Fuertes-Olivera, Pedro A. (Ed.).** 2010. *Specialised Dictionaries for Learners*. Berlin/New York: Walter de Gruyter.
- Giacomini, Laura and Giovanni Rovere.** 2009. Italienische Lernerwörterbücher aus soziolinguistischer Sicht. *Lexicographica. International Annual for Lexicography* 25: 21-45.
- Hartmann, R.R.K. and Gregory James.** 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Hartmann-Petersen, Preben, Gus Gerrans and Rasmus Hartmann-Petersen.** 2001. *Sasol Science & Technology Resource. A Book Resource for Learners, Students and Trainees in Science and Technology*. Observatory/Claremont: Infosource.
- Hausmann, Franz Josef.** 1990. Das Kinderwörterbuch. Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta (Eds.). *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography II*: 1365-1368. Berlin/New York: Walter de Gruyter.
- Hernández, Humberto.** 1989. *Los diccionarios de orientación escolar. Contribución al estudio de la lexicografía monolingüe española*. Lexicographica Series Maior 28. Tübingen: Max Niemeyer Verlag.
- Hernández, Humberto.** 1998. La lexicografía didáctica del español: aspectos históricos y críticos. Fuentes Morán, María Teresa and Reinhold Werner (Eds.). 1998. *Lexicografías iberorománicas: problemas, propuestos y proyectos*: 49-79. Madrid/Frankfurt am Main: Iberoamericana and Verduert.
- Lagane, René.** 1990. Les dictionnaires scolaires: enseignement de la langue maternelle. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta (Eds.). *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography II*: 1368-1378. Berlin/New York: Walter de Gruyter.
- Martínez de Sousa, José.** 1995. *Diccionario de lexicografía práctica*. Barcelona: Bibliograf.
- Sutton, Kathleen (Ed.).** 2007. *Multilingual Science Dictionary for South African Schools*. Cape Town: Longman.
- Tarp, Sven.** 2008. *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Max Niemeyer.
- Tarp, Sven.** 2010. H.A. Welker and Pedagogical Lexicography. *Lexikos* 20: 730-749.
- Tarp, Sven and Rufus Gouws.** 2010. Skoolwoordeboeke vir huistaalleerders van Afrikaans. *Lexikos* 20: 466-494.
- Welker, Herbert Andreas.** 2008. *Panorama Geral da Lexicografia Pedagógica*. Brasília: Thesaurus Editora.

- Wiegand, Herbert Ernst (Ed.).** 1998. *Perspektiven der pädagogischen Lexikographie des Deutschen. Untersuchungen anhand von Langenscheidts Großwörterbuch Deutsch als Fremdsprache.* Tübingen: Max Niemeyer.
- Wiegand, Herbert Ernst (Ed.).** 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des de Gruyter Wörterbuch Deutsch als Fremdsprache.* Tübingen: Max Niemeyer.

Theoriebedingte Wörterbuchformprobleme und wörterbuchformbedingte Benutzerprobleme I. Ein Beitrag zur Wörterbuchkritik und zur Erweiterung der Theorie der Wörterbuchform*

Herbert Ernst Wiegand, *Stellenbosch Institute for Advanced Study (STIAS), Wallenberg Research Centre at Stellenbosch University, Stellenbosch, Südafrika (hwiegand@sun.ac.za und herbert.ernst.wiegand@gs.uni-heidelberg.de)*

und

Rufus H. Gouws, *Departement Afrikaans en Nederlands, Universiteit Stellenbosch, Stellenbosch, Südafrika (rhg@sun.ac.za)*

Zusammenfassung: Der empirische Gegenstandsbereich der Wörterbuchforschung verändert sich u.a. dadurch, dass neue Wörterbücher erscheinen. Dabei ist ein Trend zu erkennen, dass zunehmend mehr neue Elemente der Wörterbuchform im Vorspann, im Wörterverzeichnis und im Nachspann auftreten. Die Wörterbuchform wird dadurch komplexer. Für die Theorie der Wörterbuchform ergeben sich dadurch theoriebedingte Wörterbuchformprobleme: Will die Theorie die neuen Entwicklungen berücksichtigen, muss sie erweitert werden und wird dadurch zunehmend komplexer. Wird sie erweitert, erkennt man im Licht der Theorie die Stärken und Schwächen der neuen Formelemente; das führt zur Wörterbuchkritik. Besonders die Schwächen führen zu wörterbuchformbedingten Benutzerproblemen. In diesem ersten Teil des Beitrags werden theoriebedingte Artikelformprobleme behandelt, die sich durch die Berücksichtigung eines bisher nicht untersuchten Artikeltyps ergeben, nämlich den Typ des komplexen Wörterbuchartikels, in dem zwei oder mehr formgleiche Wörter lexikographisch bearbeitet werden, die zu zwei oder mehr Wortklassen gehören. Die theoriebedingten Artikelformprobleme bestehen darin, dass es keine Strukturkonzeption für Artikel dieses Typs, keine systematische Terminologie, keine Typologie und z.T. keine Darstellungsmittel und –methoden gibt. Die Probleme werden dadurch gelöst, dass die Theorie der Wörterbuchform um die fehlenden Elemente systematisch erweitert und die Heuristik ergänzt wird.

* Dieser Beitrag wurde im November/Dezember 2010 während des Aufenthaltes von H.E. Wiegand am Stellenbosch Institute for Advanced Study (STIAS) verfasst. Für die Einladung bedankt sich Fellow H.E. Wiegand hiermit herzlich. — Dieser Beitrag besteht aus zwei Teilen. Der Teil II erscheint in Lexikos 22. 2012 und behandelt Probleme der Datenakzessivität sowie Probleme der Datendistribution bei Wörterbüchern mit distribuiertem Datenspeicher.

Stichwörter: ARTIKELFENSTER, ARTIKELFOLGE MIT FENSTER, ARTIKELSTRUKTURSHEMA, BASALER KOMPLEXER WÖRTERBUCHARTIKEL, EINFACH KOMPLEXER WÖRTERBUCHARTIKEL, FENSTERARTIKEL, KOMPLEXER WÖRTERBUCHARTIKEL, KOMPLEXES ARTIKELTEXTSTRUKTURBILD, SUBSTANTIV-VERB-ADJEKTIV-ARTIKEL, TEILARTIKEL, TEILARTIKELEXTERNER FORMKOMMENTAR, TEILARTIKELEXTERN ERWEITERTER KOMPLEXER WÖRTERBUCHARTIKEL, TEILARTIKELINTERN ERWEITERTER KOMPLEXER WÖRTERBUCHARTIKEL, THEORIEBEDINGTES WÖRTERBUCHFORMPROBLEM, WÖRTERBUCHFORMBEDINGTES BENUTZERPROBLEM, WORTFAMILIENFENSTER, ZWEIFACH KOMPLEXER WÖRTERBUCHARTIKEL

Abstract: Theory-determined Dictionary Structure Problems and Dictionary Structure Determined User Problems I. A Contribution to Dictionary Criticism and the Theory of Dictionary Structures. The empirical domain of the subject matter of dictionary research changes, among others, due to the publication of new dictionaries. A noticeable trend seems to be the increasing occurrence of more new elements regarding dictionary structures in the front matter, central list and back matter. This makes dictionary structure more complex. For the theory of dictionary structures this leads to theory-determined dictionary structure problems: If the theory wants to take cognizance of the new developments it has to be expanded and becomes increasingly complex. If it is expanded one can recognise, in the light of the theory, the strengths and weaknesses of the new structural elements. This leads to dictionary criticism. The weaknesses, especially, lead to dictionary structure-determined user problems. In this first part of the contribution theory-determined article structure problems are treated that result from an investigation of an article type not previously examined, i.e. the complex dictionary articles in which two or more structurally identical words belonging to two or more part of speech classes are lexicographically treated. The theory-determined article structure problems prevail because there are no structural concept for articles of this type, no systematic terminology, no typology and partially no means and methods of presentation. The problems are thereby solved that the theory of dictionary structures is systematically expanded and the heuristics extended to provide for the missing elements.

Keywords: ARTICLE SEQUENCE WITH WINDOW, ARTICLE STRUCTURE SCHEME, ARTICLE WINDOW, BASIC COMPLEX DICTIONARY ARTICLE, COMPLEX ARTICLE STRUCTURAL IMAGE, COMPLEX DICTIONARY ARTICLE, DICTIONARY STRUCTURE DETERMINED USER PROBLEM, NOUN-VERB-ADJECTIVE ARTICLE, PARTIAL ARTICLE, PARTIAL ARTICLE EXTERNAL COMMENT ON FORM, PARTIAL ARTICLE EXTERNAL EXTENDED COMPLEX DICTIONARY ARTICLE, PARTIAL ARTICLE INTERNAL EXTENDED COMPLEX DICTIONARY ARTICLE, SINGLE COMPLEX DICTIONARY ARTICLE, THEORY-DETERMINED DICTIONARY STRUCTURE PROBLEM, TWOFOLD COMPLEX DICTIONARY ARTICLE, WINDOW ARTICLE, WORDFAMILY WINDOW

1. **Zur Einführung: Was sind theoriebedingte Wörterbuchformprobleme und wörterbuchformbedingte Benutzerprobleme?**

Für jedes gedruckte Sprachwörterbuch gilt: Es hat

— einen Wörterbuchgegenstandsbereich

- einen Wörterbuchgegenstand und
- eine Wörterbuchform.¹

Der *Wörterbuchgegenstandsbereich* ist der sprachliche Bereich, aus dem diejenigen sprachlichen Ausdrücke stammen, die hinsichtlich einer oder mehrerer Eigenschaftsausprägungen in einem Sprachwörterbuch lexikographisch bearbeitet sind (vgl. Wiegand 1998: 303). Welche Eigenschaften jeweils Berücksichtigung finden, bestimmt in erster Linie der jeweilige Wörterbuchtyp. Mit Bezug auf die Berücksichtigung der Anzahl der sprachlichen Eigenschaften unterscheidet man den Typ des *monoinformativen* vom Typ des *polyinformativen* Wörterbuchs. Zum erstgenannten Wörterbuchtyp gehören z.B. die allermeisten alphabetischen Aussprachewörterbücher (wie z.B. Duden-6 1990 und WDA 1967). In einem Aussprachewörterbuch werden nur die Eigenschaftsausprägungen *einer* sprachlichen Eigenschaft bei allen mit den Lemmata erwähnten Lemmazeichen lexikographisch bearbeitet, nämlich die lemmazeichenspezifischen Ausprägungen der Eigenschaft 'Aussprache'; andere Bearbeitungseinheiten, die sporadisch auftreten können, dienen ausschließlich dazu, die lexikographische Bearbeitung jeweiliger Lemmazeichen eindeutig zu machen. Dass die Eigenschaft 'Aussprache' komplex ist, so dass z.B. der Worthaupt- und der Wortnebenakzent, weiterhin Akzentvarianten, die Vokalqualität und -quantität u.a.m. lexikographisch bearbeitet werden müssen, damit die Aussprache eines Lemmazeichens bearbeitet ist, spielt bei der Unterscheidung von monoinformativen und polyinformativen Wörterbüchern keine Rolle.

Vom Wörterbuchgegenstandsbereich muss der *Wörterbuchgegenstand* deutlich unterschieden werden (vgl. Wiegand 1998: 302). Nimmt man alphabetische Aussprachewörterbücher als Beispiel, dann gilt: Der *Wörterbuchgegenstand eines Aussprachewörterbuchs* ist die Menge aller Eigenschaftsausprägungen der Eigenschaft 'Aussprache' bei allen Lemmazeichen des Aussprachewörterbuchs. Der *Wörterbuchgegenstand eines polyinformativen einsprachigen Wörterbuchs*, wie z.B. der des CD 2009, des HAT 2009, des BW 1980–84 oder der des Duden-³GW, des MEDAL² 2007, des MWALED 2008 und des CALD³ 2008 ist die Menge der in diesen Wörterbüchern lexikographisch bearbeiteten Eigenschaftsausprägungen von mindestens zwei (meistens aber mehreren) sprachlichen Eigenschaften einer bestimmten Menge der in diesen Wörterbüchern erwähnten sprachlichen Ausdrücke, und die erwähnten Lemmazeichen bilden nur eine Teilmenge dieser Menge von Ausdrücken, da in polyinformativen Wörterbüchern nicht nur lemmatische, sondern auch nichtlemmatische linksgerichtete Angabenadressierung auftritt, so dass in vielen Wörterbuchartikeln ein Themenwechsel erfolgt.

Unter der Wörterbuchform eines gedruckten Sprachwörterbuchs versteht man — grob gesprochen — die Art und Weise der textuellen Präsentation der lexikographischen Text- und Bilddaten und ihrer Beziehungen untereinander im zweidimensionalen Druckraum der Wörterbuchseite sowie im *n*-fach gebrochenen Druckraum des gesamten Wörterbuchs, wobei die Zahl *n* so groß ist

wie die Seitenzahl minus 1 (z.B. liegen zwischen 3 Wörterbuchseiten zwei Umbrüche).

Wir verfügen jetzt — wenn auch nur in recht allgemeiner Weise — über die Begrifflichkeiten, um auf einer oberbegrifflichen Ebene angeben zu können, was unter einem *wörterbuchformbedingten Benutzerproblem* zu verstehen ist: Ein wörterbuchformbedingtes Benutzerproblem liegt vor genau dann, wenn ein kundiger Benutzer-in-actu (i.S.v. Wiegand 1998: 505ff), also einer der die Benutzungshinweise sowie gegebenenfalls andere wörterbuchinterne Metatexte zur lexikographischen Bearbeitung kennt (wie z.B. die Texte in der „Reference Section“ des OALD⁸ 2010), durch jeweils genau bestimmbare Teile der Wörterbuchform eines bestimmten Wörterbuchs entweder nicht ausreichend unterstützt oder sogar behindert wird, wenn er die jeweils erforderliche Benutzungshandlung eines bestimmten Typs (i.S.v. Wiegand 1998: 268ff) ausführen möchte, um sein Benutzerziel zu erreichen; das Benutzerziel besteht bei der Ausführung von Konsultationshandlungen — oberbegrifflich ausgedrückt — stets darin, auf eine konsultative Suchfrage anhand mindestens eines Elements des Wörterbuchgegenstands etwas über den Wörterbuchgegenstandsbereich zu erfahren, das für den Benutzer-in-actu als eine Antwort auf seine Suchfrage gelten kann. Wir werden im Teil II sehen, dass sich verschiedene Typen von wörterbuchformbedingten Benutzerproblemen unterscheiden lassen.

Monoinformative Wörterbücher weisen in der Regel eine Wörterbuchform auf, die so gestaltet ist, dass wörterbuchformbedingte Benutzerprobleme nicht auftreten. Dies sei an einem Beispiel verdeutlicht.

wa₁- wa₁₂: **Aristoteli**ker aristo'te:likə
aristotelisch, A... aristo'te:ɰf
Aristotelismus aristote'lismus
Aristow russ. 'aristef
Aristoxenos aris'toksenos
Arita jap. 'a.rita
Arithmetik arit'me:tik
Arithmetiker arit'me:tikə
arithmetisch arit'me:tɰf
Arithmograph aritmo'gri:f
Arithmologie aritmo'lo:gi:
Arithmomanie aritmoma'ni:,
 -n ...:ən

Abb. 1: Nichtfunktionale Artikelteilstrecke artstr₁ aus Duden-6 1990 mit zwölf Wörterbuchartikeln wa₁ – wa₁₂

Die Artikelteilstrecke in Abb. 1 besteht aus zwölf Kurzartikeln vom Typ des rudimentären Formartikels (vgl. Wiegand 2003: 264 u. Wiegand/Feinauer 2010), die keinen Zweitkommentar aufweisen; davon sind die meisten Einzeilenartikel, und es gibt in allen Artikeln nach der Lemmazeichengestaltangabe und damit auch nach dem Lemma nur Angaben vom Typ der lemma-nahen Angaben. Alle Ausspracheangaben sind an die Lemmazeichengestaltangabe linksadressiert; alle verdichteten Sprachenidentifizierungsangaben (wie z.B. *russ.* und *jap.*) sind an die Aussprachenangaben adjazent rechtsadressiert, und alle unten erweiternden durch Unterstriche oder Unterpunkte realisierten

funktionale Angabezusätze vom Typ der Wortakzentkennzeichnung, die zugleich eine Vokalquantitätskennzeichnung (entweder zu Kürze oder zur Länge des Vokals) sind, sind internlemmatisch hinaufadressiert. Die (i.S.v. Wiegand 2008: 287f) verweisundurchlässige Hauptzugriffsstruktur des Duden-6 1990 ist glattalphabetisch und mit der Makrostruktur strukturell identisch (vgl. Wiegand 2008 u. Gouws/Wiegand 2010a). Das Wörterverzeichnis ist artikelhomogen (d.h.: alle Artikel gehören zum gleichen Artikeltyp). Ein kundiger Benutzer, der den Duden-6 1990 benutzt, muss lediglich eine externe Zugriffshandlung erfolgreich ausführen und kann dann — wenn er die phonetische Umschrift beherrscht — die gesuchte Information zur Aussprache anhand des gesamten Kurzauftrags sozusagen „mit einem Blick“ erschließen (vgl. Wiegand 1998: 331ff; dort wird dies handlungstheoretisch genauer ausgeführt). Wörterbuchformbedingte Benutzerprobleme sind bei der Benutzung monoinformativer Wörterbücher nicht zu erwarten. Ein Benutzer, der in einem monoinformativen Wörterbuch die gesuchte Information nicht erschließen kann, hat entweder einen Handlungsausführungsfehler (i.S.v. Wiegand 1998: 519) gemacht, oder er hat einen sprachlichen Ausdruck gesucht, der kein Element der Makrostruktur ist, so dass die makrostrukturelle Decke des benutzten Wörterbuchs für seine Suchfrage nicht ausreichend ist, weil eine Lemmalücke vorliegt. Weil bei der Benutzung von monoinformativen Wörterbüchern wörterbuchformbedingte Benutzerprobleme in der Regel nicht auftreten, werden Wörterbücher dieses Typs im Folgenden nicht mehr berücksichtigt.

Theoriebedingte Wörterbuchformprobleme sind lexikographietheoretische Probleme, die bei der schrittweisen Erarbeitung einer Theorie der Wörterbuchform und ihrer Erweiterung auftreten (vgl. Anm. 2). Es sei daran erinnert, dass nicht jedes metalexikographische Problem ein lexikographietheoretisches Problem ist, da es u.a. auch lexikographiehistorische Probleme gibt, die ebenfalls metalexikographische Probleme sind. Es gibt zwei Gründe für das Auftreten lexikographietheoretischer Probleme vom Typ theoriebedingter Wörterbuchformprobleme. Der *erste* Grund besteht darin, dass der empirische Gegenstandsbereich der Wörterbuchforschung sich dadurch verändert, dass neue Wörterbücher erscheinen und neue Auflagen von Wörterbüchern herausgegeben werden. Der *zweite* Grund besteht darin, dass der empirische Gegenstandsbereich, nämlich die (wohl niemand bekannte) Anzahl aller Wörterbücher — qualitativ betrachtet — außerordentlich reichhaltig und — quantitativ gesehen — sehr umfangreich ist, und zwar selbst dann, wenn er von vornherein auf Wörterbücher beschränkt wird, in denen Sprachen und Sprachvarietäten lexikographisch bearbeitet sind, deren zugehörige Texte von links nach rechts und auf Zeilen von oben nach unten geschrieben sind. Der große Umfang des empirischen Gegenstandsbereichs und seine Reichhaltigkeit zwingen zum Beginn der theoretischen Arbeit dazu, dass mit einem Ausschnitt aus dem empirischen Gegenstandsbereich gearbeitet werden muss. Liegt der erste Entwurf einer Theorie der Wörterbuchform vor (vgl. Hausmann/Wiegand 1989; Wiegand 1989, 1989a, 1989b), kann der empirische Ausschnitt

erweitert werden (vgl. Anm. 2). Die Wörterbücher dieses erweiterten Ausschnittes werden allerdings dann bereits zwangsläufig im Lichte des ersten Entwurfes der Theorie hinsichtlich ihrer Wörterbuchform betrachtet. Da die Betrachtung eines empirischen Gegenstandsbereichs im Lichte einer Theorie die Wahrnehmung eines Forschers erheblich differenziert und vor allem auch verfeinert, denn zu jeder empirischen Theorie und auch zu jedem Entwicklungsstadium einer Theorie gehört bekanntlich ein bestimmter Auflösungsgrad für die theorielevanten empirischen Phänomene, werden zunehmend weitere Aspekte, Teile und Eigenschaften der Wörterbuchform wahrgenommen. Sollen diese in der Theorie Berücksichtigung finden, können theoriebedingte Wörterbuchformprobleme auftreten. Diese können, wie wir noch genauer sehen werden, von unterschiedlicher Art sein. Solange keine empirischen Phänomene auftreten, die als solche der Wörterbuchform zu betrachten sind und die die Grundlagen der Theorie der Wörterbuchform infrage stellen, sind theoriebedingte Wörterbuchformprobleme durch auf Widerspruchsfreiheit überprüfte Erweiterungen der Theorie, angemessene Spezifizierungen der zugehörigen Heuristik sowie partielle Modifizierungen und systematische Ergänzungen bei der Terminologie lösbar. Dies wird im Folgenden gezeigt.²

Schließlich muss einleitend noch kurz auf die Frage eingegangen werden, welche Zusammenhänge es rechtfertigen, dass zwei so verschiedene Problemsorten wie wörterbuchformbedingte Benutzerprobleme und theoriebedingte Wörterbuchformprobleme zusammen in einer Abhandlung betrachtet werden. Hierzu sei das Folgende festgestellt: Besonders im letzten Jahrzehnt ist bei Wörterbüchern, die in Verlagen erarbeitet werden und die zu bestimmten Wörterbuchtypen gehören, insbesondere bei den Lernerwörterbüchern, bei den Rechtschreibwörterbüchern und auch z. Tl. bei den allgemeinen einsprachigen Wörterbüchern ein zunehmend sich verstärkender Trend zu beobachten, der wohl mindestens zum Teil auch als Ausdruck eines Konkurrenzkampfes der Verlage um Marktanteile zu werten ist und der die Wörterbuchform betrifft. Offensichtlich ist man in den Verlagen der Auffassung, dass sich Wörterbücher um so besser verkaufen lassen, je bunter, abwechslungsreicher und boxenüberladener sie in den Wörterverzeichnissen gestaltet sind und je mehr vor- und nachspanninterne Umtexte sowie illustrierte Einschübe im Wörterverzeichnis auftreten. Die alphabetischen gedruckten Lernerwörterbücher werden mit Formelementen angereichert und zum Teil sogar überhäuft, die früher nur in distinktiven und kumulativen Synonymiken, in Stilwörterbüchern, Wörterbüchern der Sprachschwierigkeiten, Wortfamilienwörterbüchern und Bildwörterbüchern üblich waren; gleichzeitig erweitert sich auch das Datenklassensortiment, das zu den Wörterbüchern gehört. Was die Wörterbuchgegenstände betrifft, hat dieser Trend durchaus positive Aspekte, weil dadurch sprachliche und kulturelle Zusammenhänge sichtbar gemacht werden können, die in einem alphabetischen Wörterbuch, das keinen distribuierten Datenspeicher, sondern nur ein Wörterverzeichnis mit Wörterbuchartikeln aufweist, kaum Berücksichtigung finden können. Mit dem genannten Trend bei den alphabe-

tischen Wörterbüchern, insbesondere bei denen der englischen Verlagslexikographie, den wir als *Trend zum lexikographischen Facelifting* bezeichnen wollen, geht eine theoriefeindliche Einstellung Hand in Hand, wie man z.B. an folgendem Zitat (vgl. dazu auch die Kritik von Tarp 2010: 451ff) sehen kann: „This is not a book about 'theoretical lexicography' — for the very good reason that we do not believe that such a thing exists“ (Atkins/Rundell 2008: 4). Welche negativen Folgen diese bedenkliche Theoriefeindlichkeit hat, werden wir besonders in Teil II dieses Beitrages in Lexikos 22. 2012 zeigen.

Aufgrund des Trends zum lexikographischen Facelifting ergibt sich, was die Wörterbuchform betrifft, das folgende Dilemma: Die Verlagslexikographen erarbeiten Wörterbücher bestimmter Typen so, dass fast in jeder neuen Auflage neuartige ad hoc erfundene Wörterbuchformelemente auftreten: Die Wörterbuchform ihrer Wörterbücher wird dadurch, ohne dass dafür lexikographische Gründe genannt werden, zunehmend komplexer und z. T. auch benutzerunfreundlicher. Weil aber viele — wie z.B. „The Oxford Guide to Practical Lexicography“ (Atkins/Rundell 2008) durchgehend zeigt — von der Theorie der Wörterbuchform reichlich wenig verstehen, treten *erstens* zunehmend mehr wörterbuchformbedingte Benutzerprobleme auf, so dass diese Entwicklung in der Verlagslexikographie zu Lasten der Benutzer geht. *Zweitens* treten dadurch jedoch auch zunehmend neue theoriebedingte Wörterbuchformprobleme dann auf, wenn Metalexikographen sich bemühen, die neuen Entwicklungen im empirischen Gegenstandsbereich zu berücksichtigen, was nun dazu führt, dass die Theorie der Wörterbuchform, da sie als empirische Theorie konzipiert ist, zunehmend weitere Formelemente berücksichtigen muss und dadurch komplexer wird. Da im Lichte der Theorie der Wörterbuchform die wörterbuchformbedingten Benutzerprobleme, z.B. die bei der Benutzung der neuesten englischen Lernerwörterbüchern entstehenden, die von den Lexikographen zu verantworten sind, ebenso wie die für den Benutzer hilfreichen Formveränderungen stets leicht zu erkennen sind (vgl. z.B. Wiegand 2009b), führt dies zwangsläufig entweder zur negativen oder zur positiven Wörterbuchkritik.

Damit dürfte der Zusammenhang der beiden Problemsorten „theoriebedingte Wörterbuchformprobleme“ und „wörterbuchformbedingte Benutzerprobleme“ sowie die Formulierung des Untertitels verständlich geworden sein.

2. Artikelformprobleme

Besonders bei der Analyse von *kondensierten* Wörterbuchartikeln (i.S.v. Wiegand 2003: 207; vgl. auch Bustos Plaza/Wiegand 2005/2006 u. Wiegand 1998b) hinsichtlich ihrer Form können theoriebedingte Artikelformprobleme auftreten und bei der Benutzung von kondensierten Wörterbüchern artikelformbedingte Benutzerprobleme. Im Folgenden werden zunächst Beispiele für den zuerst genannten Problemtyp behandelt.

2.1 Artikelformprobleme für die Theorie der Wörterbuchform I: komplexe Wörterbuchartikel

Bei den lexikographischen akzessiven Einträgen sind die Wörterbuchartikel die wichtigsten (vgl. WLWF-1 2010). Daher wurde in Wiegand 2003 und in Wiegand/Feinauer 2010 ein typologisches System für Wörterbuchartikel erarbeitet, das u.a. aus den folgenden elf wörterbuchformbezogenen Typologien besteht, die sich durch ihr oberstes Typologiekriterium (= TK) unterscheiden.

- Artikeltypologie mit Bezug auf die innere Textkondensierung:
Oberstes TK: *Vorhandensein von standardisierter Textkondensierung*
- Artikeltypologie mit Bezug auf Makrostrukturen und Registerzugriffsstrukturen
Oberstes TK: *Anzahl der Zugriffspfade*
- Artikeltypologie mit Bezug auf initialalphabetische Makrostrukturen
Oberstes TK: *Zugehörigkeit zu einer funktionalen Artikelteilstrecke*
- Artikeltypologie mit Bezug auf die Datendistribution
Oberstes TK: *Bezug der artikelinternen Daten zum artikelzugehörigen Nachschlagegegenstand*
- Artikeltypologie mit Bezug auf mediostrukturelle Aspekte (ohne mediostrukturelle Orientierung)
Oberstes TK: *Verhältnis von Wörterbuchartikeln zu mediostrukturellen Textsegmenten*
- Artikeltypologie mit Bezug auf mediostrukturelle Aspekte (mit mediostruktureller Orientierung)
Oberstes TK: *genuiner Zweck von verweisvermittelnden Wörterbuchartikeln*
- Artikeltypologie nach Art, Anzahl und Reihenfolge der Kommentare eines Wörterbuchartikels
Oberstes TK: *Art, Anzahl und Reihenfolge der Kommentare*
- Artikeltypologie mit Bezug auf die Datendistribution im Zweitkommentar
Oberstes TK: *Typ der artikelinternen Datendistribution*
- Artikeltypologie mit Bezug auf Textarchitekturen
Oberstes TK: *Vorhandensein einer Textarchitektur*
- Artikeltypologie mit Bezug auf den Artikelumfang
Oberstes TK: *relativer Artikelumfang*
- Artikeltypologie mit Bezug auf die Standardisierung
Oberstes TK: *Vorhandensein von Standardisierung.*

Insgesamt wurden mehr als 250 artikelformbedingte Typen von Wörterbuchartikeln unterschieden. Ein solches typologisches System ist nicht nur für die Wörterbuchanalyse von Interesse, sondern ein Lexikograph hat dadurch die Chance, für die meisten lexikographischen Vertextungsprobleme, die unter Berücksichtigung der Wörterbuchfunktionen (i.S.v. Wiegand 2001b) bei der Planung und Gestaltung von Wörterbuchartikeln auftreten können, ein geeigneten Artikeltyp zu finden. Der Ausschnitt aus dem empirischen Gegenstandsbereich, der in Wiegand 2003 berücksichtigt wurde, betrug 1200 Wörterbücher. 2008 wurde dieser empirische Gegenstandsbereich um ca. 100 Wörterbücher erweitert, darunter alle neueren englischen Lernerwörterbücher (vgl. Wiegand 2009b). Im OALD⁷ 2005 (jetzt auch im OALD⁸ 2010), im OSACD² 2010, im CALD³ 2008, im CD 2009, im COED 2009 und im POED⁶ 2006 fanden sich dann Wörterbuchartikel wie z.B. die in Abb. 2.

- Wa₁₃: **blind-fold** /'blaɪnfəʊld; NAmE -foʊld/ *noun, verb, adj., adv.*
 ■ **noun** something that is put over sb's eyes so they cannot see
 ■ **verb** ~ sb to cover sb's eyes with a piece of cloth or other covering so that they cannot see: *The hostages were tied up and blindfolded.*
 ■ **adj., adv.** (BrE) (also blind-fold-eɪʃ BrE, NAmE) with the eyes covered: *The reporter was taken blindfold to a secret location.* ◊ *I knew the way home blindfold* (= because it was so familiar). ◊ *I could do that blindfold* (= very easily, with no problems).
- Wa₁₄: **leap** /li:p/ *verb, noun*
 ■ **verb** (leapt, leapt /lept/) or (leaped, leaped) 1 [I, T] to jump high or a long way: + **adv./prep.** *A dolphin leapt out of the water.* ◊ *We leapt over the stream.* ◊ ~ **sth** *The horse leapt a five-foot wall.* 2 [I] + **adv./prep.** to move or do sth suddenly and quickly: *She leapt out of bed.* ◊ *He leapt across the room to answer the door.* ◊ *I leapt to my feet* (= stood up quickly). ◊ *They leapt into action immediately.* ◊ (figurative) *She was quick to leap to my defence* (= speak in support of me). ◊ *The photo seemed to leap off the page* (= it got your attention immediately). ◊ *His name leapt out at me* (= I saw it immediately). 3 [I] ~ (in sth) (from ...) (to...) to increase suddenly and by a large amount **SYD** shoot up: *The shares leapt in value from 476p to close at 536p.*
 4 [I] look before you leap (saying) used to advise sb to think about the possible results or dangers of sth before doing it ◊ more at CONCLUSION, HEART
 5 [I] leap at sth to accept a chance or an opportunity quickly and with enthusiasm **SYD** jump at: *I leapt at the chance to go to France.*
 ■ **noun** 1 a long or high jump: *a leap of six metres* ◊ *She took a flying leap and landed on the other side of the stream.* ◊ (figurative) *His heart gave a sudden leap when he saw her.* ◊ (figurative) *Few people successfully make the leap from television to the movies.* 2 ~ (in sth) a sudden large change or increase in sth: *a leap in profits* ◊ see also QUANTUM LEAP
 3 [I] by/in leaps and bounds very quickly; in large amounts: *Her health has improved in leaps and bounds.*
 a leap in the dark an action or a risk that you take without knowing anything about the activity or what the result will be
- Wa₁₅: **con-comi-tant** /kən'kɒmɪtənt; NAmE -'kɒ:m- / *adj., noun*
 ■ **adj.** (formal) happening at the same time as sth else, especially because one thing is related to or causes the other
 ■ **noun** (formal) a thing that happens at the same time as sth else

- wa₁₆: **for-mat** **AM** /'fɔ:mæt; NAmE 'fɔ:mæt/ *noun, verb*
 ■ **noun** **1** the general arrangement, plan, design, etc. of sth: *The format of the new quiz show has proved popular.*
2 the shape and size of a book, magazine, etc: *They've brought out the magazine in a new format.* **3** {computing} the way in which data is stored or held to be worked on by a computer
 ■ **verb** (-!-) **1** ~ sth to prepare a computer disk so that data can be recorded on it **2** ~ sth {technical} to arrange text in a particular way on a page or a screen
- wa₁₇: **blindfold** ■ **n.** a piece of cloth tied around the head to cover someone's eyes. ■ **v.** deprive of sight with a blindfold. ■ **adv.** with a blindfold covering the eyes.
 -ORIGIN C16: from obsolete *blindfell* 'strike blind, blindfold', from Old English *geblindfeilan*.
- wa₁₈: **thatch** /θætʃ/ *noun, verb*
 ■ **noun** **1** [U, C] dried STRAW, REEDS, etc. used for making a roof; a roof made of this material: *a roof made of thatch* ◦ *The thatch was badly damaged in the storm.* **2** [sing.] ~ of hair {informal} thick hair on sb's head
 ■ **verb** ~ sth to cover the roof of a building with thatch
 > *thatched* *adj.*: *They live in a thatched cottage.* ↗ VISUAL VOCAB page V14

Abb. 2: Wörterbuchartikel wa₁₃ – wa₁₆ und wa₁₈ aus OALD⁸ 2010; wa₁₇ aus OSACD² 2010

Wörterbuchartikel, wie die in Abb. 2, wurden in Wiegand 2003 und Wiegand/Feinauer 2010 nicht berücksichtigt; dadurch ergibt sich eine Reihe von theoriebedingten Artikelformproblemen, deren Lösung im Folgenden am Beispiel von wa₁₃ bis wa₁₈ in Abb. 2, wa₁₉ – wa₂₃ in Abb. 11 und wa₂₄ – wa₂₈ in Abb. 15 erläutert wird, einschließlich einiger Folgen, die die gewählten Lösungen für andere Teile der Theorie der Wörterbuchform und für die zugehörigen Darstellungsmethoden (i.S.v. Wiegand 2010c: 290ff) haben.

Bisher galt in der Theorie der Wörterbuchform für Wörterbuchartikel das Folgende: Jeder kondensierte Wörterbuchartikel hat obligatorisch einen Formkommentar und einen Zweitkommentar (z.B. einen semantischen Kommentar oder einen Kommentar zur Form und Semantik) und kann fakultativ einen oder mehrere weitere Kommentare aufweisen. Ein Wörterbuchartikel, der nur einen Form- und einen Zweitkommentar aufweist, ist ein *basaler* Wörterbuchartikel; weist er einen weiteren oder mehrere weitere Kommentare auf, liegt ein *erweiterter* Wörterbuchartikel vor; es gibt deutlich über 150 Typen von erweiterten (nichtkomplexen) Wörterbuchartikeln. Eine Ausnahme bilden nur die rudimentären Formartikel (i.S.v. Wiegand 2003: 264ff); diese weisen nur einen Formkommentar auf und finden sich z.B. in finalalphabetischen Wörterbüchern und besonders in älteren Rechtsschreibewörterbüchern (vgl. Wiegand 1994). Alle Kommentare sind — daran sei hier erinnert — unmittelbare Textkonstituenten eines Wörterbuchartikels, also die Artikelkonstituenten, und damit Angaben oder Angebotexte, die genau diesen formbestimmten Kommentarstatus haben. Mit wa₁₃ z.B. liegt dagegen ein Wörterbuchartikel vor, für den gilt: Er hat als unmittelbare Textkonstituenten einen Formkommentar sowie drei weitere „Artikelteile“, die *mit* oder — je nach der metalexikographischen Forminterpretation — *direkt nach* einem schwarzen Quadrat „■“

beginnen. Diese „Artikelteile“ erfüllen nicht alle Definitionskriterien für einen Kommentar. Denn für jeden Kommentar gilt: Mit ihm dürfen nur bestimmte, sprachtheoretisch festgelegte Eigenschaften des Lemmazeichens lexikographisch bearbeitet werden. Die „Artikelteile“ sind auch keine Subartikel, denn diese weisen stets entweder ein Volllemma oder ein Teillemma auf, die „Artikelteile“ aber nicht. Der Terminus, der wohl am besten trifft, worum es sich bei den drei „Artikelteilen“ handelt, lautet *Teilartikel*. Ein Teilartikel ist dann offensichtlich eine unmittelbare Textkonstituente eines komplexen Wörterbuchartikels, der stets zusammen mit mindestens einem weiteren Teilartikel auftritt; für den Terminus *Teilartikel* spricht weiterhin, dass sich ein Auslagerungsprozess beschreiben lässt, wie z.B. wa₁₃ aus drei Wörterbuchartikeln entstanden sein könnte, nämlich einem Substantiv-, einem Verb- und einem Adjektiv/Adverb-Artikel, deren Lemmata Wortklassenindices aufweisen, wie dies z.B. in MEDAL² 2007 und in LDCE⁵ 2009 der Fall ist, und zwar wie folgt: Zwei Lemmata und zwei der drei um eine verdichtete Sprachenidentifizierungsangabe (nämlich „NAme“) binnenerweiterte Aussprachenvariantenangaben werden nach links ausgelagert und bilden dann einen terminologisch so zu nennenden *teilartikelexternen Formkommentar* (FK.taex), in dessen textuellem Skopus drei Teilartikel liegen. Dass dieser Formkommentar dann allerdings um die Wortartangaben „noun“ und „verb“ sowie um die verdichteten Wortartangaben „adj.“ und „adv.“ erweitert wird, ist überflüssig (vgl. auch wa₁₇ in Abb. 2) und stellt nicht nur eine sinnlose Druckraumverschwendung dar, sondern schafft auch unklare Skopusbeziehungen. Der neue Artikeltyp, zu dem wa₁₃, wa₁₅ und wa₁₆ gehören, sowie weitere Wörterbuchartikel in den Abb. 11 u. 15, heißt oberbegrifflich *komplexer Wörterbuchartikel* und ist mit diesem Artikeltypennamen auch terminologisch von allen anderen Artikeltypen deutlich getrennt. Artikeltypen, deren zugehörige Artikel keine Teilartikel aufweisen, können dann oberbegrifflich *nichtkomplexe Wörterbuchartikel* genannt werden; dies ist allerdings nur in Kontexten erforderlich, in denen die Unterscheidung von *komplexen* und *nichtkomplexen* Wörterbuchartikeln argumentativ erforderlich ist. Komplexe Wörterbuchartikel mit zwei Teilartikeln (wie z.B. wa₁₅ und wa₁₆ in Abb. 2) weisen — wenn der Grad der Komplexität durch die Anzahl der Teilartikel bestimmt wird — den geringsten Komplexitätsgrad auf; sie heißen daher *einfach komplexe Wörterbuchartikel*. Den einfach komplexen stehen die *mehrfach komplexen Wörterbuchartikel* gegenüber; bei diesen unterscheidet man die *zweifach*, die *dreifach* und die *vierfach komplexen Wörterbuchartikel*. Zweifach komplexe Wörterbuchartikel weisen drei Teilartikel auf (wie z.B. wa₁₃ und wa₁₄ in Abb. 2). Dreifach komplexe Wörterbuchartikel weisen vier und vierfach komplexe Wörterbuchartikel weisen fünf Teilartikel auf; im OALD⁸ 2010 kommen dreifach komplexe Artikel nur relativ sehr selten vor, z.B. der Artikel zu *save*; sehr selten sind vierfach komplexe Wörterbuchartikel. Ein Beispiel ist der Artikel zu *double*.

Im Folgenden wird die Artikelform von komplexen Wörterbuchartikeln näher betrachtet, und zwar zuerst anhand von wa₁₃. Dazu wird zu-

nächst das komplexe reduzierte und erweiterte allgemeine Artikelkonstituentenstrukturbild (vgl. dazu WLWF-I 2010) von wa₁₃ in Abb. 3 angegeben.

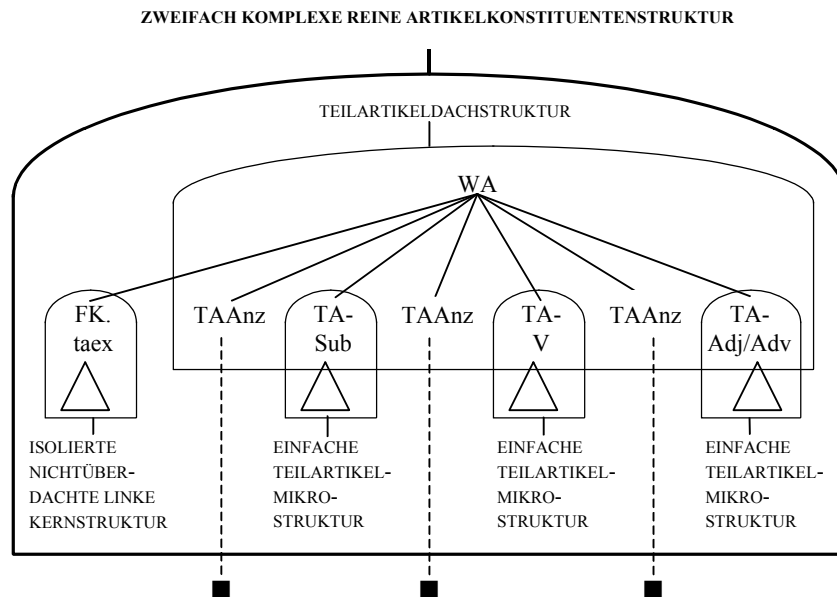


Abb. 3: Komplexes reduziertes und erweitertes allgemeines Artikelkonstituentenstrukturbild für hierarchische zweifach komplexe reine Artikelkonstituentenstrukturen, die zum Typ des zweifach komplexen Wörterbuchartikels gehören, wie wa₁₃ in Abb 2

Abkürzungen: FK.taex = teilartikelexterner Formkommentar; TA-Sub = Teilartikel zum Substantiv, TA-V = Teilartikel zum Verb; TA-Adj/Adv = komplexer Teilartikel zum Adjektiv und Adverb; TAAnz = Teilartikelanzeiger

Das komplexe reduzierte und erweiterte allgemeine Artikelkonstituentenstrukturbild in Abb. 3, das — um nur wenige zu nennen — auch für die Artikel zu den Lemmazeichen *crash* und *cross* im OALD⁸ 2010 gilt, lässt deutlich erkennen, dass die durch die geschwärzten Quadrate realisierten Teilartikelanzeiger nicht als Teile der Teilartikel gelten, sondern als Teile von wa₁₃, die zwischen den Teilartikeln stehen. Die Teilartikelanzeiger sind mithin keine mikrostrukturellen Textkonstituenten, sondern nur Textkonstituenten der hierarchischen Artikelkonstituentenstruktur von komplexen Wörterbuchartikeln, deren Substruktur die hierarchische Artikelmikrostruktur ist. Demgemäß hat das komplexe allgemeine Mikrostrukturbild die Form, wie sie sich in Abb. 4 zeigt.

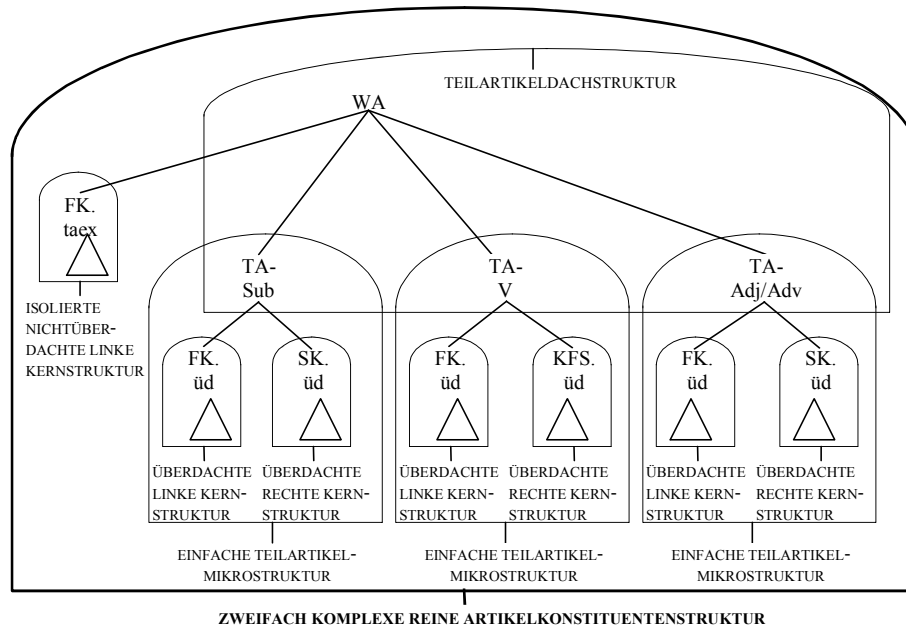


Abb. 4: Komplexes allgemeines Mikrostrukturbild für hierarchische zweifach komplexe reine Artikelmikrostrukturen, die zum Typ des zweifach komplexen Wörterbuchartikels (mit drei Teilartikeln) gehören, wie wa₁₃ in Abb. 2; *Anm.:* Teilartikel sind stets überdacht. In den Klassensymbolen für Teilartikel muss daher „üd“ nicht auftreten

Im Folgenden wird öfter von dem formalen Darstellungsmittel der Artikelstrukturschemata Gebrauch gemacht (vgl. dazu u.a. Wiegand 2010c: 313 u. WLWF-1 2010). In diesen werden Klassensymbole für mikrostrukturelle Textkonstituenten sowie folgende Zeichen verwendet: „:“ für *besteht aus* und „<“ für *geht voraus*.

Anhand komplexer allgemeiner Mikrostrukturbilder (wie z.B. dem in Abb. 4) lässt sich stets das einfache und das erweiterte Artikelstrukturschema ablesen. Ein *einfaches Artikelstrukturschema* berücksichtigt ausschließlich die unmittelbaren mikrostrukturellen Textkonstituenten. Bei nichtkomplexen Wörterbuchartikeln (wie z.B. wa₃₀ und wa₃₁) sind das nur Kommentare. Bei komplexen Wörterbuchartikeln sind dies dagegen stets der teilartikelexterne Formkommentar (FK.taex) sowie die Teilartikel (TA). Ein *erweitertes Artikelstrukturschema* berücksichtigt bei nichtkomplexen Wörterbuchartikeln die Subkommentare des Zweitkommentars und bei komplexen Wörterbuchartikeln die unmittelbaren Textkonstituenten der Teilartikel und damit die überdachten Kommentare, die auch in komplexen allgemeinen Mikrostrukturbildern von komplexen Wörterbuchartikeln Berücksichtigung finden (vgl. z.B. Abb. 4). Damit aber auch die teilartikelinternen Teilmikrostrukturen in den überdachten Zweikommentaren, nämlich die der überdachten semantischen Kommentare (SK.üd) und die der überdachten Kommentare zur

Form und Semantik (KFS.üd) berücksichtigt werden können, wird ein weiterer Typ von Artikelstrukturschema benötigt, der *expandiertes Artikelstrukturschema für komplexe Wörterbuchartikel* (kurz: *expandiertes Artikelstrukturschema*) heißen soll, und in dem die überdachten semantischen Subkommentare (SSK.üd) und die überdachten Subkommentare zur Form und Semantik (SKFS.üd) Berücksichtigung finden. Das expandierte Artikelstrukturschema für wa₁₆ (das anhand des komplexen expandierten allgemeinen Mikrostruktur- und Mikroarchitekturbilds in Abb. 7 abgelesen werden kann) hat die folgende Form WA : vFK.taex < vLzWK.taex < hFK.taex < TA-Sub [FK.üd < SK.üd [SSK.üd < SSK.üd < SSK.üd]] < TA-V [FK.üd < KFS.üd [SKFS.üd < SKFS.üd]].

Ein Forscher, der die Theorie der Wörterbuchform und ihre zugehörigen Methoden kennt, muss Artikelstrukturschemata natürlich nicht an allgemeinen Mikrostrukturbildern ablesen, sondern kann sie auch ohne diese konstruieren. So hat z.B. das expandierte Artikelstrukturschema für wa₁₄ in Abb. 2 die folgende Form: WA : FK.taex < TA-V [FK.üd < KFS.üd [SKFS.üd < SKFS.üd < SKFS.üd]] < PostK.komp [PostSK:Phras.üd < PostSK:PhrV.üd]] < TA-Sub [FK.üd < KFS.üd [SKFS.üd < SKFS.üd]] < PostK:Phras]].

(*Abkürzungen*: vFK.taex = teilartikelexterner vorderer Formkommentar; lZwK.taex = teilartikelexterner linker Zwischenkommentar; hFK.taex = teilartikelexterner hinterer Formkommentar; PostK.komp = komplexer Postkommentar; PostSK:Phras.üd = überdachter Postsubkommentar zur Phraseologie; PostSK:PhrV.üd = überdachter Postsubkommentar zu phrasal verbs)

Wenn anhand des Kontextes und/oder anhand des Artikelstrukturschemas selbst klar ist, dass ein Teilartikel entweder basal oder erweitert ist, können die einfachen Teilartikelklassensymbole, wie z.B. TA-Sub und TA-V u.a., sowohl für basale als auch für erweiterte Teilartikel Verwendung finden. Ist dies aber nicht klar, müssen erweiterte Teilartikelklassensymbole verwendet werden, wie z.B. für einen basalen Verb-Teilartikel das Symbol TA-V.ba und für einen rechts-erweiterten Substantiv-Teilartikel das Symbol TA-Sub.re. Es ist auch zugelassen, dass man die einfachen Symbole für Teilartikelklassen nur oberbegrifflich verwendet und in den Artikelstrukturschemata stets die erweiterten (vgl. dazu Wiegand 2011).

Bei der Beschreibung des formalen Darstellungsmittels der Artikelstrukturschemata müssen daher zukünftig auch die Klassensymbole für Teilartikel und überdachte Kommentare sowie Subkommentare Berücksichtigung finden (vgl. Wiegand 2010c: 313f u. 2011). Das einfache Artikelstrukturschema für wa₁₃ kann an dem allgemeinen Mikrostrukturbild in Abb. 4 abgelesen werden und hat folgende Form: WA : FK.taex < TA-Sub < TA-V < TA-Adj/Adv; das erweiterte Artikelstrukturschema für wa₁₃ weist folgende Form auf: WA : FK.taex < TA-Sub[FK.üd < SK.üd] < TA-V[FK.üd < KFS.üd] < TA-Adj/Adv[FK.üd < SK.üd].

Der Artikel wa₁₃ ist weiterhin vertikalarchitektonisch ausgebaut (i.S.v. Wiegand 2001a: 191ff), und zwar im deutlichen Unterschied zum ersten Teil in wa₁₇, der aus den Teilartikeln besteht. Die Trägermenge für die abstrakte Arti-

kelarchitektur von wa_{13} — sie heie $M_{Arch}(wa_{13})$ — ist elementenheterogen, da sie als Elemente die Klasse der teilartikelexternen Formkommentare (FK.taex) und drei Klassen von Folgen von Artikelkonstituenten aufweist. Sie kann daher wie folgt angegeben werden, mit „<“ fr *geht voraus*:

$$M_{Arch}(wa_{13}) = \{FK.taex, (TAAanz < TA-Sub), (TAAanz < TA-V), (TAAanz < TA-Adj/Adv)\}.$$

Definiert man auf $M_{Arch}(wa_{13})$ eine zweistellige textarchitektonische Relation vom Typ *oberhalb*-Relation mit dem Relationsterm *x ist (als Ganzes) oberhalb von y*, mit „x“ und „y“ als Variablen fr artikelinterne Textblcke, ist die abstrakte vertikale hybride Artikelkonstituentenarchitektur erhltlich, die zu wa_{13} gehrt, so dass wa_{13} zum Typ des hybrid vertikalarchitektonisch ausgebauten komplexen Wrterbuchartikels gehrt (vgl. Wiegand 2009a). Die abstrakte vertikale hybride Artikelkonstituentenarchitektur ist in Abb. 5 dargestellt.

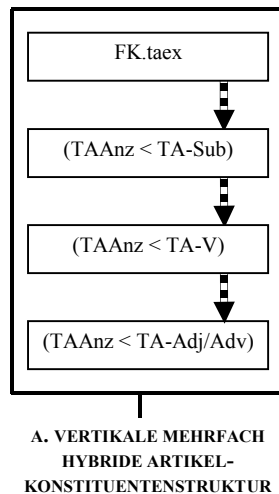


Abb. 5: *Komplexes allgemeines Artikelkonstituentenarchitekturbild fr die abstrakte vertikale mehrfach hybride Artikelkonstituentenarchitektur, die zum Typ des basalen Substantiv-Verb-Adjektiv / Adverb-Artikels gehren wie wa_{13} in Abb. 2 gehrt; Darstellungskonventionen: „x \dashrightarrow y“ bedeutet soviel wie *x ist oberhalb von y* (mit „x“ und „y“ als Variablen fr artikelinterne Textblcke); Abkrzungen: A = ABSTRAKT*

Kombiniert man das komplexe reduzierte und erweiterte allgemeine Artikelkonstituentenstrukturbild in Abb. 3 mit dem komplexen allgemeinen Artikelkonstituentenarchitekturbild in Abb. 5, verfgt man ber das komplexe reduzierte und erweiterte allgemeine Artikelkonstituentenstruktur- und Artikelkonstituentenarchitekturbild, das in Abb. 6 dargestellt ist.

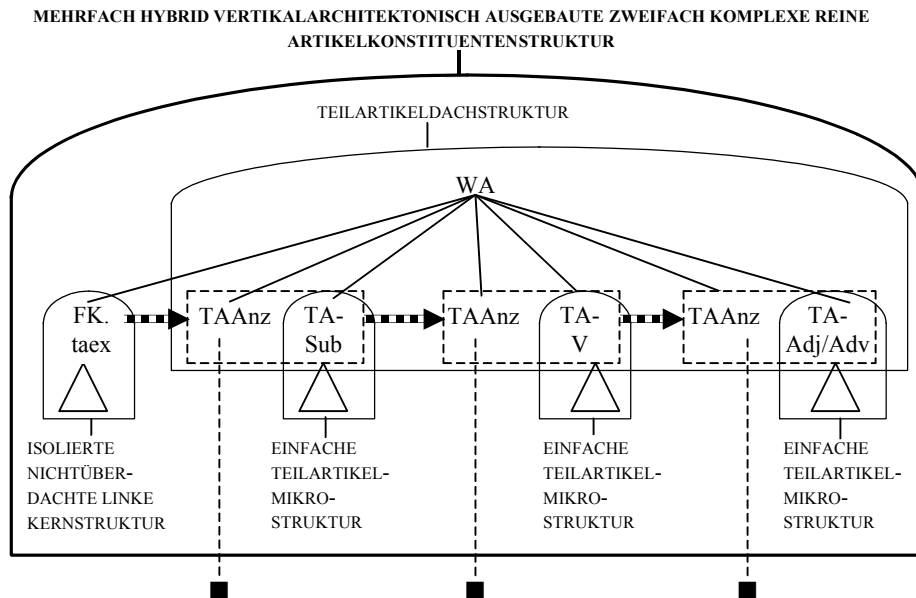


Abb. 6: *Komplexes reduziertes und erweitertes allgemeines Artikelkonstituentenstruktur- und Artikelkonstituentenarchitekturbild für die mehrfach hybrid vertikalarchitektonisch ausgebaute zweifach komplexe reine Artikelmikrostruktur, die zum Typ des basalen Substantiv-Verb-Adjektiv/Adverb-Artikels gehört wie wa₁₃ in Abb. 2*

Bis zur oben erwähnten Erweiterung des Gegenstandsbereichs der Theorie der Wörterbuchform galt das Folgende: Eine Artikelkonstituentenarchitektur gilt als hybrid, wenn in ihrer Trägermenge nicht nur Artikelkonstituenten, sondern auch Folgen von Artikelkonstituenten auftreten, so dass eine elementenheterogene Trägermenge gegeben ist (vgl. z.B. Wiegand 2010a) und weiterhin gilt: Eine Mikroarchitektur gilt als hybrid, wenn ihre Trägermenge in dem Sinne elementenheterogen ist, dass sowohl mikrostrukturelle Textkonstituenten als auch Folgen von solchen zu ihren Elementen gehören. In der Trägermenge $M_{Arch}(wa_{13})$ tritt nun aber eine mikrostrukturelle Textkonstituente auf, nämlich FK.taex, und Folgen, die aus einer Textkonstituente der Artikelkonstituentenstruktur, nämlich aus TAAanz, und aus einer mikrostrukturellen Textkonstituente bestehen, so dass sich z.B. eine Folge wie „TAAanz < TA-Sub“ ergibt (vgl. auch Abb. 5). Es liegt also in wa₁₃ eine vertikale Textarchitektur vor, die in anderer Weise hybrid ist, so dass ein neuer Textarchitekturtyp eingeführt werden muss, der zu komplexen Wörterbuchartikeln gehört und *vertikale mehrfach hybride Artikelkonstituentenarchitektur* heißen soll. Die architektonisch ausgebaute hierarchische Artikelmikrostruktur, die in konkreter und abstrakter Ausprägung bei wa₁₃ auftritt, gehört daher zum Typ der mehrfach hybrid ver-

tikalarchitektonisch ausgebauten zweifach komplexen reinen Artikelmikrostruktur.

In komplexen Wörterbuchartikeln, wie in wa₁₃ bis wa₁₇, treten — wie wir gesehen haben — keine Zweitkommentare auf, sondern stattdessen in deren Artikelposition (was später noch relevant werden wird) mindestens zwei Teilartikel, die — sieht man von den Teilartikelnanzeigern ab — unmittelbar aufeinander folgen. Das hat auch zur Folge, dass dann einige andere Typen von allgemeinen Strukturbildern (vgl. zu diesen u.a. Wiegand 2010c: 293ff und WLWF-1 2010) festgelegt werden müssen, wenn ein Typ zu denjenigen Typen von allgemeinen Artikelstrukturbildern gehört, die im Typennamen das Prädikat *expandiert* aufweisen, weil die Expandierung eines allgemeinen Artikelstrukturbildes für Textkonstituentenstrukturen von Wörterbuchartikeln darin besteht, dass auch die einfach mittelbaren Textkonstituenten des jeweiligen Zweitkommentars, also z.B. die verschiedenen Subkommentare zweiter Stufe, berücksichtigt werden. Entsprechend müssen, wenn expandierte Artikelstrukturbilder von komplexen Wörterbuchartikeln konstruiert werden, die einfach mittelbaren Textkonstituenten der Teilartikel berücksichtigt werden, also die überdachten semantischen Subkommentare und überdachten Subkommentare zur Form und Semantik (vgl. Abb. 7). Die neuen Typen von Artikelstrukturbildern für komplexe Wörterbuchartikel sind dann die Folgenden:

- komplexes expandiertes allgemeines Artikelkonstituentenstrukturbild
- komplexes expandiertes allgemeines Mikrostrukturbild
- komplexes expandiertes allgemeines Suchbereichsstrukturbild
- komplexes erweitertes und expandiertes allgemeines Artikelkonstituentenstrukturbild
- komplexes erweitertes und expandiertes allgemeines Mikrostrukturbild
- komplexes erweitertes und expandiertes allgemeines Suchbereichsstrukturbild
- komplexes expandiertes allgemeines Artikelkonstituentenstruktur- und Artikelkonstituentenarchitekturbild
- komplexes expandiertes allgemeines Mikrostruktur- und Mikroarchitekturbild (vgl. Abb. 7, Abb. 16 u. Abb. 17)
- komplexes expandiertes allgemeines Suchbereichsstruktur- und Suchbereichsarchitekturbild
- komplexes erweitertes und expandiertes allgemeines Artikelkonstituentenstruktur- und Artikelkonstituentenarchitekturbild
- komplexes erweitertes und expandiertes allgemeines Mikrostruktur- und Mikroarchitekturbild

- komplexes erweitertes und expandiertes allgemeines Suchbereichsstruktur- und Suchbereichsarchitekturbild.

Hier hat man damit ein deutliches Beispiel dafür, wie eine theoriesensitive Methodologie durch die Theorieerweiterung eine angemessene Erweiterung bei Darstellungsmethoden erfährt. Im übrigen gilt, dass alle Typennamen von allgemeinen Strukturbildern, von allgemeinen Struktur- und Architekturbildern sowie von allgemeinen Architekturbildern von Strukturen und Architekturen von komplexen Wörterbuchartikeln aus Gründen der einheitlichen terminologischen Differenzierung im Typennamen als ersten Namenteil eine Form des Prädikates *komplex* aufweisen.

Im Folgenden betrachten wir den Artikel wa_{16} in Abb. 2; wa_{16} unterscheidet sich von wa_{13} in der Artikelform vor allem dadurch, dass nur zwei Teilartikel gegeben sind, so dass ein einfach komplexer Wörterbuchartikel vorliegt, und weiterhin dadurch, dass das Lemmazeichen *format* sowohl als Substantiv als auch als Verb als polysem interpretiert ist. Im Unterschied zu wa_{13} weist wa_{16} eine hierarchische mehrfach hybrid vertikalarchitektonisch ausgebaute einfach komplexe reine Artikelkonstituentenstruktur (in konkreter und abstrakter Ausprägung) auf und als deren Substruktur eine hierarchische vertikalarchitektonisch ausgebaute einfach komplexe reine Artikelmikrostruktur (in konkreter und abstrakter Ausprägung). Dass die Artikelmikrostrukturen von wa_{13} und wa_{16} nicht mehrfach hybrid vertikalarchitektonisch, sondern vertikalarchitektonisch ausgebaut sind, liegt daran, dass die Teilartikelanzeiger keine mikrostrukturellen Textkonstituenten, sondern Textkonstituenten der Artikelkonstituentenstruktur sind, und die Teilartikel als mikrostrukturelle Textkonstituenten gelten, so dass die Trägermenge — sie heiße $M_{Arch}(wa_{16})$ — elementenhomogen ist und für die abstrakte Mikroarchitektur mit drei Elementen, die alle Klassen von mikrostrukturellen Textkonstituenten sind, wie folgt angegeben werden kann:

$$M_{Arch}(wa_{16}) = \{FK.taex, TA-Sub, TA-V\}.$$

Definiert man auf $M_{Arch}(wa_{16})$ eine Relation vom Typ der *oberhalb*-Relation, erhält man die abstrakte vertikale Mikroarchitektur von wa_{16} , die man — wie die von wa_{13} — in einem allgemeinen Mikroarchitekturbild darstellen kann.

Bereits bei der Betrachtung von wa_{13} wurden die jeweiligen aus mehreren methodischen Schritten bestehenden Strukturbildungsprozesse nicht beschrieben (vgl. zu diesen u.a. Wiegand 2010c u. Wiegand/Fuentes Morán 2009), sondern es wurden lediglich deren Ergebnisse mittels der Darstellungsmethoden dargestellt, die zur Theorie der Wörterbuchform gehören. Auf diese Weise wird (u.a. aus Platzgründen) auch weiterhin verfahren. Um eine erste rasche Übersicht über die hierarchische Artikelmikrostruktur von wa_{16} zu erhalten, wird in Abb. 7 zunächst das komplexe expandierte allgemeine Mikrostruktur- und Mikroarchitekturbild von wa_{16} präsentiert.

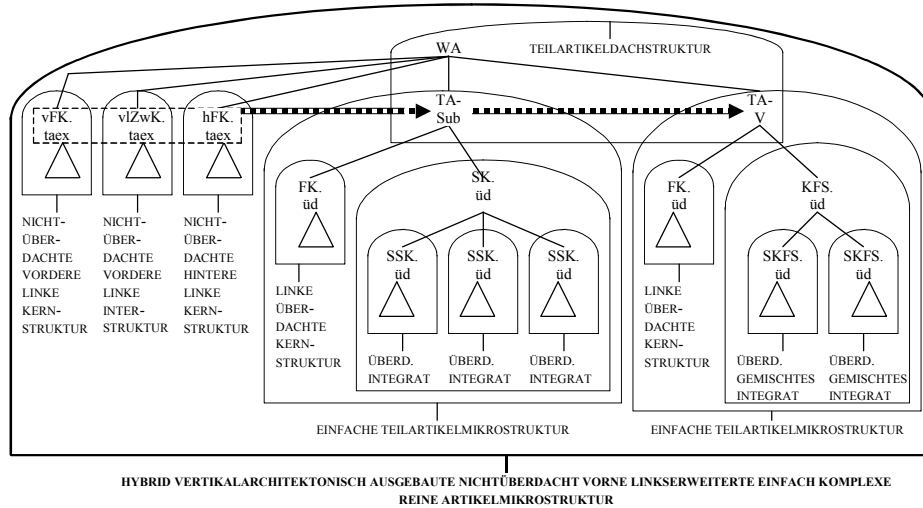


Abb. 7: Komplexes expandiertes allgemeines Mikrostruktur- und Mikroarchitekturbild für hierarchische hybrid vertikalarchitektonisch ausgebaute nichtüberdacht vorne linkerweiterte einfach komplexe reine Artikel mikrostrukturen, die zum Typ des teilartikelextern links binnenerweiterten basalen Substantiv-Verb-Artikels gehören, wie wa₁₆ in Abb. 2

Abkürzungen: überd. = überdacht. Alle weiteren Abkürzungen sind Klassennamen; vFK.taex = teilartikelexterner vorderer Formkommentar; vlZwK.taex = teilartikelexterner vorderer linker Zwischenkommentar; hFK.taex = teilartikelexterner hinterer Formkommentar; TA-Sub = Teilartikel zum Substantiv; TA-V = Teilartikel zum Verb; FK.üd = überdachter Formenkommentar; SK.üd = überdachter semantischer Kommentar; SSK.üd = überdachter semantischer Subkommentar; KFS.üd = überdachter Kommentar zur Form und Semantik; SKFS.üd = überdachter Subkommentar zur Form und Semantik; *Darstellungskonvention:* „x → y“ bedeutet soviel wie x ist oberhalb von y, mit „x“ und „y“ als Variable für mikrostrukturelle Textkonstituenten

Das komplexe expandierte allgemeine Mikrostruktur- und Mikroarchitekturbild in Abb. 7 zeigt deutlich die typkonstitutiven Grundstrukturen eines einfach komplexen OALD⁸-Wörterbuchartikels zu einem Wort, das angeblich (vgl. dazu 2.3) ein „academic word“ (AK) ist, zur zwei Wortarten (noun, verb) gehört (Sic!) und sowohl als „noun“ als auch als „verb“ als polysem interpretiert ist. Es gibt zahlreiche OALD⁸-Artikel, die eine analoge Artikel mikrostruktur und -architektur aufweisen und sich nur quantitativ bei den überdachten semantischen Subkommentaren unterscheiden, weil das Lemmzeichen eine andere Polysemiestruktur aufweist, z.B. die Artikel zu *contrast* und zu *feature*.

Das einfache Artikelstrukturschema für den Wörterbuchtyp, zu dem wa₁₆ gehört, das man an dem komplexen expandierten allgemeinen Mikrostruktur- und Mikroarchitekturbild in Abb. 7 ablesen kann, hat folgende Form: WA :

vFK.taex < vLzWk.taex < hFK.taex < TA-Sub < TA-V. Das expandierte Artikelstrukturschema für wa₁₆ ist bereits bekannt.

Im Folgenden wird die abstrakte hierarchische und isomorphe konkrete hybrid vertikalarchitektonisch ausgebaute nichtüberdacht vorne links erweiterte einfach komplexe reine Artikelmikrostruktur von wa₁₆ in einem (aus Platzgründen) zweigeteilten Strukturgraphen dargestellt, in dem der geordnete Baumgraph für die abstrakte Struktur und der für die isomorphe konkrete Struktur aufeinander abgebildet sind. Das komplexe expandierte allgemeine Mikrostruktur- und Mikroarchitekturbild in Abb. 7 ist ein nach festen Regeln hergestellter Auszug aus dem Baumgraph für die abstrakte Struktur in Abb. 8a und 8b.

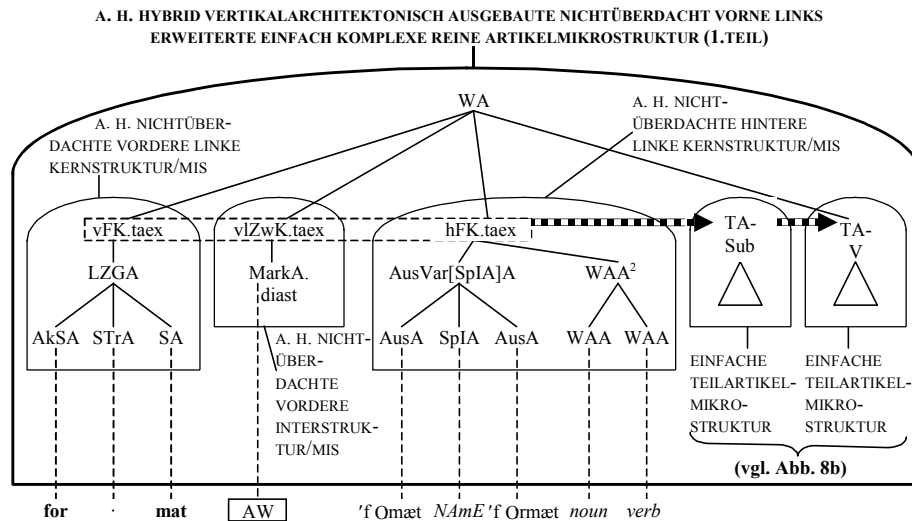


Abb. 8a: Nicht vollständig ausgeführter einfach kommentierter und architektonisch angereicherter Strukturgraph zur abstrakten (und isomorphen konkreten) hierarchischen hybrid vertikalarchitektonisch ausgebauten nichtüberdacht vorne linkerweiterten einfach komplexen reinen Artikelmikrostruktur, die zum Typ des teilartikelextern vorne links binnerweiterter Substantiv-Verb-Artikels gehört, wie wa₁₆ in Abb. 2

Abkürzungen: A = ABSTRAKT; H = HIERARCHISCH; _iAB_j = Angabeblock (für Nullangaben in der konkreten Struktur); MIS = MIKROSTRUKTUR (TEILARTIKEL-DACHSTRUKTUR/MIS ist zu lesen wie *Teilartikeldachstruktur als Teilstruktur der Mikrostruktur*). Alle folgenden Abkürzungen sind Klassensymbole; LZGA = Lemmzeichengestaltangabe; AkSA = Akzentsilbenangabe; STrA = Silbentrennungsangabe; SA = Silbenangabe; MarkA.diast = diastratische Markierungsangabe; Aus[SpIA]A = um eine Sprachenidentifizierungsangabe binnerweiterte Aussprachevariantenangabe; die eckige Klammer zeigt an, dass es sich um eine nichtelementare Angabe handelt, die funktional-positional segmentierbar ist (vgl. u. a. Wiegand 1991 u. 2005). AusA = Ausspracheangabe; WAA² = homosegmentäre Wortartenangabe (vgl. auch Abb. 7)

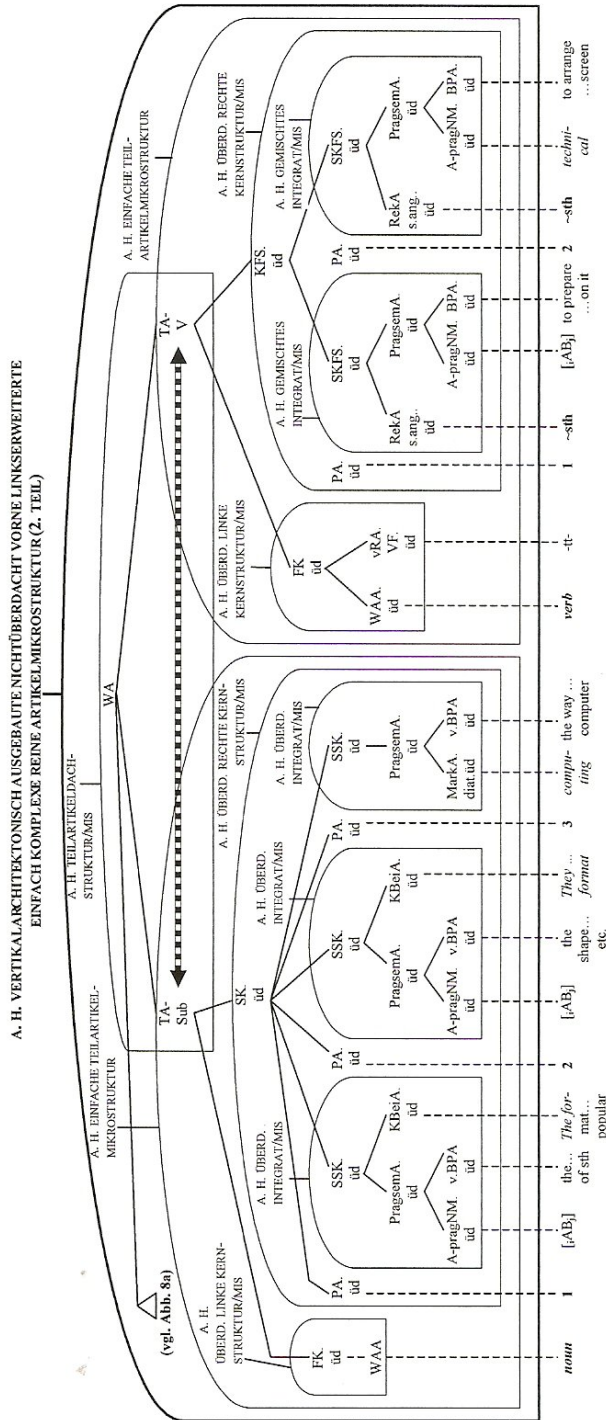


Abb. 8b: Nicht vollständig ausgeführter einfach kommentierter und architektonisch angereicherter Strukturgraph zur abstrakten (und isomorphen konkreten) hierarchischen hybrid vertikalarchitektonisch ausgebauten nichtüberdacht vorne links erweiterter einfach komplexen reinen Artikelmikrostruktur, die zum Typ des teilarthiklextern vorne links binnerweiteren Substantiv-Verb-Artikel gehört, wie wa₁₆ in Abb. 2

Abkürzungen: PA.üd = Polysemieangabe (als Teil eines n-Polysemieangabentupels mit $n > 2$; PragsemA.üd = überdachte pragmatisch-semantische Angabe; KBeiA.üd = überdachte Kompetenzbeispielangabe; A-pragNM.üd = überdachte Angabe zur pragmatischen Nullmarkierung; v.BPA.üd = überdachte Bedeutungsparaphrasenangabe; Mark.diat.üd = überdachte diatechnische Markierungsangabe; FK.üd = überdachter Formkommentar; RA-VF.üd = überdachte Rechtschreibangabe für Vergangenheitformen; v.RekA.s.ang.üd = überdachte semantisch angereicherte verdichtete Rektionsangabe

Im Folgenden wird erläutert, warum sowohl in wa₁₃ als auch in wa₁₆ im Teilartikel zum Substantiv ein semantischer Kommentar (SK; vgl. Abb. 7 und Abb. 8b) und im Teilartikel zum Verb ein Kommentar zur Form und Semantik auftritt. Das Substantiv *blindfold* ist „zählbar“ (a countable noun). Also müsste eigentlich nach der Wortartangabe „*noun*“ „*C*“, die verdichtete Angabe der Zählbarkeit, stehen. Diese wurde nicht etwa vergessen, sondern in der so genannten „Reference Section“ im Nachspann des OALD⁸ 2010 (die z.T. als ein lexikographisch gut getarnter, weil nicht akzessiver Teil der Benutzungshinweise im Vorspann verstanden werden muss!) findet man auf der Seite R11, auf die nicht extern zugriffen werden kann, sondern die man (sozusagen wie Ostereier) suchen muss (vgl. Wiegand 2009b), folgenden Textausschnitt: „Countable nouns are the most common type of noun. If they have only one meaning [wie in wa₁₃], or if all the meanings are countable [wie in wa₁₆] they are just marked *noun*“.

Diese ausgesprochen benutzerunfreundliche Lösung, mit der offenbar Druckraum eingespart werden soll, steht im auffälligen Kontrast zu der Druckraumverschwendung durch die homosegmentären Wortartangaben in den teilartikelexternen Formkommentaren. Für die Theorie der Wörterbuchform bedeutet dies, dass zwei Typen von Teilartikeln zu Substantiven unterschieden werden müssen. Ein Typ mit semantischem Kommentar und ein Typ mit Kommentar zur Form und Semantik, der an wa₁₈ in Abb.2 ausgeprägt ist. Bei den Teilartikeln zum Verb finden sich dagegen stets Kommentare zur Form und Semantik, weil sich entweder eine Rektionsangabe (z.B. „*I*“ oder „*T*“) findet, oder eine semantisch angereicherte Rektionsangabe mit Kohäsionsanweisung (wie z.B. ~*sb* oder ~*sth*), oder eine Angabe der Anschlusspräposition, (wie z.B. ~ *out*), oder eine grammatische Angabe eines anderen Typs.

Es folgen nun noch weitere Erläuterungen zu den Abb. 8a und 8b. Die diastatische Markierungsangabe „*AW*“ ist keine Angabe zur Form des (bzw. der!) Lemmazeichen(s). Sie kann deswegen auch nicht in einem Formkommentar stehen; vielmehr wird durch sie der teilartikelexterne Formkommentar (FK.taex) in einen vorderen und hinteren teilartikelexternen Formkommentar aufgespalten, so dass „*AW*“ den vorderen linken teilartikelexternen Zwischenkommentar bildet. Eine solche Formkommentaraufspaltung ist auch in zahlreichen Wörterbüchern in Artikeln, die keine komplexen Wörterbuchartikel sind, üblich, und es können Angaben, die zu recht verschiedenen Angabetypen gehören (öfters solche zur Herkunft oder zur Etymologie), im linken Zwischenkommentar von nichtkomplexen Wörterbuchartikeln stehen. Die Klasse der Angabe zur pragmatischen Nullmarkierung ist eine Unterklasse der Klasse der Nullangaben; das sind solche Angaben, deren Angabegegenstand (i.S.v. Wiegand 2005: 215f u. WLWF-I 2010) ein kundiger Benutzer anhand eines Leerzeichens in einer bestimmten Artikelposition erschließen kann. So ist in den semantischen Subkommentaren der Teilartikel zu Substantiven die Artikelposition für diasystematische Markierungsangaben (zu diesen vgl. Wiegand 2010b) die Position nach der Polysemieangabe (vgl. z.B. „*computing*“ in 3. SSK.üd in wa₁₆), und in Teilartikeln zu Verben ist es die Artikelposition nach der semantisch angereicherten Rektionsangabe, die zu den Angaben mit Kohäsionsanweisung gehört (vgl. z.B.

„technical“ in 2. SSK.üid in wa₁₆). Steht in den genannten Artikelpositionen keine diasystematische Markierungsangabe, dann kann der kundige Benutzer-in-actu schließen, dass das Lemmazeichen, wenn es die im überdachten semantischen Subkommentar paraphrasierte Bedeutung aufweist, in allen im OALD⁸ 2010 berücksichtigten Markierungsdimensionen nullmarkiert ist. Für diesen Fall steht als terminale Konstituente der konkreten Artikelmikrostruktur ein Angabeblank (AB); dieser weist Nachbarschaftsvariablen „i“ und „j“ auf, die mit den Klassennamen der linken und rechten Nachbarangaben belegt werden können (z.B. PA.üidABv.BPA.üid), wodurch die Angabeblanks positional eindeutig identifiziert sind, was vor allem für den Fall, dass es mehrere in einem Artikel gibt, unbedingt erforderlich ist. Die Variablenbelegung wurde in Abb. 8b (aus Platzgründen) nicht vorgenommen. Dass Benutzer aus leeren Artikelpositionen über Eigenschaftsausprägungen beim Lemmazeichen gültige Schlüsse ziehen können, ist in den Benutzungshinweisen und den „Reference Section“ nicht erklärt und zeigt, dass die Artikelautoren entweder ihre eigene Artikelform nicht verstanden haben oder den Benutzer nicht ausreichend informieren. Auch die verdichteten semantisch angereicherten Rektionsangaben von Typ „~sth“ und „~sb“ sind in den Benutzungshinweisen (im „Key to dictionary entries“ im Vorspann) nicht erklärt.

Üblicherweise gilt für Wörterbuchartikel, die nicht zu den komplexen Wörterbuchartikeln gehören, dass alle formkommentarinternen Angaben skopuseröffnende Angaben sind. Ihr textueller Skopus ist der Zweitkommentar; das bedeutet, dass das, was mit einer skopuseröffnenden Angabe im Formkommentar angegeben wird, auch im Zweitkommentar gilt, es sei denn, es treten skopusbeschränkende Angaben auf (vgl. zu diesen Wiegand 2005: 283ff). Beispielsweise gelten die Aussprachen von *format*, die mit den beiden Ausspracheangaben in wa₁₆ angegeben sind, auch für *format* in der Kompetenzbeispielangabe im zweiten überdachten semantischen Subkommentar des ersten Teilartikels. Auch die Angabe „AW“ gilt für beide Teilartikel. Anders ist dies allerdings mit den beiden Wortartangaben „noun“ und „verb“ in wa₁₃. Sie haben — je nach Argumentation — keinen Skopus oder einen heterogenen; für den letzteren Fall beschränkt dann „noun“ im ersten Teilartikel, den Skopus von „verb“ im teilartikelexternen Formkommentar und „verb“ im zweiten Teilartikel den Skopus von *noun* im teilartikelexternen Formkommentar, so dass man je nach ihrer Position von skopuseröffnenden und skopusbeschränkenden Wortartangaben sprechen muss, woran man erkennen kann, wie unprofessionell diese lexikographische Bearbeitung ist.

Wahrscheinlich ist es aber wohl angemessener, sogar von einem *systematischen Artikelformfehler* zu sprechen, zumal die Wortartangaben im teilartikelexternen Formkommentar überflüssig sind. Man braucht in den Benutzungshinweisen nur festzustellen, dass wenn Teilartikelanzeiger der Form „■“ am Zeilenanfang auftreten, mit dem Lemma die Form von mindestens zwei Lemmazeichen genannt wird, die zu verschiedenen Wortklassen gehören (vgl. dazu 2.3).

Ein weiteres Problem besteht darin, dass die Reihenfolge der Teilartikel von komplexen Wörterbuchartikeln wechselt. In wa₁₃ kommt erst der Teilartikel zum Substantiv, dann der zum Verb und zuletzt der zum Adjektiv und Adverb. In wa₁₄ ist die Reihenfolge so: Verb-Teilartikel, Substantiv-Teilartikel und in wa₁₅

kommt zuerst der Adjektiv- und dann der Substantiv-Teilartikel. Alle anderen möglichen Reihenfolgen sind im OALD⁸ auch vertreten. Warum sich die Reihenfolge ändert, ist nirgends (oder möglicherweise an einer lexikographisch gut getarnten Stelle) erklärt. Sollte es (wie wir vermuten) die Frequenz sein, die die jeweilige Reihenfolge bestimmt, dann ist das benutzerunfreundlich. Kein Benutzer — das weiß man aus der Wörterbuchbenutzungsforschung — der punktuelle Konsultationshandlungen ausführt, hat einen Vorteil von einer stets wechselnden Anordnung der Teilartikel, und die Frequenz interessiert den Benutzer überhaupt nicht. In mehreren englischen Lernerwörterbüchern herrscht ein für Benutzer *nutzloser Frequenz-Fetichismus*. Hier — und nicht nur hier, wie wir noch genauer sehen werden — wird gegen ein lexikographisches Grundprinzip verstoßen, nämlich das Prinzip der durchgängigen gleichartigen Standardisierung für kondensierte Wörterbuchartikel (vgl. Wiegand 1997). Die größtmögliche Gleichmäßigkeit der Form bei gleichartigen Wörterbuchgegenstandsaspekten ist das, was den Benutzer zum erfolgreichen Wörterbuchbenutzer macht, und weiterhin ist dies die Eigenschaft, die es dem Lexikographen ermöglicht, konsistente Benutzungshinweise zu formulieren, die einfach sind und für alle Fälle gelten.

Bevor nun noch weitere Beispiele von komplexen Wörterbuchartikeln betrachtet werden, sei dargestellt, dass und wie mittels der Theorie der Wörterbuchform *sämtliche denkbaren theoriebedingten Wörterbuchformprobleme*, die im Zusammenhang mit dem Typ des komplexen Wörterbuchartikels überhaupt auftreten können, nicht nur, wenn sie tatsächlich auftreten, systematisch gelöst werden können, sondern es gilt: Welche theoriebedingten Wörterbuchformprobleme überhaupt auftreten können, kann systematisch vorausgesagt werden. In dieser Hinsicht hat die Theorie der Wörterbuchform sogar prognostische Kraft. Um dies zu zeigen, wird zunächst auf die Artikeltypologie nach Art, Anzahl und Reihenfolge der Kommentare eines Wörterbuchartikels zurückgegriffen (vgl. oben 2.1, Wiegand 2003: 259-282 und Wiegand/Feinauer 2010).

Dieser Rückgriff erfolgt so, dass die komplexen Wörterbuchartikel teilweise analog zu den nichtkomplexen Artikeln typologisiert werden und weiterhin so, dass gezeigt wird, dass die Teilartikel eines komplexen Artikels im Prinzip genauso typologisiert werden können wie die nichtkomplexen kondensierten Wörterbuchartikel, also wie die, die keine Teilartikel haben, sondern nur Kommentare aufweisen, und zwar mit folgender Einschränkung: Als Zweitkommentare werden nur semantische Kommentare und Kommentare zur Form und Semantik berücksichtigt, weil die Kommentare der anderen vier Typen (relativ zur derzeitigen Kenntnislage), die als Zweitkommentare in nichtkomplexen Wörterbuchartikeln auftreten können, in komplexen Wörterbuchartikeln nicht auftreten. Das bedeutet, dass bei den basalen Wörterbuchartikeln, also denen, die aus nur zwei Kommentaren bestehen, nämlich einem Form- und einem Zweitkommentar, nur die zwei Typen berücksichtigt werden, für die die beiden einfache Artikelstrukturschemata $WA : FK < SK$ und $WA : FK < KFS$ gelten.

Nach Wiegand (2003: 259ff) gibt es zu jedem Typ von basalen Wörterbuchartikeln 31 Typen von erweiterten Wörterbuchartikeln. Für den basalen Wörterbuchtyp mit dem Artikelstrukturschema $WA : FK < SK$ sind alle 31 Typen in der europä-

schen Lexikographie in Wörterbuchartikeln realisiert und inzwischen systematisch analysiert. Die Zahl 31 zählt nur die Untertypen zweiter Stufe des Typs des erweiterten Wörterbuchartikels und nicht die (drei) Untertypen erster Stufe (vgl. Abb. 9).

In Abb. 9 ist der Typologieausschnitt präsentiert. Die Zahlen 1-31 verweisen auf die einfachen Artikelstrukturschemata, die auf die Abb. 9 folgen.



Abb. 9: *Partiell kommentierter Typologiegraph zu einem Ausschnitt aus einer Typologie von kondensierten Wörterbuchartikeln, die Teil eines typologischen Systems für Wörterbuchartikel ist*

Abkürzungen: TK = Typologiekriterium; WA = Wörterbuchartikel; SK = semantischer Kommentar; KFS = Kommentar zur Form und Semantik; *Darstellungskonventionen:* „x — y“ bedeutet (von unten nach oben gelesen) soviel wie *x ist ein Untertyp von y* mit „x“ und „y“ als Bezeichnungen von Wörterbuchtypen; „ \leftarrow “ bedeutet soviel wie *die Anwendung des TK führt zu der Unterteilung*

Es folgen nun die zur Abb. 9 gehörenden 31 einfachen Artikelstrukturschemata für erweiterte Wörterbuchartikel, die Erweiterungen eines basalen Wörterbuchartikels sind, für den das Artikelstrukturschema WA : FK < SK gilt (mit „<“ für *geht voraus* und mit „:“ für *besteht aus*; vgl. Wiegand 2003: 259ff).

- 1 WA : PräK < FK < SK
- 2 WA : FK < SK < PostK
- 3 WA : PräK < FK < SK < PostK
- 4 WA : vFK < lZwK < hFK < SK
- 5 WA : FK < mZwK < SK
- 6 WA : FK < vSK < rZwK < hSK
- 7 WA : vFK < lZwK < hFK < mZwK < SK
- 8 WA : FK < mZwK < vSK < rZwK < hSK
- 9 WA : vFK < lZwK < hFK < vSK < rZwK < rSK
- 10 WA : vFK < lZwK < hFK < mZwK < vSK < rZwK < hSK
- 11 WA : PräK < vFK < lZwK < hFK < SK
- 12 WA : PräK < FK < mZwK < SK
- 13 WA : PräK < FK < vSK < rZwK < hSK
- 14 WA : PräK < vFK < lZwK < hFK < mZwK < SK
- 15 WA : PräK < FK < mZwK < vSK < rZwK < rSK
- 16 WA : PräK < vFK < lZwK < hFK < vSK < rZwK < rSK
- 17 WA : PräK < vFK < lZwK < hFK < mZwK < vSK < rZwK < hSK
- 18 WA : FK < vSK < rZwK < hSK < PostK
- 19 WA : FK < mZwK < SK < PostK
- 20 WA : vFK < lZwK < hFK < SK < PostK
- 21 WA : FK < mZwK < vSK < rZwK < hSK < PostK
- 22 WA : vFK < lZwK < hFK < mZwK < SK < PostK

- 23 WA : vFK < lZwK < hFK < vSK < rZwK < hSK < PostK
 24 WA : vFK < lZwK < hFK < mZwK < vSK < lZwK < hSK < PostK
 25 WA : PräK < vFK < lZwK < hFK < SK < PostK
 26 WA : PräK < FK < vSK < rZwK < hSK < PostK
 27 WA : PräK < FK < mZwK < SK < PostK
 28 WA : PräK < vFK < lZwK < hFK < mZwK < SK < PostK
 29 WA : PräK < FK < mZwK < vSK < rZwK < hSK < PostK
 30 WA : PräK < vFK < lZwK < hFK < vSK < rZwK < hSK < PostK
 31 WA : PräK < vFK < lZwK < hFK < mZwK < vSK < rZwK < hSK < PostK.

Die Abkürzungen für Kommentare sind Klassensymbole und wie folgt aufzulösen: PräK = Präkommentar; FK = Formkommentar; vFK = vorderer Formkommentar; hFK = hinterer Formkommentar; lZwK = linker Zwischenkommentar; mZwK = mittlerer Zwischenkommentar; rZwK = rechter Zwischenkommentar; SK = semantischer Kommentar; vSK = vorderer semantischer Kommentar; hSK = hinterer semantischer Kommentar; PostK = Postkommentar

Wenn man in den 31 einfachen Artikelstrukturschemata für Artikeltypen, die den basalen Wörterbuchtyp mit dem Artikelstrukturschema WA : FK < SK erweitern, alle Artikelstrukturschemata, in denen das Klassensymbol „SK“ auftritt, so verändert, dass „SK“ durch „KFS“, das Klassensymbol für die Klasse der Kommentare zur Form und Semantik, ersetzt wird, verfügt man über alle Artikelstrukturschemata von erweiterten Typen von Wörterbuchartikeln, die den basalen Artikeltyp, für den das Artikelstrukturschema WA : FK < KFS gilt, erweitern. Das sind dann ebenfalls 31 Artikelstrukturschemata, z.B. die beiden Folgenden:

WA : PräK < FK < KFS

WA : FK < KFS < PostK

Wir verfügen bereits über die beiden einfachen Artikelstrukturschemata für wa₁₃ und wa₁₆, also über die beiden Folgenden:

WA : FK.taex < TA-Sub < TA-V < TA-Adj/Adv (für wa₁₃),

WA : vFK.taex < lZwK.taex < hFK.taex < TA-Sub < TA-V (für wa₁₆).

Es wurde weiterhin oben festgestellt, dass die Teilartikel als mikrostrukturelle Textkonstituenten gelten und dass unmittelbare Folgen von mindestens zwei Teilartikeln (wie z.B. „TA-Sub < TA-V“) so betrachtet werden können, dass sie in der Position des Zweitkommentars eines Wörterbuchartikels stehen, der keine Teilartikel aufweist, und mithin kein komplexer Wörterbuchartikel ist.

Dies bedeutet nun das Folgende: Der Artikel wa_{13} in Abb. 2 lässt sich als *basaler zweifach komplexer Wörterbuchartikel* und wa_{15} in Abb. 2 lässt sich als *basaler einfach komplexer Wörterbuchartikel* auffassen. Diese beiden Wörterbuchtypen lassen sich wie folgt definieren:

- Def. 1:** Ein *basaler einfach komplexer Wörterbuchartikel* liegt vor genau dann, wenn ein Wörterbuchartikel aus einem teilartikelexternen Formkommentar besteht, auf den zwei basale Teilartikel folgen.
- Def. 2:** Ein *basaler zweifach komplexer Wörterbuchartikel* liegt vor genau dann, wenn ein Wörterbuchartikel aus einem teilartikelexternen Formkommentar besteht, auf den drei basale Teilartikel folgen.

Für basale dreifach und basale vierfach komplexe Wörterbuchartikel lassen sich analoge Definitionen formulieren. Die in Def. 1 und Def. 2 definierten Artikeltypen sind Untertypen des Artikeltyps des basalen komplexen Wörterbuchartikels, der wie folgt definiert ist:

- Def. 3:** Ein *basaler komplexer Wörterbuchartikel* liegt vor genau dann, wenn ein Wörterbuchartikel aus einem teilartikelexternen Formkommentar besteht, auf den mindestens zwei basale Teilartikel folgen.

Man kann nun das Folgende erkennen: Bei wa_{16} handelt es sich offenbar um einen *erweiterten* einfach komplexen Wörterbuchartikel. Vergleicht man nämlich sein einfaches Artikelstrukturschema

$$WA : \nu FK.taex < \nu ZwK.taex < hFK.taex < TA-Sub < TA-V$$

mit dem Artikelstrukturschema Nr. 4 für nichtkomplexe Wörterbuchartikel, die zum Typ der links binnenerweiterten Wörterbuchartikel gehören (vgl. Abb. 9), nämlich mit $WA : \nu FK < LZwK < hFK < SK$, und substituiert die Teilartikelfolge „TA-Sub < TA-V“ für „SK“ (weil ja die Teilartikelfolgen in der Position des Zweitkommentars stehen) und geht weiterhin von νFK zu $\nu FK.taex$, von $LZwK$ zu $\nu ZwK.taex$ und von hFK zu $hFK.taex$ über, ist das einfache Artikelstrukturschema von wa_{16} gegeben, so dass man nun — in Analogie zu dem Artikeltyp des links binnenerweiterten Wörterbuchartikels, der bei nichtkomplexen Wörterbuchartikeln ausgeprägt ist, sagen kann: der komplexe Wörterbuchartikel wa_{16} gehört zum Artikeltyp des teilartikelextern vorne links binnenerweiterten einfach komplexen Wörterbuchartikels; *vorne links* muss es deswegen heißen, weil es auch *hinten links* binnenerweiterte komplexe Wörterbuchartikel gibt, wie wir noch genauer sehen werden. Der Artikeltyp des teilartikelextern vorne links binnenerweiterten einfach komplexen Wörterbuchartikels ist wie folgt definiert:

- Def. 4:** Ein *teilartikelextern vorne links binnenerweiterter einfach komplexer Wörterbuchartikel* liegt vor genau dann, wenn ein Wörterbuchartikel einen aufgespalteten teilartikelexternen Formkommentar aufweist,

so dass auf einen teilartikelexternen vorderen Formkommentar ein teilartikelextern vorderer linker Zwischenkommentar und auf diesen ein teilartikelextern hinterer Formkommentar folgt, auf den zwei Teilartikel folgen.

Lässt man zu, dass im Definiens von Definitionen von Artikeltypen einfache Artikelstrukturschema mit Variablen für Teilartikeltypen (TA_1 , TA_2 , TA_3) zugelassen sind, dann kann statt Def. 4 folgende Def. 4a formuliert werden:

Def. 4a: Ein *teilartikelextern vorne links binnenerweiterter einfach komplexer Wörterbuchartikel* liegt vor genau dann, wenn auf einen Wörterbuchartikel folgendes einfache Artikelstrukturschema zutrifft: $WA : vFK.taex < vlZwK.taex < hFK.taex < TA_1 < TA_2$.

In der Def. 4a ist dann „ TA_1 “ eine Variable für TA-Sub, TA-V, TA-Adj, TA Adj/Adv (und weitere Teilartikeltypen) als ersten Teilartikel und „ TA_2 “ eine Variable für TA-Sub, TA-V, TA-Adj, TA-Adj/Adv (und weitere Teilartikeltypen) als zweiten Teilartikel. Diese Regelung ist für das OALD⁸ 2010 erforderlich, weil die Reihenfolge der Teilartikeltypen wechselt.

Im Folgenden wird der Frage nachgegangen, welche weiteren teilartikelextern erweiterten komplexen Wörterbuchartikel sinnvollerweise unterschieden werden können. Eine solche Fragestellung ist zunächst ganz unabhängig davon, ob solche Wörterbuchartikel im empirischen Gegenstandsbereich tatsächlich auftreten oder nicht. Allerdings gilt eher dies: Hat ein Wörterbuchforscher ein klares Konzept von einem Artikeltyp, dann findet er leichter Wörterbuchartikel, an denen dieser Typ ausgeprägt ist, als wenn er ohne Konzeption eines Artikeltyps Wörterbuchartikel betrachtet, um die Typen zu finden. Denn diese rufen dem Forscher nicht zu: „Hallo, ich bin eine Ausprägung des Typs T_1 “. Vielmehr geht es gerade umgekehrt vor sich: Der Forscher stellt fest: „Du Artikel wa_1 gehörst ab jetzt zum Typ T_1 “. Das soll kein Witz sein, sondern eine Aussage in „lockerer“ Form über das Verhältnis von Theorie und Empirie in der Wörterbuchforschung, wenn es um Aspekte der Wörterbuchform geht.

Zunächst ist klar, dass die Konstruktion einer vollständigen Analogie zu den erweiterten nichtkomplexen Wörterbuchartikeln in Abb. 9 zwar möglich wäre, aber zu wahrscheinlich unbrauchbaren Artikeltypen führen würde, denn dass z.B. die Teilartikel aufgespalten werden und dann ein teilartikelextern rechter Zwischenkommentar eingeschoben wird, führt zu keiner brauchbaren Struktur. Brauchbare hierarchische Artikelmikrostrukturen für erweiterte komplexe Wörterbuchartikel ergeben sich aber, wenn man teilartikelexterne mittlere Zwischenkommentare vorsieht.

Bei den einfach komplexen Wörterbuchartikeln gibt es dann gerade einen mittleren Zwischenkommentar, und es liegt dann ein Wörterbuchartikel vor, der zum Typ des teilartikelextern zentralerweiterten einfach komplexen Wörterbuchartikels gehört. Dieser Typ ist wie folgt definiert:

Def. 5: Ein *teilartikelextern zentralerweiterter einfach komplexer Wörterbuchartikel* liegt vor genau dann, wenn zwischen dem ersten und dem zweiten Teilartikel ein *teilartikelexterner mittlerer Zwischenkommentar* steht.

Bei den zweifach komplexen Wörterbuchartikeln müssen drei teiltexterne Erweiterungen mit einem mittleren Zwischenkommentar berücksichtigt werden:

- (1) Ein *teilartikelexterner linker mittlerer Zwischenkommentar* (lmZwK.taex) wird zwischen dem ersten und dem zweiten Teilartikel eingeschoben. Dann liegt ein *teilartikelextern links-zentralerweiterter zweifach komplexer Wörterbuchartikel* vor.
- (2) Ein *teilartikelexterner rechter mittlerer Zwischenkommentar* (rmZwK.taex) wird zwischen dem zweiten und dem dritten Teilartikel eingeschoben. Dann liegt ein *teilartikelextern rechts-zentralerweiterter zweifach komplexer Wörterbuchartikel* vor.
- (3) Zwischen dem ersten und dem zweiten Teilartikel wird ein *teilartikelexterner linker mittlerer Zwischenkommentar* eingeschoben und zwischen dem zweiten und dem dritten Teilartikel wird ein *teilartikelexterner rechter mittlerer Zwischenkommentar* eingeschoben. Dann liegt ein *teilartikelextern links- und rechts-zentralerweiterter zweifach komplexer Wörterbuchartikel* vor, der auch *teilartikelextern vollständig zentralerweiterter zweifach komplexer Wörterbuchartikel* heißt. Bei dreifach komplexen Wörterbuchartikeln wird ein erster und ein zweiter mittlerer Zwischenkommentar unterschieden.

Es sei ausdrücklich daran erinnert, dass *teilartikelexterne mittlere Zwischenkommentare* unmittelbare Textkonstituenten des Wörterbuchartikels sind. Während bei der Erweiterung durch *teilartikelexterne vordere linke Zwischenkommentare* die Teilartikeldachstruktur nicht verändert wird (vgl. Abb. 7), ist dies bei der *teilartikelexternen Erweiterung* durch *mittlere Zwischenkommentare* anders. Die Teilartikeldachstruktur wird in dem speziellen Sinn diskontinuierlich, dass auf ein dachstrukturzugehöriges Element im dachstrukturexternes Element folgt (vgl. Abb. 10).

Für den Fall, dass die *Reihenfolge* der Teilartikel so ist wie in wa_{14} in Abb. 2, nämlich so: TA-Sub < TA-V < TA-Adj (mit „<“ für *geht voraus*) und alle drei Teilartikel basale Teilartikel sind (was in wa_{14} nicht der Fall ist), so dass einfache Teilartikelmikrostrukturen gegeben sind, hat das komplexe reduzierte allgemeine Mikrostrukturbild die Form, die sich in Abb. 10 findet, und das einfache Artikelstrukturschema hat die Form: WA : FK.taex < TA-Sub < lmZwK.taex < TA-V < rmZwK.taex < TA-Adj.

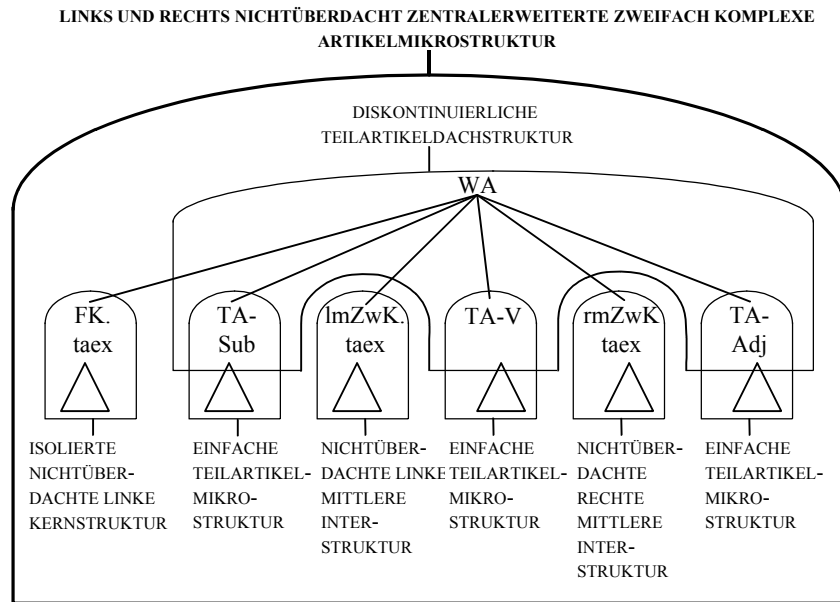


Abb. 10: Komplexes reduziertes allgemeines Mikrostrukturbild für links und rechts nichtüberdacht zentralerweiterte zweifach komplexe Artikelmikrostrukturen, die bei teilartikelextern links und rechts zentralerweiterten zweifach komplexen Wörterbuchartikeln auftreten, die zum Typ des Substantiv-Verb-Adjektiv-Artikel gehören

Abkürzungen: lmZwK.taex = teilartikelexterner linker mittlerer Zwischenkommentar; rmZwK.taex = teilartikelexterner rechter mittlerer Zwischenkommentar

Die Frage, welche lexikographische Daten in einem linken oder rechten mittleren Zwischenkommentar präsentiert werden, bezieht sich nicht auf die Wörterbuchform. Denkbar sind artikelzugehörige Abbildungen oder bei kultursensitiven Lemmazeichen enzyklopädische Angabetexte zum kulturellen Hintergrund.

Da aufgespaltene Teilartikel nicht auftreten, treten auch keine teilartikelexternen rechten Zwischenkommentare auf wie bei den nichtkomplexen Wörterbuchartikeln. Teilartikelexterne Postkommentare treten dagegen häufig auf.

In wa₁₇ findet sich ein teilartikelexterner Postkommentar zur Etymologie, der mit „-ORIGIN“ beginnt. Das einfache Artikelstrukturschema von wa₁₇ hat demgemäß die folgende Form: WA : FK.taex < TA-Sub < TA-V < T-Adj < PostK.Ety.taex

Im Folgenden seien einige weitere Wörterbuchartikel gegeben.

- wa₁₉: **brief** ■ **adj.** **1** of short duration. **2** concise; using few words. **3** (of clothing) not covering much of the body. ■ **n.** **1** chiefly Brit. a summary of the facts and legal points in a case given to an advocate to argue in court. ▶ a piece of work for an advocate. ▶ US a written summary of a case, for presentation to a court. **2** Brit. informal an attorney or advocate. **3** a set of instructions given to a person about a task. ■ **v.** instruct or inform thoroughly in preparation for a task. ▶ instruct (an advocate) by brief.
-PHRASES **hold a brief for** Brit. be retained as counsel for. **hold no brief for** not support. **in brief** in short.
-DERIVATIVES **briefless** adj. (Brit. Law). **briefly** adv. **briefness** n.
-ORIGIN Middle English: from Old French *brief*, from Latin *brevis* 'short'.
- wa₂₀: **bridle** ■ **n.** **1** the headgear used to control a horse, consisting of buckled straps to which a bit and reins are attached. **2** Nautical a mooring cable. ■ **v.** **1** put a bridle on. **2** bring under control. **3** show resentment or anger.
-PHRASES **off** (or **on**) **the bridle** see **BIT**³.
-ORIGIN Old English, of Germanic origin.
- wa₂₁: **brim** ■ **n.** **1** the projecting edge around the bottom of a hat. **2** the lip of a cup, bowl, etc. ■ **v.** (**brimmed**, **brimming**) [often as adj. **brimming**] fill or be full to the point of overflowing.
-DERIVATIVES **brimful** adj. **brimless** adj. **-brimmed** adj.
-ORIGIN Middle English: perhaps rel. to German *Bräme* 'trimming'.
- wa₂₂: **grey** ^{ɔːrɪ} (especially BrE) (NAmE usually *gray*) /greɪ/
adj., noun, verb
■ **adj.** **1** ^{ɔːrɪ} having the colour of smoke or ASHES: *grey eyes/hair* ◊ *wisps of grey smoke* ◊ *a grey suit* **2** ^{ɔːrɪ} [not usually before noun] having grey hair: *He's gone very grey.* **3** ^{ɔːrɪ} (of the sky or weather) dull; full of clouds: *grey skies* ◊ *I hate these grey days.* **4** (of a person's skin colour) pale and dull, because they are ill/sick, tired or sad **5** without interest or variety; making you feel sad: *Life seems grey and pointless without him.* **6** (*disapproving*) not interesting or attractive: *The company was full of faceless grey men who all looked the same.* **7** [only before noun] connected with old people: *the grey vote* ◊ *grey power* ▶ *grey-ness* (especially BrE) (NAmE usually *gray-ness*) *noun* [U, sing.]
■ **noun** **1** ^{ɔːrɪ} [U, C] the colour of smoke or ASHES: *the dull grey of the sky* ◊ *dressed in grey* **2** [C] a grey or white horse: *She's riding the grey.*
■ **verb** [I] (of hair) to become grey: *His hair was greying at the sides.* ◊ *a tall woman with greying hair*
- wa₂₃: **fancy** ■ **v.** (**-ies**, **-ied**) **1** informal, chiefly Brit. feel a desire for. ▶ find sexually attractive. **2** regard as a likely winner. **3** imagine. ▶ used to express surprise: *fancy that!* ■ **adj.** (**-ier**, **-iest**) elaborate or highly decorated. ▶ chiefly N. Amer. (especially of foodstuffs) of high quality. ▶ (of an animal) bred to develop particular points of appearance. ■ **n.** (pl. **-ies**) **1** a superficial or transient feeling of liking or attraction. **2** the faculty of imagination. ▶ an unfounded or tentative belief or idea. **3** (also **fancy cake**) a small iced cake or biscuit.
-PHRASES **take** (or **catch**) **someone's fancy** appeal to someone. **take a fancy to** become fond of, especially without an obvious reason.
-DERIVATIVES **fanciable** adj. (informal). **fanciness** n.
-ORIGIN Middle English: contraction of **FANTASY**.

Abb. 11: Wörterbuchartikel wa₁₉ – wa₂₁ und wa₂₃ aus OSACD² 2010; wa₂₂ aus OALD⁸ 2010

Bei den Artikeln wa₁₉ – wa₂₃ kann die Frage auftreten, ob ein komplexer Wörterbuchartikel nur einen teilartikelexternen Postkommentar oder mehrere aufweisen kann. Diese Frage muss in Übereinstimmung mit den terminologischen Regelungen für nichtkomplexe Wörterbuchartikel beantwortet werden, weil es keinen Grund gibt, in diesem Fall besondere Regelungen für komplexe Wörterbuchartikel einzuführen. Für die Theorie der Wörterbuchform gilt daher durchgängig — und d.h. für nichtkomplexe und für komplexe Wörterbuchartikel — Folgendes: Es werden zwei Typen von Kommentaren unterschieden:

- (1) Solche Typen, die durch ihre artikelinterne Position und zusätzlich inhaltlich festgelegt sind, nämlich der Formkommentar und alle Typen von Zweitkommentaren, wie z.B. der semantische Kommentar und der Kommentar zur Form und Semantik.
- (2) Solche Typen, die nur durch ihre Artikelposition bestimmt sind, nämlich der Präkommentar, der linke, mittlere und rechte Zwischenkommentar, der vordere linke, der hintere linke, der linke mittlere, der rechte mittlere, der erste mittlere und der zweite mittlere Zwischenkommentar sowie der Postkommentar. Die Kommentare dieses Typs können in verschiedenen Wörterbüchern Daten aufweisen, die zu unterschiedlichen Datentypen gehören. Sie können thematisch homogen sein; dann wird z.B. ein Postkommentar zur Phraseologie (PostK:Phras) oder ein Postkommentar zur Wortbildung (PostK:Wbdg) und weitere unterschieden. Die nur durch die Artikelposition bestimmten Kommentare können aber auch thematisch inhomogen sein. Dann weist ein Postkommentar entsprechende thematisch homogene Postsubkommentare auf. Das gilt dann auch für teilartikelexterne Postkommentare. In wa₁₉ finden sich drei, in wa₂₀ und wa₂₁ je zwei thematisch homogene Postsubkommentare. Alle Postkommentare in OSACD² 2010, die Postsubkommentare aufweisen, sind vertikalarchitektonisch ausgebaut; die zugehörigen komplexen Wörterbuchartikel gehören deswegen zu den partiell vertikalarchitektonisch ausgebauten Wörterbuchartikeln, und bei diesen zu den postkommentarintern vertikalarchitektonische ausgebauten Wörterbuchartikeln. Das komplexe allgemeine Mikrostruktur- und Mikroarchitekturbild für wa₁₉ in Abb. 11 findet sich in Abb. 12. Komplexe Wörterbuchartikel, die durch teilartikelexterne Postkommentare rechtserweitert sind, die Postsubkommentare aufweisen, gelten als komplex rechtserweitert.

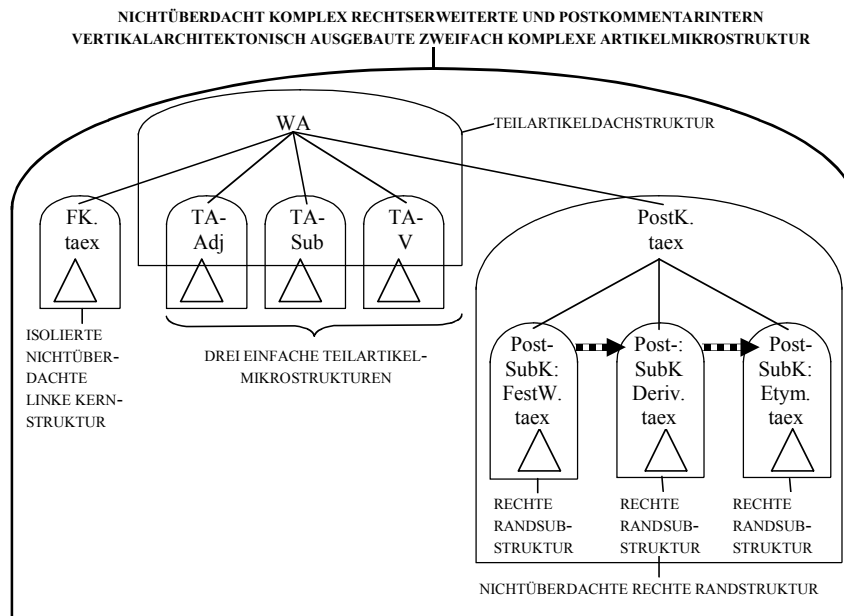


Abb. 12: Komplexes allgemeines Mikrostruktur- und Mikroarchitekturbild für den Typ der nichtüberdacht komplex rechtserweiterten und postkommentarintern vertikalarchitektonisch ausgebauten zweifach komplexen Artikelstruktur, die zum Typ des nichtüberdacht komplex rechtserweiterten und postkommentarintern vertikalarchitektonisch ausgebauten zweifach komplexen Wörterbuchartikels vom Typ des Adjektiv-Substantiv-Verb-Artikels gehört, wie wa₁₉ in Abb. 11

Es ist klar, dass die erläuterten verschiedenen teilartikelexternen textuellen Erweiterungen in allen möglichen Kombinationen zusammen auftreten können (was nicht bedeutet, dass sie in einem Wörterbuch auftreten müssen). Im Folgenden wird in der Form eines Typologieausschnittes eine Übersicht über alle sinnvollen Möglichkeiten gegeben. Wegen der verschiedenen Reihenfolgen der Teilartikel werden in den Artikelstrukturschemata die Variablen TA₁, TA₂ und TA₃ verwendet. Diese können dann durch die Klassensymbole TA-Sub, TA-V, TA-Adj und wenn nötig durch weitere ersetzt, so dass z.B. WA : FK.taex < TA₁ < TA₂ < TA₃ zu WA : FK.taex < TA-Sub < TA-Adj < TA-V wird.

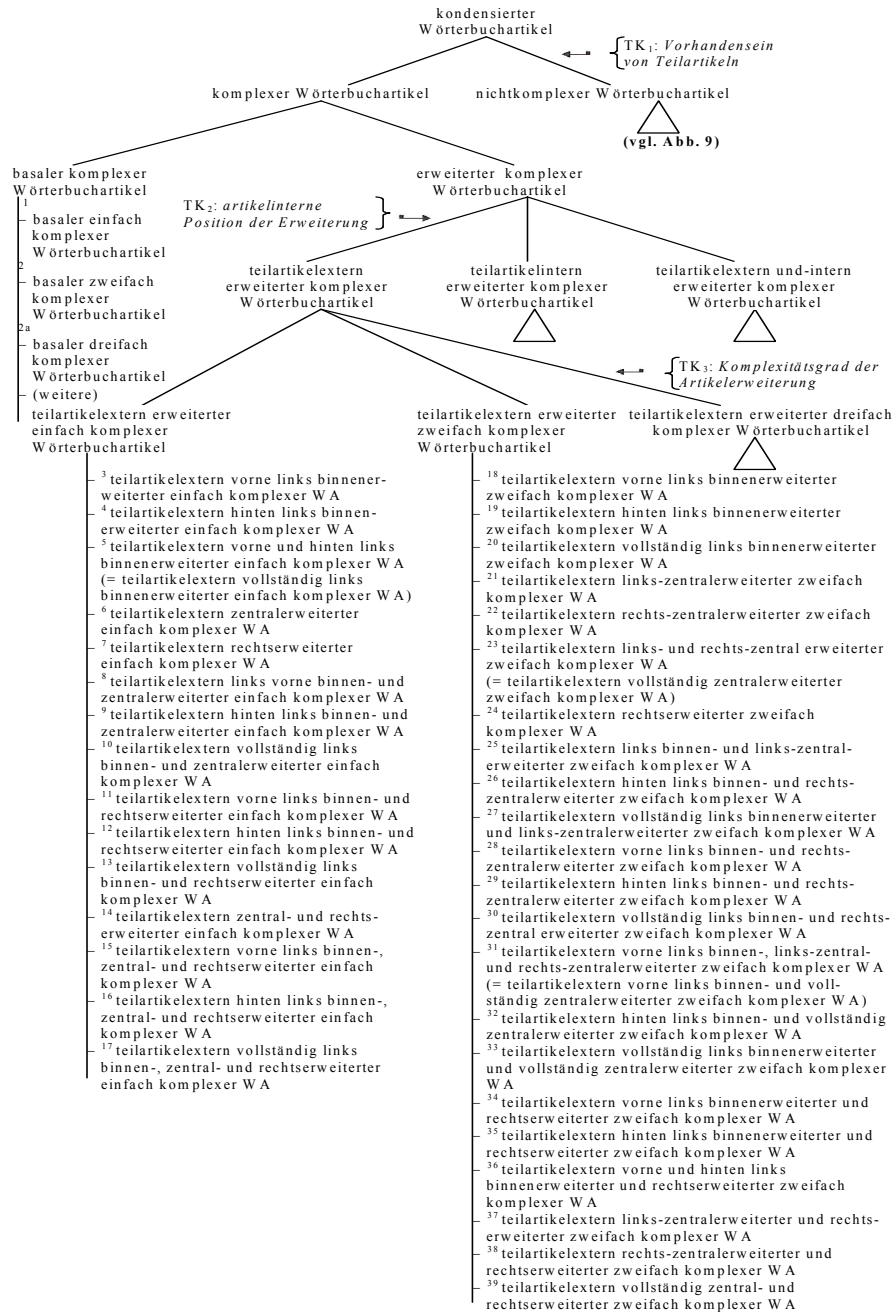


Abb. 13: Partiiell kommentierter Typologiegraph zu einem Ausschnitt einer Typologie von komplexen Wörterbuchartikeln

Im Folgenden werden die 39 einfachen Artikelstrukturschemata unter der Verwendung der Variablen „TA₁“, „TA₂“ und „TA₃“ für die Klassensymbole für Klassen von Teilartikeln wie TA-Sub, TA-Adj, TA-V und andere angegeben („<“ = *geht voraus*; „:“ = *besteht aus*).

I: Einfache Artikelstrukturschemata für basale komplexe Wörterbuchartikel und für teilartikelextern erweiterte einfach komplexe Wörterbuchartikel in Auswahl

- (1) WA : FK.taex < TA₁ < TA₂
- (2) WA : FK.taex < TA₁ < TA₂ < TA₃
- (3) WA : FK.taex < TA₁ < TA₂ < TA₃ < TA₄
- (4) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂
- (5) WA : FK.taex < hlZwK.taex < TA₁ < TA₂
- (6) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK.taex < TA₁ < TA₂
- (7) WA : FK.taex < TA₁ < mZwK.taex < TA₂
- (8) WA : FK.taex < TA₁ < TA₂ < PostK.taex
- (9) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < mZwK.taex < TA₂
- (10) WA : FK.taex < hlZwK < TA₁ < mZwK.taex < TA₂
- (11) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < mZwK.taex < TA₂
- (12) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < mZwK.taex < TA₂ < PostK.taex
- (13) WA : FK.taex < hlZwK.taex < TA₁ < TA₂ < PostK.taex
- (14) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK.taex < TA₁ < TA₂ < PostK.taex
- (15) WA : FK.taex < TA₁ < mZwK.taex < TA₂ < PostK.taex
- (16) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂ < TA₃
- (17) WA : FK.taex < hlZwK.taex < TA₁ < mZwK.taex < TA₂ < PostK.taex
- (18) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK.taex < TA₁ < TA₂

II: Einfache Artikelstrukturschemata für teilartikelextern erweiterte zweifach komplexe Wörterbuchartikel in Auswahl

- (19) WA : FK.taex < hlZwK.taex < TA₁ < TA₂ < TA₃

- (20) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂ < TA₃
- (21) WA : FK.taex < TA₁ < lmZwK.taex < TA₂ < TA₃
- (22) WA : FK.taex < TA₁ < TA₂ < rmZwK.taex < TA₃
- (23) WA : FK.taex < TA₁ < lmZwK.taex < TA₂ < rmZwK.taex < TA₃
- (24) WA : FK.taex < TA₁ < TA₂ < TA₃ < PostK.taex
- (25) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < lmZwK.taex < TA₂ < TA₃
- (26) WA : FK.taex < hlZwK.taex < TA₁ < lmZwK.taex < TA₂ < TA₃
- (27) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK < TA₁ < lmZwK.taex < TA₂ < TA₃
- (28) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂ < rmZwK.taex < TA₃
- (29) WA : FK.taex < hlZwK.taex < TA₁ < TA₂ < rmZwK.taex < TA₃
- (30) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂ < rmZwK.taex < TA₃
- (31) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < lmZw.taex < TA₂ < rmZwK.taex < TA₃
- (32) WA : FK.taex < hlZwK.taex < TA₁ < lmZwK.taex < TA₂ < rmZwK.taex < TA₃
- (33) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK.taex < TA₁ < lmZwK.taex < TA₂ < rmZwK.taex < TA₃
- (34) WA : vFK.taex < vlZwK.taex < hFK.taex < TA₁ < TA₂ < TA₃ < PostK.taex
- (35) WA : FK.taex < hlZwK.taex < TA₁ < TA₂ < TA₃ < PostK.taex
- (36) WA : vFK.taex < vlZwK.taex < hFK.taex < hlZwK.taex < TA₁ < TA₂ < TA₃ < PostK.taex
- (37) WA : FK.taex < TA₁ < lmZwK.taex < TA₂ < TA₃ < PostK.taex
- (38) WA : FK.taex < TA₁ < TA₂ < rmZwK.taex < TA₃ < Post.taex
- (39) WA : FK.taex < TA₁ < lmZwK.taex < TA₂ < rmZwK < TA₃ < PostK.taex

Abkürzungen: vlZwK = vorderer linker Zwischenkommentar ; hlZwK = hinterer linker Zwischenkommentar; mZwK = mittlerer Zwischenkommentar ; lmZwK = linker mittlerer Zwischenkommentar ; rmZwK = rechter mittlerer Zwischenkommentar ; PostK = Postkommentar; taex = teilartikelextern

Zu jedem der 39 einfachen Artikelstrukturschemata kann ein komplexes reduziertes allgemeines Mikrostrukturmodell angegeben werden, also ein solches, das ausschließlich die unmittelbaren Textkonstituenten des Wörterbuchartikels berücksichtigt. Das Darstellungsmittel des komplexen reduzierten allgemeinen

Mikrostrukturbildes ist gerade für Analysesituationen, wie die hier vorliegende, geschaffen, in denen man über die unmittelbaren Textkonstituenten der Zweitkommentare bzw. der Teilartikel nicht verfügt.

Um zu zeigen, welche Form die komplexen reduzierten allgemeinen Mikrostrukturbilder (mit Variablen) aufweisen, wird das Mikrostrukturbild für das einfache Artikelstrukturschema (31) in Abb. 14 präsentiert.

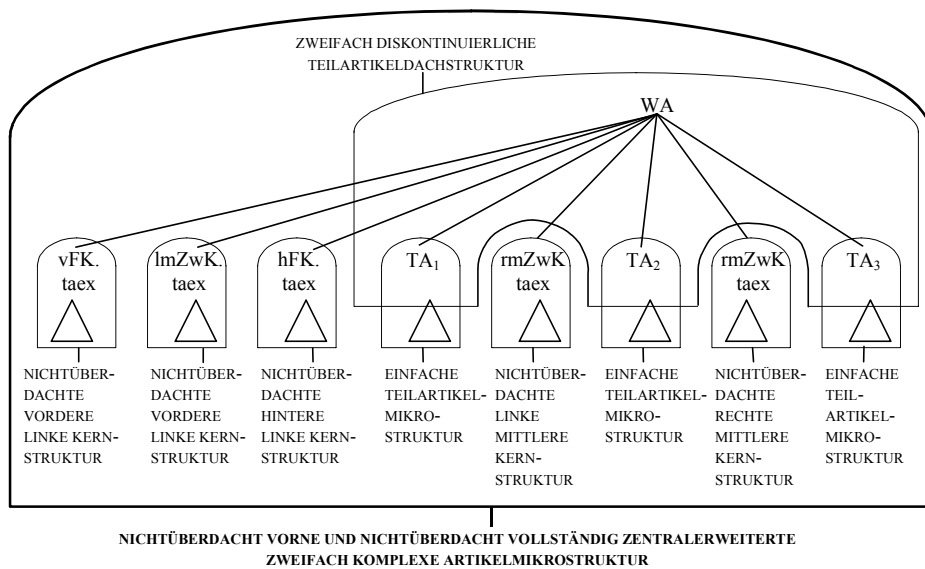


Abb. 14: Komplexes reduziertes allgemeines Mikrostrukturbild (mit Variablen TA₁, TA₂ und TA₃) für den Typ der nichtüberdacht vorne und nichtüberdacht vollständig zentralerweiterten zweifach komplexen Artikelmikrostruktur

Betrachtet man die Termini für die Artikeltypen in Abb. 13 genauer, dann bemerkt man, dass außer den vier oberbegrifflichen Termini *komplexer Wörterbuchartikel*, *basaler komplexer Wörterbuchartikel*, *erweiterter komplexer Wörterbuchartikel* und *teilartikelextern erweiterter komplexer Wörterbuchartikel* alle Termini entweder mit „*einfach komplexer Wörterbuchartikel*“ enden oder mit „*zweifach komplexer Wörterbuchartikel*“. Anhand dieser letzten Teile der Termini erfährt man, ob ein komplexer Wörterbuchartikel entweder zwei oder drei Teilartikel aufweist. Angenommen, es treten in einem Wörterbuch folgende Teilartikeltypen auf: TA-Sub, TA-V und TA-Adj und alle Reihenfolgen sind zugelassen, dann hat man bei den einfach komplexen Wörterbuchartikeln folgende sechs Möglichkeiten für die Reihenfolge der Teilartikel:

TA-Sub < TA-V ; TA-Sub < TA-Adj ; TA-V < TA-Sub ; TA-V < TA-Adj ;

TA-Adj < TA-Sub ; TA-Adj < TA-Verb.

Die Wörterbuchforschung arbeitet bereits z.B. mit folgenden Termini für nicht-komplexe Wörterbuchartikel: *Substantivartikel*, *Adjektivartikel*, *Verbartikel* und weiteren analog gebildeten (vgl. z.B. WLWF-I 2010). Nichts spricht dagegen, dass sie auch mit folgenden Termini für einfach komplexe Wörterbuchartikel arbeiten kann, anhand derer man die Reihenfolge der artikelinternen Teilartikel sofort erkennt, nämlich mit: *Substantiv-Verb-Artikel*, *Substantiv-Adjektiv-Artikel*; *Verb-Substantiv-Artikel*; *Verb-Adjektiv-Artikel*; *Adjektiv-Substantiv-Artikel*; *Adjektiv-Verb-Artikel*. Diese Termini kann man bei ihrer schriftlichen Verwendung auch kürzen, z.B. wie folgt: *Sub-V-Artikel*; *Sub-Adj-Artikel*; *V-Sub-Artikel*; *V-Adj-Artikel*; *Adj-Sub-Artikel*; *Adj-V-Artikel*. Anhand aller dieser Termini wird nicht nur die Teilartikelreihenfolge deutlich, sondern auch der Komplexitätsgrad, nämlich dass ein so bezeichneter Artikel einfach komplex ist. Man kann jetzt z.B. gut verständlich feststellen: wa₁₆ und wa₁₈ in Abb. 2 sind Substantiv-Verb-Artikel, wa₁₅ in Abb. 2 ist ein Adjektiv-Substantiv-Artikel oder wa₂₀ in Abb. 2 ist ein Substantiv-Verb-Artikel. Weiterhin kann man feststellen: wa₁₅ ist ein basaler Adjektiv-Substantiv-Artikel und wa₁₆ ist ein teilartikelextern erweiterter Substantiv-Verb-Artikel.

Schließlich kann man in allen Termini für teilartikelextern erweiterte Wörterbuchartikel, in denen *einfach komplexer Wörterbuchartikel* vorkommt, diesen Terminusteil jeweils durch einen der sechs Möglichkeiten ersetzen und erhält dann sechs verschiedene Termini: Dies sei an einem der Termini gezeigt, und zwar am Terminus *teilartikelextern vorne links binnenerweiterter Wörterbuchartikel*. Die Substitutionen führen zu folgenden sechs Termini:

- *teilartikelextern vorne links binnenerweiterter Substantiv-Verb-Artikel*
- *teilartikelextern vorne links binnenerweiterter Substantiv-Adjektiv-Artikel*
- *teilartikelextern vorne links binnenerweiterter Verb-Substantiv-Artikel*
- *teilartikelextern vorne links binnenerweiterter Verb-Adjektiv-Artikel*
- *teilartikelextern vorne links binnenerweiterter Adjektiv-Substantiv-Artikel*
- *teilartikelextern vorne links binnenerweiterter Adjektiv-Verbartikel.*

Zu diesen Artikeltypen gehören dann Definitionen, die nach dem Muster der Def. 6 formuliert sind.

Def. 6: Ein *teilartikelintern vorne links binnenerweiterter Substantiv-Verb-Artikel* liegt vor genau dann, wenn der teilartikelexterne Formkommentar durch einen vorderen linken Zwischenkommentar aufge-

spalten ist und wenn der erste Teilartikel ein Substantiv-Teilartikel ist, auf den ein Verb-Teilartikel unmittelbar folgt.

Insgesamt lassen auf diese Weise bei den teilartikelextern erweiterten einfach komplexen Wörterbuchartikeln zahlreiche Wörterbuchartikeltypen unterscheiden, und für jeden dieser Artikel gibt es einen systematisch gebildeten Terminus.

Man kann jetzt z.B. feststellen: wa₁₆ in Abb. 2 ist ein teilartikelextern vorne links binnenerweiterter basaler Substantiv-Verb-Artikel. Will man dessen architektonischen Ausbau berücksichtigen, dann kann man sagen: wa₁₆ ist ein hybrid vertikalarchitektonisch ausgebauter (und) teilartikelextern vorne links binnenerweiterter basaler Substantiv-Verb-Artikel (vgl. Abb. 7).

Bei den Termini für die teilartikelextern erweiterten zweifach komplexen Wörterbuchartikel kann analog verfahren werden. Der Artikel wa₁₃ in Abb. 2 ist dann z.B. ein basaler Substantiv-Verb-Adjektiv/Adverb-Artikel, und der Artikel wa₂₂ in Abb. 11 ist ein basaler Adjektiv-Substantiv-Verb-Artikel. Wenn man mit drei Teilartikeltypen rechnet, gibt es sechs verschiedene Reihenfolgen. Da in Abb. 13 25 Artikeltypen für teilartikelextern erweiterte zweifach komplexe Wörterbuchartikel vorgesehen sind, sind bei Berücksichtigung der Reihenfolge 150 teilartikelextern erweiterte zweifach komplexe Wörterbuchartikeltypen unterscheidbar und für jeden lässt sich ein eigener Terminus angeben. Nimmt man den Terminus *teilartikelextern rechtserweiterter zweifach komplexer Wörterbuchartikel* als Beispiel, dann lauten die Termini wie folgt:

- *teilartikelextern rechtserweiterter Substantiv-Verb-Adjektiv-Artikel*
- *teilartikelextern rechtserweiterter Substantiv-Adjektiv-Verb-Artikel*
- *teilartikelextern rechtserweiterter Verb-Substantiv-Adjektiv-Artikel*
- *teilartikelextern rechtserweiterter Verb-Adjektiv-Substantiv-Artikel*
- *teilartikelextern rechtserweiterter Adjektiv-Substantiv-Verb-Artikel*
- *teilartikelextern rechtserweiterter Adjektiv-Verb-Substantiv-Artikel.*

Der Artikel wa₂₃ in Abb. 11 gehört zum Typ des teilartikelextern rechtserweiterten Verb-Adjektiv-Substantiv-Artikels, und zwar zum Untertyp des teilartikelextern zweifach komplex rechtserweiterten V-Adj-Sub-Artikels.

Komplexe Wörterbuchartikel sind nicht nur teilartikelextern erweitert, sondern auch teilartikelintern. Gegeben seien die Artikel in Abb. 15.

wa₂₄: **bəˈrɡaɪn** BrE /'bɑːɡən; NAmE 'bɑːrɡən/ *noun, verb*
 ■ **noun** **1** ⇨ a thing bought for less than the usual price: *I picked up a few good bargains in the sale.* ◊ *The car was a bargain at that price.* ◊ *bargain prices* ⇨ **COLLOCATIONS** at **SHOPPING** **2** ~ (with sb) an agreement between two or more people or groups, to do sth for each other: *He and his partner had made a bargain to tell each other everything.* ◊ *I've done what I promised and I expect you to keep your side of the bargain* (= do what you agreed in return). ◊ *Finally the two sides struck a bargain* (= reached an agreement).
 ⇨ **go into the bargain** (BrE) (NAmE *in the bargain*) (used to emphasize an extra piece of information) also; as well: *Volunteers learn a lot and enjoy themselves into the bargain.* ⇨ more at **HARD** *adj.*, **STRIKE** *v.*
 ■ **verb** [I] to discuss prices, conditions, etc. with sb in order to reach an agreement that is acceptable **SYD** **negotiate**: ~ (with sb) (about/over/for sth) *In the market dealers were bargaining with growers over the price of coffee.* ◊ *He said he wasn't prepared to bargain.*
 ⇨ **bargain sth away** to give sth away and not get sth of equal value in return: *They felt that their leaders had bargained away their freedom.* **bargain for/on sth** (usually in negative sentences) to expect sth to happen and be prepared for it: *We hadn't bargained for this sudden change in the weather.* ◊ *When he agreed to answer a few questions, he got more than he bargained for* (= he got more questions, or more difficult ones, than he had expected). ◊ ~ **doing sth** *I didn't bargain on finding them here as well.* ◊ ~ **sb doing sth** *I hadn't bargained on them being here.*

wa₂₅: **ˈbɑːrəl** /'bærəl/ *noun, verb*
 ■ **noun** **1** a large round container, usually made of wood or metal, with flat ends and, usually, curved sides: *a beer/wine barrel **2** the contents of or the amount contained in a barrel; a unit of measurement in the oil industry equal to between 120 and 159 litres: *They got through two barrels of beer.* ◊ *Oil prices fell to \$9 a barrel.* **3** the part of a gun like a tube through which the bullets are fired
 ⇨ **a barrel of laughs** (often *ironic*) very amusing; a lot of fun: *Life hasn't exactly been a barrel of laughs lately.* (get/have sb) **over a barrel** (*informal*) (to put/have sb) in a situation in which they must accept or do what you want: *They've got us over a barrel. Either we agree to their terms or we lose the money.* ⇨ more at **LOCK** *n.*, **SCRAPE** *v.*, **SHOOT** *v.*
 ■ **verb** (-H, NAmE -!) [I] + **adv./prep.** (NAmE, *informal*) to move very fast in a particular direction, especially in a way that you cannot control: *He came barreling down the hill and smashed into a phone booth.**

wa₂₆: **riːˈmʊv** BrE /rɪ'muːv/ *verb, noun*
 ■ **verb** **1** ⇨ to take sth/sb away from a place: ~ **sth/sb** *illegally parked vehicles will be removed.* ◊ ~ **sth/sb from sth/sb** *He removed his hand from her shoulder.* ◊ *Three children were removed from the school for persistent bad behaviour.* **2** ⇨ ~ **sth** to take off clothing, etc. from the body: *She removed her glasses and rubbed her eyes.* **3** ⇨ to get rid of sth unpleasant, dirty, etc; to make sth disappear: ~ **sth** *She has had the tumour removed.* ◊ *to remove problems/obstacles/objections* ◊ ~ **sth from sb/sth** *The news removed any doubts about the company's future.* **4** ⇨ ~ **sb from sth** to dismiss sb from their position or job: *The elections removed the government from power.*
 ⇨ **once, twice, etc. re'moved** (of a cousin) belonging to a different generation: *He's my cousin's son so he's my first cousin once removed.* **be far/further/furthest removed from sth** to be very different from sth; to not be connected with sth: *Many of these books are far removed from the reality of the children's lives.*
 ■ **noun** [C, U] (*formal*) an amount by which two things are separated: *Charlotte seemed to be living at one remove from reality.*

- wa₂₇: **shrink** /ʃrɪŋk/ *verb, noun*
 ■ **verb** (shrank /ʃræŋk/, shrunk /ʃrʌŋk/) or (shrunk, shrunk) **1** [I, T] ~ (sth) to become smaller, especially when washed in water that is too hot; to make clothes, cloth, etc. smaller in this way: *My sweater shrank in the wash.* **2** [I, T] to become or to make sth smaller in size or amount: *The tumour had shrunk to the size of a pea.* ◊ *The market for their products is shrinking.* ◊ ~ **sth** *There was a movie called 'Honey, I Shrunk the Kids'.* ◊ *Television in a sense has shrunk the world.* ⇨ see also SHRUNKEN **3** [I] + **adv./prep.** to move back or away from sth because you are frightened or shocked **SYN** cower: *He shrank back against the wall as he heard them approaching.*
 [I] a **shrinking violet** (humorous) a way of describing a very shy person
 [I] **shrink from sth** to be unwilling to do sth that is difficult or unpleasant: *We made it clear to them that we would not shrink from confrontation.* ◊ ~ **doing sth** *They did not shrink from doing what was right.*
 ■ **noun** (slang, humorous) a PSYCHIATRIST OR PSYCHOLOGIST
- wa₂₈: **jack** /dʒæk/ *noun, verb, adj.*
 ■ **noun** **1** [C] a device for raising heavy objects off the ground, especially vehicles so that a wheel can be changed **2** [C] an electronic connection between two pieces of electrical equipment **3** [C] (in a PACK/DECK of cards) a card with a picture of a young man on it, worth more than a ten and less than a queen: *the jack of clubs* ⇨ VISUAL VOCAB page V32 **4** [C] (in the game of BOWLS) a small white ball towards which players roll larger balls **5** **jacks** [pl.] a children's game in which players BOUNCE a small ball and pick up small metal objects, also called jacks, before catching the ball **6** (also **jack 'shit, taboo**) [U] (NAME, slang) (usually used in negative sentences) anything or nothing at all: *You don't know jack.* ⇨ see also BLACKJACK, FLAPJACK, UNION JACK
 [I] a **jack of all trades** a person who can do many different types of work, but who perhaps does not do them very well ⇨ more at ALL RIGHT *adj.*, WORK *n.*
 ■ **verb** ~ **sth** | ~ **sb** (for sth) (NAME, informal) to steal sth from sb, especially sth small or of low value: *Someone jacked my seat.*
 [I] **jack sb a'round** (NAME, informal) to treat sb in a way that is deliberately not helpful to them or wastes their time: *Let's go. We're being jacked around here.* **jack 'in/ into sth** (informal) to connect to a computer system: *I'm jacking into the Internet now.* **jack sth** ⇨ in (BrE, informal) to decide to stop doing sth, especially your job: *After five years, he decided to jack it all in.* **jack 'off** (taboo, slang) (of a man) to MASTURBATE **jack 'up** (informal) to INJECT an illegal drug directly into your blood: *Drug users were jacking up in the stairwells.* **jack sth** ⇨ up **1** to lift sth, especially a vehicle, off the ground using a jack **2** (informal) to increase sth, especially prices, by a large amount
 ■ **adj.** [not before noun] ~ of sb/sth (AUSTRAL) tired of or bored with sb/sth

Abb. 15: Wörterbuchartikel wa₂₄ – wa₂₈ aus OALD⁸ 2010

Prinzipiell ist es möglich, dass alle Artikelerweiterungen, die ein basaler nicht-komplexer Wörterbuchartikel erfahren kann und die in Abb. 9 anhand der Termini zu erkennen sind, auch bei einem Teilartikel eines komplexen Wörterbuchartikels als teilartikelinterne Erweiterungen auftreten. Tatsächlich ist dies aber nicht der Fall. Vielmehr treten im OALD⁸ 2010 nur überdachte Postkommentare auf und damit nur teilartikelintern rechtserweiterte Teilartikel wie z.B.

in wa₂₄ – wa₂₈ in Abb. 15. Um möglichst solche Artikeltypbezeichnungen zu erhalten, anhand derer man erkennen kann, welcher Artikeltyp genau vorliegt, wenn die teilartikelinterne Erweiterung berücksichtigt wird, sei zunächst das Folgende erläutert: Bei einfach komplexen Artikeln, wie z.B. bei dem Substantiv-Verb-Artikel wa₂₄ in Abb. 15, wird auf den ersten Teilartikel mit *vorne* Bezug genommen und auf den zweiten mit *hinten*. Dann ergibt sich für den Wörterbuchtyp, zu dem der einfach komplexe Wörterbuchartikel wa₂₄ gehört, folgender Terminus: *teilartikelintern vorne und hinten rechtserweiterter Substantiv-Verb-Artikel* oder der synonyme Terminus *teilartikelintern vollständig rechtserweiterter Substantiv-Verb-Artikel*. Anhand beider Termini erfährt man, dass sowohl der Substantiv-Teilartikel als auch der Verb-Teilartikel intern rechtserweitert ist. Für wa₂₅ in Abb. 15 lautet der Terminus dann: *teilartikelintern vorne rechtserweiterter Substantiv-Verb-Artikel*. Anhand dieses Terminus erfährt man, dass nur Substantiv-Teilartikel intern rechtserweitert ist. Der Artikel wa₂₇ weist einen thematisch inhomogenen komplexen Postkommentar auf, der aus zwei thematisch homogenen Postsubkommentaren besteht; der erste ist ein Postsubkommentar zur Phraseologie, der zweite einer zu phrasal verbs. Komplexe Postkommentare dieser Art treten im OALD⁸ 2010 nur in Verb-Teilartikeln auf. Für wa₂₇ lautet der Terminus des zugehörigen Artikeltyps: *teilartikelintern vorne komplex rechtserweiterter Verb-Substantiv-Artikel*. Der Artikel wa₂₈ in Abb. 15 ist ein zweifach komplexer Wörterbuchartikel, und zwar ein Substantiv-Verb-Adjektiv-Artikel; der Substantiv-Teilartikel und der in der Mitte situierte Verbteilartikel sind beide rechtserweitert. Entsprechend lautet der Terminus für den zugehörigen Artikeltyp: *teilartikelintern vorne und zentral rechtserweiterter Substantiv-Verb-Adjektiv-Artikel*.

Die Definition des zweifach komplexen Wörterbuchartikels, der mit dem zuletzt genannten Terminus bezeichnet wird, lautet wie folgt:

Def. 7: Ein *teilartikelintern vorne und zentral rechtserweiterter Substantiv-Verb-Adjektiv-Artikel* liegt vor genau dann, wenn die beiden ersten Teilartikel, und zwar der zum Substantiv sowie der zum Verb, durch je einen teilartikelinternen Postkommentar rechtserweitert sind.

Der Artikel wa₂₆ in Abb. 15 ist sowohl teiltartikelextern als auch teilartikelintern erweitert. Die teiltartikelexterne Erweiterung entspricht der von wa₁₆ in Abb. 2. Demgemäß lautet der Terminus für wa₂₆ wie folgt: *teiltartikelextern vorne links binnenerweiterter und teilartikelintern vorne rechtserweiterter Verb-Substantiv-Artikel*.

Die Definition des einfach komplexen Wörterbuchartikels, der mit dem zuletzt genannten Terminus bezeichnet wird, lautet wie folgt:

Def. 8: Ein *teiltartikelextern vorne links binnenerweiterter und teilartikelintern vorne rechtserweiterter Verb-Substantiv-Artikel* liegt vor genau dann,

wenn ein einfach komplexer Wörterbuchartikel einen teilartikelexternen vorderen linken Zwischenkommentar aufweist und der erste Teilartikel, nämlich der zum Verb, durch einen teilartikelinternen Postkommentar rechtserweitert ist.

Nach diesen Erklärungen zur Bildung der Termini für teilartikelintern erweiterte einfach und zweifach komplexe Wörterbuchartikel sowie für teilartikelextern und teilartikelintern erweiterte Wörterbuchartikel ist es nicht nötig alle Termini, die für komplexe erweiterte Wörterbuchartikel auftreten können, aufzuzählen. Vielmehr können alle nicht genannten Termini nach den gleichen Regeln systematisch konstruiert werden.

Damit die textuellen Strukturen und Architekturen von teilartikelintern erweiterten und von sowohl teilartikelintern als auch teilartikelextern erweiterten komplexen Wörterbuchartikel anschaulicher werden, werden im Folgenden noch zwei komplexe allgemeine Struktur- und Architekturbilder präsentiert und mit kurzen Erläuterungen versehen. Zunächst wird in Abb. 16 das komplexe allgemeine Mikrostruktur- und Mikroarchitekturbild zu dem Artikeltyp präsentiert, zu dem wa₂₄ aus Abb. 15 gehört.

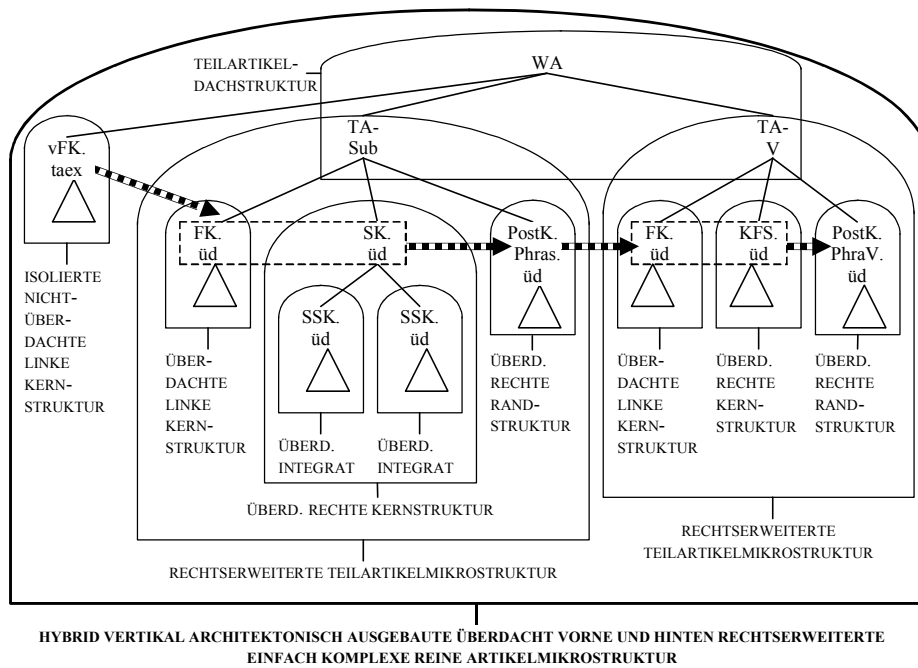


Abb. 16: Komplexes expandiertes allgemeines Mikrostruktur- und Mikroarchitektur-

bild für hybrid vertikalarchitektonisch ausgebaute überdacht vorne und hinten rechtserweiterte einfach komplexe reine Artikelmikrostrukturen, die Wörterbuchartikel vom Typ des teilartikelintern vorne und hinten rechtserweiterte Substantiv-Verb-Artikel aufweisen, wie wa₂₄ in Abb. 15

Abkürzungen: PostK:Phras.üd = überdachter Postkommentar zur Phraseologie; PostK:PhraV.üd = überdachter Postkommentar zu phrasal verbs.

Allgemeine Artikeltextstrukturbilder sowie allgemeine Artikeltextstruktur- und Artikelarchitekturbilder — daran sei hier erinnert — dienen dazu, einen raschen Überblick über die textuellen Strukturen und Architekturen von Wörterbuchartikeln zu gewinnen; dies gilt ebenso für komplexe allgemeine Artikeltextstruktur- und Artikelarchitekturbilder (vgl. Wiegand 2010c). In Abb. 16 ist die Artikelarchitektur von wa₂₄ vollständig abgebildet; wa₂₄ weist fünf Textblöcke auf: Den obersten, aus einer Artikelzeile bestehenden Textblock bildet der teilartikelexterne Formkommentar (FK.taex). Dieser steht oberhalb des Textblockes, den die Folge aus überdachtem Formkommentar und überdachtem semantischen Kommentar (FK.üd < SK.üd) bildet. Darunter folgt als nächster Textblock der überdachte Postkommentar zur Phraseologie (PostK:Phras.üd). Unter diesem folgt der erste Textblock des Verb-Teilartikels, bestehend aus der Folge von überdachtem Formkommentar und überdachtem Kommentar zur Form und Semantik (FK.üd < KFS.üd). Den letzten fünften Textblock in wa₂₄ bildet der überdachte Postkommentar zu phrasal verbs (PostK:PhrV.üd). Anhand des komplexen expandierten allgemeinen Mikrostruktur- und Mikroarchitekturbildes in Abb. 16 lässt sich folgendes einfache architektonisch angereichertes Artikelstrukturschema ablesen:

WA : FK.taex < /ob [FK.üd < SK.üd] < /ob PostK:Phras.üd < /ob [FK.üd < KFS.üd] < /ob Postk:PhraV.üd. (mit „:“ für *besteht aus*, „<“ für *geht voraus* und „< /ob“ für *geht voraus und ist oberhalb von*).

Im Folgenden betrachten wir anhand des komplexen expandierten allgemeinen Mikrostruktur- und Mikroarchitekturbildes den Artikel wa₂₆ in Abb. 15 als Beispiel für einen einfach komplexen Wörterbuchartikel, der sowohl teilartikelextern als auch teilartikelintern erweitert ist (vgl. Abb. 17).

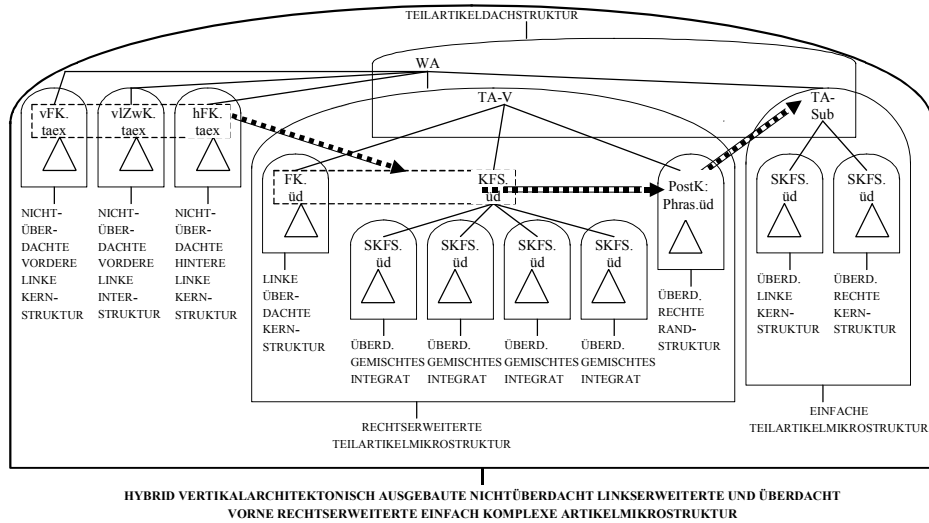


Abb. 17: Komplexes expandiertes allgemeines Mikrostruktur- und Mikroarchitekturbild für den Typ der hybrid vertikalarchitektonisch ausgebauten nichtüberdacht links erweiterten und überdacht vorne rechtserweiterten einfach komplexen reinen Artikelmikrostruktur, die — wie wa₂₆ in Abb. 15 — zum Typ des teilartikelextern vorne links binnererweiterten und teilartikelintern vorne rechtserweiterten Verb-Substantiv-Artikel gehören; Darstellungskonventionen: „x \dashrightarrow y“ bedeutet soviel wie x ist oberhalb von y

Auch in Abb. 17 ist der vertikalarchitektonische Artikelausbau, soweit er die Textblöcke betrifft, vollständig abgebildet. Der Artikel wa₂₆ weist vier übereinanderliegende Textblöcke auf; demgemäß gibt es für die Repräsentation der direkten oberhalb-Beziehung drei Pfeile der Form „ \dashrightarrow “. Nicht abgebildet wird demgemäß u.a., dass die Folge von Kommentaren vFK.taex < vZwK.taex < hFK.taex auch oberhalb des genannten Verb-Teilartikels liegt, weil dies aus der Transitivität der oberhalb-Relation folgt.

Nach den exemplarischen Analysen der 16 komplexen Wörterbuchartikel wa₁₃ – wa₁₈ in Abb. 2, wa₁₉ – wa₂₃ in Abb. 11 und wa₂₄ – wa₂₈ in Abb. 15 seien die unterschiedlichen Artikeltypen, zu der die 16 komplexen Wörterbuchartikel gehören, zusammenfassend genannt; dabei wird der vertikalarchitektonische Ausbau nicht berücksichtigt. Hinzugefügt werden bei den basalen komplexen Wörterbuchartikeln die einfachen und bei den erweiterten komplexen Wörterbuchartikeln die erweiteren Artikelstrukturschemata.

Komplexe Wörterbuchartikel in Abb. 2:

- wa₁₃: Typ des basalen Substantiv-Verb-Adjektiv / Adverb-Artikels;
WA : FK.taex < TA-Sub < TA-V < TA-Adj/Adv
- wa₁₄: Typ des teilartikelintern vorne komplex rechtserweiterten und hinten rechtserweiterten Verb-Substantiv-Artikels;
WA : FK.taex < TA-V [FK.üd < KFS.üd < PostK.inh.üb] < TA-Sub [FK.üd < SK.üd]
- wa₁₅: Typ des basalen Adjektiv-Substantiv-Artikels
WA : FK.taex < TA-Adj < TA-Sub
- wa₁₆: Typ des teilartikelextern vorne links binnenerweiterten Substantiv-Verb-Artikels
WA : vFK.taex < vlZwK.taex < hFK.taex < TA-Sub [FK.üd < SK.üd] < TA-V [FK.üd < KFS.üd]
- wa₁₇: Typ des teilartikelextern rechtserweiterten Substantiv-Verb-Adjektiv-Artikels
WA : FK.taex < TA-Sub [FK.üd < SK.üd] < TA-V [FK.üd < SK.üd < TA-Adv [FK.üd < SK.üd] < PostK:Etym.taex
- wa₁₈: Typ des basalen Substantiv-Verb-Artikels
WA : FK.taex < TA-Sub < TA-V

Komplexe Wörterbuchartikel in Abb. 11:

- wa₁₉: Typ des teilartikelextern komplex rechtserweiterten Adjektiv-Substantiv-Verb-Artikels
WA : FK.taex < TA-Adj [FK.üd < SK.üd] < TA-Sub [FK.üd < SK.üd] < TA-V [FK.üd < SK.üd] < PostK.inh.üd
- wa₂₀: Typ des teilartikelextern komplex rechtserweiterten Substantiv-Verb-Artikels
WA : FK.taex < TA-Sub [FK.üd < SK.üd] < TA-V [FK.üd < SK.üd] < PostK.inh.üd
- wa₂₁: Typ des teilartikelextern komplex rechtserweiterten Substantiv-Verb-Artikels
WA : FK.taex < TA-Sub [FK.üd < SK.üd] < TA-V [FK.üd < KFS.üd] < PostK.inh.üd
- wa₂₂: Typ des basalen Adjektiv-Substantiv-Verb-Artikels
WA : FK.taex < TA-Adj < TA-Sub < TA-V
- wa₂₃: Typ des teilartikelextern komplex rechtserweiterten Verb-Adjektiv-Substantiv-Artikels
WA : FK.taex < TA-V [FK.üd < KFS.üd] < TA-Adj [FK.üd < KFS.üd] < TA-Sub [FK.üd < KFS.üd] < PostK.inh.üd

Komplexe Wörterbuchartikel in Abb. 15:

- wa₂₄: Typ des teilartikelintern vorne rechts- und hinten rechtserweiterten Substantiv-Verb-Artikels
WA : FK.taex < TA-Sub [FK.üd < KFS.üd < PostK:Phras.üd] < TA-V [FK.üd < KFS.üd < PostK:PhrV.üd]
- wa₂₅: Typ des teilartikelintern vorne rechtserweiterten Substantiv-Verb-Artikels
WA : FK.taex < TA-Sub [FK.üd < SK.üd < PostK:Phras.üd] < TA-V [FK.üd < KFS.üd]
- wa₂₆: Typ des teilartikelextern vorne links binnererweiterten und teilartikelintern vorne rechtserweiterten Verb-Substantiv-Artikels
WA : vFK.taex < vlZwK.taex < hFK.taex < TA-V [FK.üd < KFS.üd < PostK:Phras.üd] < TA-Sub [FK.üd < KFS.üd]
- wa₂₇: Typ des teilartikelintern vorne komplex rechtserweiterte Verb-Substantiv-Artikels
WA : FA.taex < TA-V [FK.üd < KFS.üd < PostK:komp.üd] < TA-Sub [FK.üd < SK.üd]
- wa₂₈: Typ des teilartikelintern vorne und zentral rechtserweiterten Substantiv-Verb-Adjektiv-Artikels
WA : FK.taex < TA-Sub [FK.üd < KFS.üd < PostK:Phras.üd] < TA-V [FK.üd < KFS.üd < PostK:PhrasV.üd] < TA-Adj [FK.üd < KFS.üd]

(*Abkürzungen*: PostK:komp.üd = überdachter komplexer Postkommentar; PostK:PhrasV = Postkommentar zu phrasal verbs; „<“ = geht voraus; „:“ = besteht aus)

Die 16 genannten Typen von komplexen Wörterbuchartikeln sind nur ein kleiner Ausschnitt aus der Menge der Wörterbuchtypen, die an Wörterbuchartikeln ausgeprägt sind, und wie im CALD³ 2008, CD 2009, COED 2008, OALD⁸ 2010, im OSACD² 2010 und im POED⁶ 2006 tatsächlich auftreten, und die in diesen Wörterbüchern realisierten Artikeltypen sind nur ein sehr kleiner Ausschnitt einer nach dem erläuterten System konstruierbaren vollständigen Typologie für Typen von komplexen Wörterbuchartikeln.

Im Folgenden wird zur besseren Übersicht über den genauen typologischen Ort der 16 komplexen Wörterbuchartikel in Abb. 18 ein Ausschnitt aus einer Typologie für komplexe Wörterbuchartikel präsentiert, der genau diejenigen Typen von komplexen Wörterbuchartikeln aufweist, zu denen die 16 Artikel wa₁₃ – wa₂₈ gehören. Der Typologiegraph wird partiell extensionalisiert: Die 16 komplexen Wörterbuchartikel werden mittels der β -Relation, also der Typzugehörigkeitsrelation, zu der der Relationsterm $R_{\beta}^T = x$ gehört zum Artikeltyp y gehört, ihren jeweiligen Artikeltypen zugeordnet (vgl. Wiegand 2010c: 317); dabei ist „ x “ eine Variable für Wörterbuchartikel und „ y “ eine Variable für Artikeltypen.

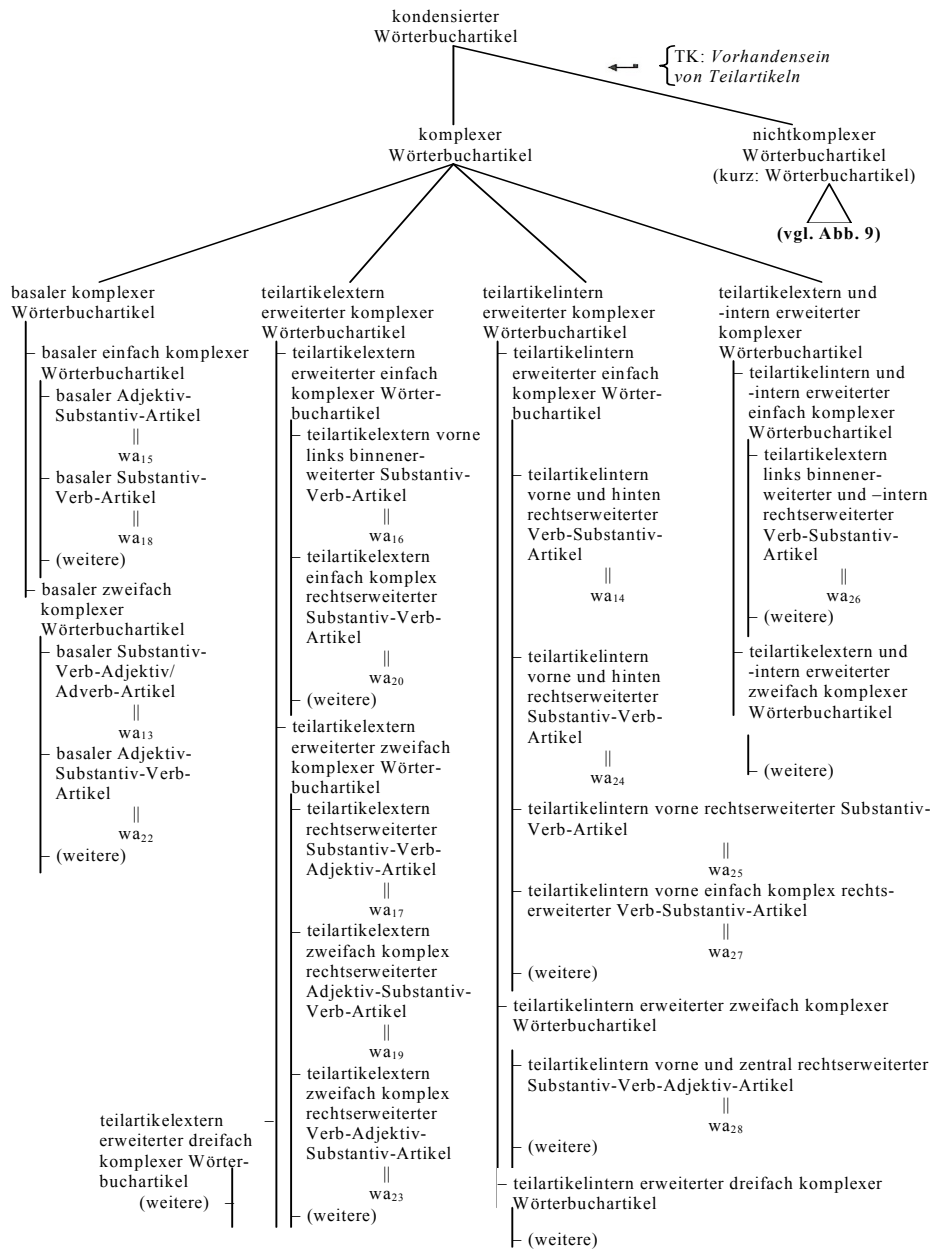


Abb. 18: Partiiell kommentierter und extensionalisierter Typologiegraph zu einem Ausschnitt aus einer Typologie von komplexen Wörterbuchartikeln; Darstellungskonvention: „x || y“ bedeutet soviel wie x gehört zum Artikeltyp y und repräsentiert genau eine Typzugehörigkeitsbeziehung

Damit sind die theoriebedingten Wörterbuchformprobleme, die durch die Berücksichtigung der komplexen Wörterbuchartikel entstanden waren, gelöst, und zwar *erstens* durch Erweiterung einer bereits (in Wiegand 2003: 259ff) entworfenen Typologie für nichtkomplexe kondensierte Wörterbuchartikel mit dem obersten Typologiekriterium *Art, Anzahl, und Reihenfolge der Kommentare* und *zweitens* durch eine dazugehörige systematisch aufgebaute Terminologie sowie *drittens* durch die systematische Ergänzung der Darstellungsmittel der Artikelstrukturschemata und der allgemeinen Artikelstruktur- und Artikelarchitekturbilder.

2.2 Artikelformprobleme für die Theorie der Wörterbuchform II: Fensterartikel und Artikelfolgen mit Fenster

Fensterartikel sind Wörterbuchartikel, zu denen als Teil des Textblockes eines Wörterbuchartikels eine rechteckige Fläche gehört, die farbig (im OALD⁸ 2010 hellblau) unterlegt ist und in der thematisch zusammengehörige Daten präsentiert werden, die in einer Beziehung zum Lemmazeichen stehen. Die rechteckige Fläche nennen wir *Artikelfenster* (kurz: *Fenster*) und sprechen demgemäß von *Fensterartikeln*. Diese wurden bisher in der Theorie der Wörterbuchform nicht berücksichtigt. Im OALD⁸ 2010 werden in Fenstern Wortfamilien zusammengestellt (vgl. Abb. 19). Daher sprechen wir von *Wortfamilienfenstern*, so dass es die Möglichkeit gibt, terminologisch verschiedene Fenstertypen hinsichtlich ihrer Thematik zu unterscheiden. Was man unter einer Wortfamilie versteht, wird nirgends erklärt; ein Lemma **word family** fehlt. Es wird mit einem sehr eng gefassten Wortfamilienbegriff gearbeitet. Die Selektion ist willkürlich: Warum es in manchen Wörterbuchartikeln Wortfamilienfenster gibt und in anderen Artikeln, deren Lemmazeichen ebenfalls zu einer Wortfamilie gehört, Wortfamilienfenster fehlen, bleibt ein Rätsel. Eine äußere Zugriffsstruktur für die Wortfamilienfenster fehlt, so dass die fensterinternen Daten nichtakzessive Textdaten sind. Dies hat als ein schwerer lexikographischer Fehler zu gelten, woran man erneut erkennen kann, dass die Oxford-Lexikographen zu wenig von der Wörterbuchform verstehen. Nach dem im OALD⁸ 2010 gepflegten Wörterbuchstil, dessen Nachahmung — wegen der damit z.T. verbundenen ungeschickten Datendistribution — niemanden empfohlen werden kann, hätte der Benutzer wenigstens erwarten dürfen, dass in der so genannten „Reference Section“ im Nachspann eine alphabetische Liste mit erwähnten Lemmata aller Fensterartikel mit Wortfamilien zu finden ist. Diese hätte er allerdings auch nur durch Herumsuchen und nicht durch systematisches Zugreifen finden können. Denn das Inhaltsverzeichnis des OALD⁸ 2010 ist eine Karikatur eines lexikographischen Inhaltsverzeichnisses, das — um funktionsfähig zu sein — bestimmte Eigenschaften aufweisen muss, die die Oxford-Lexikographen offenbar nicht kennen (vgl. Wiegand 2009b u. Teil II). Gegeben seien nun drei Fensterartikel mit Wortfamilienfenster in Abb. 19.

- wa29: **sus·pect** 0-^W *verb, noun, adj.* **WORD FAMILY**
^W **verb** /sə'spekt/ (not used in the progressive tenses)
1 ^W [I, I] to have an idea that sth is probably true or likely to happen, especially sth bad, but without having definite proof: ~ (sth) *If you suspect a gas leak, do not strike a match or even turn on an electric light.* ◊ *Suspecting nothing, he walked right into the trap.* ◊ *As I had suspected all along, he was not a real policeman.* ◊ ~ (that)... *I began to suspect (that) they were trying to get rid of me.* ◊ *It is suspected that... It was suspected that the drugs had been brought into the country by boat.* ◊ ~ **sb/sth to be/have sth** *She suspected him to be an impostor.* **2** ^W [I] to have an idea that sb is guilty of sth, without definite proof: ~ **sb/sth of sth** *He resigned after being suspected of theft.* ◊ ~ **sb/sth of doing sth** *The drug is suspected of causing over 200 deaths.* ◊ ~ **sb/sth** *Whom do the police suspect?* **3** [I] ~ **sth** to be suspicious about sth; to not trust sth: *I suspected her motives in offering to help.* ▶ **sus·pect·ed** *adj.*: a suspected broken arm ◊ suspected tax evasion ◊ suspected terrorists
^W **noun** ^W /'sʌspekt/ a person who is suspected of a crime or of having done sth wrong: a murder suspect ◊ He is the **prime suspect** in the case.
^W **adj.** /'sʌspekt/ **1** that may be false and that cannot be relied on **SYN** **questionable**: *Some of the evidence they produced was highly suspect.* **2** that you suspect to be dangerous or illegal **SYN** **suspicious**: a suspect package (= one that may contain drugs, a bomb, etc.)
- wa30: **able** 0-^W /'eɪbl/ *adj.* **WORD FAMILY**
1 ^W ~ **to do sth** (used as a modal verb) to have the skill, intelligence, opportunity, etc. needed to do sth: *You must be able to speak French for this job.* ◊ A viral illness left her barely able to walk. ◊ *I didn't feel able to disagree with him.* ◊ *Will you be able to come?* **SYN** **unable** ↗ note at **CAN** **2** (**able**: /'eɪbl(r)/, **ablest**: /'eɪblɪst/) intelligent; good at sth: *the ablest student in the class* ◊ *We aim to help the less able in society to lead an independent life.* ↗ see also **ABLY**
able *adj.* (≠ unable)
ably *adv.*
ability *noun* (≠ inability)
disabled *adj.*
disability *noun*
- wa31: **sat·is·fac·tion** 0-^W /sə'tɪs'fækʃn/ *noun*
1 ^W [U, C] the good feeling that you have when you have achieved sth or when sth that you wanted to happen does happen; sth that gives you this feeling; to **gain/get/derive satisfaction** from sth ◊ a look/smile of satisfaction ◊ *She looked back on her career with great satisfaction.* ◊ *He had the satisfaction of seeing his book become a best-seller.* ◊ *She didn't want to give him the satisfaction of seeing her cry.* ◊ *The company is trying to improve customer satisfaction.* ◊ *He was enjoying all the satisfactions of being a parent.* ↗ see also **DISSATISFACTION** **2** [U] the act of **FULFILLING** a need or desire: *the satisfaction of sexual desires* ◊ *the satisfaction of your ambitions* **3** [U] (*formal*) an acceptable way of dealing with a complaint, a debt, an injury, etc.: *I complained to the manager but I didn't get any satisfaction.* **WORD FAMILY**
satisfaction *noun* (≠ dissatisfaction)
satisfactory *adj.* (≠ unsatisfactory)
satisfy *verb*
satisfying *adj.* (≠ unsatisfying)
satisfied *adj.* (≠ dissatisfied) (≠ unsatisfied)
SYN to sb's **satisfaction** **1** if you do sth to sb's **satisfaction**, they are pleased with it: *The affair was settled to the complete satisfaction of the client.* **2** if you prove sth to sb's **satisfaction**, they believe or accept it: *Can you demonstrate to our satisfaction that your story is true?*

Der Artikel *wa*₂₉ ist ein zweifach komplexer Wörterbuchartikel vom Typ des basalen Verb-Substantiv-Adjektiv-Artikels mit Wortfamilienfenster. Auch komplexe Wörterbuchartikel anderer Typen treten mit Wortfamilienfenster auf. Beispielsweise ist der Artikel zum Lemmazeichen *broad* ein teilartikelintern vorne rechtserweiterter Adjektiv-Substantiv-Artikel mit Wortfamilienfenster, der Artikel zu *care* ist ein artikelintern vorne rechts- und hinten komplex rechtserweiterter Substantiv-Verb-Artikel, und der Artikel zu *pity* ein artikelintern vorne rechtserweiterter Substantiv-Verb-Artikel mit Wortfamilienfenster. Der Artikel zu *clear* ist ein diskontinuierlicher (dazu vgl. Teil II) artikelintern vorne rechtserweiterter, zentral komplex rechtserweiterter und hinten komplex rechtserweiterter Adjektiv-Verb-Adverb-Artikel mit Wortfamilienfenster. Der Artikel zu *deep* ist ein artikelintern vorne rechtserweiterter und zentral rechtserweiterter Adjektiv-Adverb-Substantiv-Artikel mit Wortfamilienfenster. Weitere komplexe Fensterartikel mit Wortfamilienfenster sind die Artikel zu den Lemmazeichen: *comfort, fat, force, long, produce, repeat, seperate, stable, value* und *wide*. Insgesamt gibt es 16 komplexe Fensterartikel mit Wortfamilienfenster im OALD⁸ 2010.

Nicht nur komplexe Wörterbuchartikel sind Fensterartikel mit Wortfamilienfenster, sondern auch nichtkomplexe wie *wa*₃₀ und *wa*₃₁ in Abb. 19 und weiterhin die Artikel zu folgenden Lemmazeichen: *deceive, defy, deny, destroy, discreet, divide, explain, explode, grateful, happy, intend, marry, nature, perceive, pronounce, rely, satisfaction, seperate, speak, strong, understand, vary* und *wit*. Insgesamt gibt es im OALD⁸ 2010 25 nichtkomplexe Fensterartikel mit Wortfamilienfenster und damit insgesamt 41 Fensterartikel mit Wortfamilienfenster. An dieser sehr kleinen Zahl erkennt man, dass lexikographisches Facelifting wenig mit seriöser und verantwortungsbewusster Sprachlexikographie zu tun hat. Denn es wurde eine vollständig willkürliche Auswahl aus der Menge der Wortfamilien getroffen.

Die Struktur von Fensterartikeln lässt sich mit den bisher in der Wörterbuchforschung entwickelten Untersuchungs- und Darstellungsmethoden nicht analysieren. Fenster sind keine „klassischen“ Textkonstituenten (vgl. aber unten), die mit den bisher eingeführten Strukturbildungsmethoden für hierarchische reine und hybride Textkonstituentenstrukturen systematisch strukturiert werden können (vgl. zu diesen Wiegand 2010c). Sie lassen sich auch nicht mit den bisher bekannten vertikalen und horizontalen Relationstypen als Teile einer Artikeltextarchitektur auffassen. Zwar befinden sich die Fenster rechts von Artikeltextteilen und auch oberhalb von Artikeltextteilen, Relationen vom Typ der *oberhalb*- und vom Typ der *rechts-von*-Relation werden aber auf Trägermengen definiert, deren Elemente Textkonstituenten sind und/oder geordnete Folgen von solchen und nicht auf Mengen von beliebigen Artikeltextteilen. Wenn man die Lage der Fenster im Artikel als einen Aspekt der Artikel- und damit der Wörterbuchform bestimmen möchte, muss man daher anders vorgehen. Dazu wird jetzt ein Vorschlag gemacht.

Zunächst kann man bei Wörterbüchern, die im Spaltensatz gedruckt sind, wegen der für diesen Satz geltenden Regeln, davon ausgehen, dass die Fenster in einem Wörterbuchartikel vom Setzer nicht so in der Spaltenmitte situiert werden, dass links und rechts des Fensters Artikeltextteile stehen. Daher kann man den Textblock, der insgesamt zu einem Wörterbuchartikel gehört, ohne dass dabei die gegebenenfalls vorhandenen artikelzugehörigen Teiltextblöcke (in wa₂₉ sind es z.B. vier) und damit der architektonische Ausbau des Wörterbuchartikels eine Rolle spielen, durch einen senkrechten Strich in eine linke und rechte Textblockhälfte teilen. Darauf zieht man zwei waagrechte Striche so, dass sechs gleichgroße Rechtecke gegeben sind und beschreibt ihre Lage im Textblock.

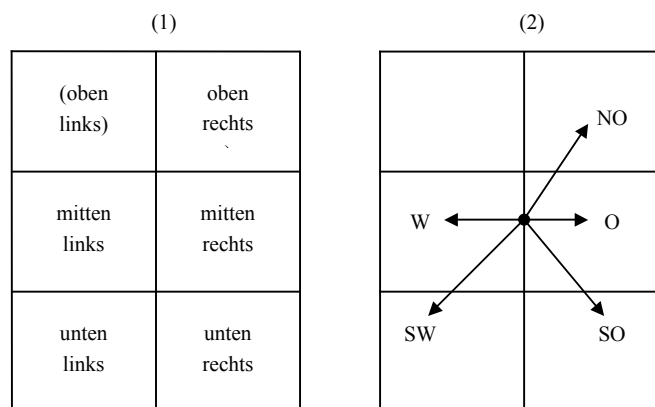


Abb. 20: Ordnungssysteme im zweidimensionalen Druckraum zur Bestimmung der artikelinternen Position von Artikelfenstern

Nach der Textblockeinteilung in Abb. 20 (1) kann man die Artikelfenster wie folgt charakterisieren:

- Ein Fenster liegt *oben rechts*, wenn kein artikelinterner Textteil über ihm und rechts von ihm steht.
- Ein Fenster liegt *mitte rechts*, wenn artikelinterne Textteile über und unter ihm, aber nicht rechts von ihm stehen.
- Ein Fenster liegt *unten rechts*, wenn kein artikelinterner Textteil unter ihm und rechts von ihm steht.
- Ein Fenster liegt *unten links*, wenn kein artikelinterner Textteil unter ihm und links von ihm steht.
- Ein Fenster liegt *mitte links*, wenn artikelinterne Textteile über und unter ihm, aber nicht links von ihm stehen.

Die oben-links-Position wird nicht bestimmt, da in dieser, wegen der dort gelegenen Lemmaposition, kein Artikelfenster auftreten kann.

Damit ist ein Ordnungssystem vorgeschlagen, nach dem man die relative artikelinterne Fensterposition bestimmen und nach der artikelinternen Lage fünf Fenstertypen unterscheiden kann; es ergeben sich dann z.B. die folgenden fünf Termini für Artikelfenster: *Oben rechts positioniertes, mitten rechts, unten rechts, unten links und mitten links positioniertes Artikelfenster* (kurz: *Fenster*).

Will man die Fensterartikel nach der Lage ihrer Fenster unterscheiden, ergeben sich Termini wie z.B. der Folgende für wa_{29} und wa_{30} : *Fensterartikel mit oben rechts positioniertem Fenster* und für wa_{31} *Fensterartikel mit mitten rechts positioniertem Fenster*. Will man zusätzlich das Fensterthema berücksichtigen (denn es sind ja verschiedene Fensterthemen denkbar), ergeben sich Termini wie z.B.: *Fensterartikel mit oben rechts positioniertem Wortfamilienfenster* (oder: z.B. *Synonymfenster*). Die genannten Termini sind zwar deskriptiv treffend, haben aber den Nachteil, dass sie ziemlich lang sind.

Aus diesem Grund und auch wegen ihrer Übersetzbarkeit in andere Sprachen bevorzugen wir das jetzt zu erläuternde System zur textblock- und artikelinternen Lagebestimmung der Artikelfenster sowie die dazu gehörenden Termini. Übernommen wird das System zur Bestimmung der geographischen Richtungen, wie es auch für Atlanten gilt und dessen Kenntnis bei den potenziellen Wörterbuchbenutzern vorausgesetzt werden kann. Das System wird nicht vollständig, sondern nur mit den Richtungen Nordosten (NO), Osten (O), Südosten (SO), Südwesten (SW) und Westen (W) übernommen. Nun kann man vom *Nordost-Artikelfenster* oder kurz z.B. vom *Nordost-Fenster* und noch kürzer z.B. vom *NO-Fenster* sprechen. Die in Abb. 10 (2) in der gleichen Position liegenden Fenster sind so festgelegt wie die in Abb. 10 (1); also ist z.B. das NO-Fenster so bestimmt wie das oben rechts positionierte Fenster. Die Artikel sind nun z.B. mit *Fensterartikel mit NO-Fenster* (kurz: *Artikel mit NO-Fenster*) eindeutig und treffend benannt. Will man das Artikelthema berücksichtigen, kann man den Terminus folgendermaßen bilden: *Fensterartikel mit NO-Fenster zur Wortfamilie*. Der Artikel wa_{29} gehört dann zum Typ des basalen Verb-Substantiv-Adjektiv-Artikels mit NO-Fenster zur Wortfamilie, wa_{30} zum Typ des (nichtkomplexen) Wörterbuchartikels mit NO-Fenster zur Wortfamilie und wa_{31} zum Typ des (nichtkomplexen) rechtserweiterten Wörterbuchartikels mit O-Fenster zur Wortfamilie.

Die Definition des Artikeltyps des basalen zweifach komplexen Wörterbuchartikels mit Wortfamilienfenster, zu dem wa_{29} in Abb. 19 gehört, lautet:

Def. 9: Ein *basaler Verb-Substantiv-Adjektiv-Fensterartikel mit NO-Fenster zur Wortfamilie* liegt vor genau dann, wenn in einem zweifach komplexen Wörterbuchartikel auf den ersten Teilartikel, den basalen Verb-Teilartikel, ein basaler Substantiv-Teilartikel und auf diesen ein basaler Adjektiv-Teilartikel unmittelbar folgt und wenn ein NO-Artikelfenster zur Wortfamilie gegeben ist.

Weiter oben wurde festgestellt, dass ein Fenster keine Textkonstituente im „klassischen“ Sinn ist. Zwar hat auch ein Fenster eine bestimmte Form und mindestens eine Funktion wie eine „klassische“ Textkonstituente, aber die Fensterposition ist keine Position in der sprachlichen Kette wie die Position einer „klassischen“ Textkonstituente, die damit immer zeilengebunden ist. Die artikelinterne Position eines Fensters ist dagegen zeilenungebunden. Man kann daher die Artikelfenster (wie übrigens auch artikelinterne Abbildungen) als *zeilenungebundene Textkonstituenten* auffassen, die nicht Teil der reinen oder hybriden Artikelkonstituentenstruktur sind, sondern eigenständige *zweidimensionale funktionale Artikelzusätze*, die, wenn sie wie Wortfamilienfenster im OALD⁸ 2010 aufgebaut sind, selbst eine bestimmte interne Struktur, und zwar die einer initialalphabetischen erweiterten Liste (i.S.v. Wiegand 1998: 198ff) aufweisen. Dass einem Artikelfenster zurecht der Status einer zeilenungebundenen Artikeltextkonstituente zukommt, ergibt sich daraus, dass seine Existenz einem bestimmten Wörterbuchartikeltyp konstituiert, nämlich den Fensterartikel.

Soll die Methode der Konstruktion von Artikelstrukturschemata (i.S.v. Wiegand 2010c: 313f) auch auf Fensterartikel anwendbar sein, muss das Symbolinventar dieser metalexikographischen Methode so ergänzt werden, das einfache erweiterte und architektonisch angereicherte sowie expandierte Artikelstrukturschemata auch für Fensterartikel angegeben werden können. Daher wird folgende Regelung eingeführt: Am Schluss jedes Artikelstrukturschemas, unabhängig davon, zu welchem Typ von Artikelstrukturschema es gehört, folgt auf ein Pluszeichen („+“) eines der folgenden fünf Symbole: NO-□, O-□, SO-□, SW-□, und W-□, die wie folgt zu lesen sind: Nordost-Artikelfenster (kurz: NO-Fenster) und sinngemäß so weiter.

Das einfache Fensterartikelstrukturschema für wa₂₉ in Abb. 19 lautet dann wie folgt:

$$WA - \square : FK.taex < TA-V < TA-Sub < T-Adj + NO-\square.$$

Das erweiterte Fensterartikelstrukturschema für wa₂₉ lautet:

$$WA - \square : FK.taex < TA-V [FK.üd < KFS.üd] < TA-Sub [FK.üd < SK.üd] < TA-Adj [FK.üd < SK.üd] + NO-\square.$$

Das einfache architektonisch angereicherte Fensterartikelstrukturschema für wa₂₉ hat die Form:

$$WA - \square : FK.taex </ob TA-V </ob TA-Sub </o TA-Adj + NO-\square.$$

Das erweiterte architektonisch angereicherte Fensterartikelstrukturschema für wa₂₉ lautet:

$$WA - \square : FK.taex </ob TA-V [FK.üd < KFS.üd] </ob TA-Sub [FK.üd < SK.üd] </o TA-Adj [FK.üd < SK.üd] + NO-\square \text{ (mit „</ob“ für geht voraus und ist oberhalb von).}$$

Das expandierte architektonisch angereicherte Fensterartikelstrukturschema für wa₂₉ lautet:

WA - □ : FK.taex </ob TA-V [FK.üd < KFS.üd [SKFS.üd < SKFS.üd < SKFS.üd]] </ob TA-Sub </ob TA-Adj [SSK.üd < SSK.üd] + NO-□.

Artikelfenster können bei der Darstellung von Textkonstituentenstrukturen von Fensterartikeln nicht berücksichtigt werden, da sie keine strukturelle Beziehung zu den Textkonstituenten aufweisen. Obwohl Artikelfenster keine Elemente der Artikelkonstituentenarchitektur sind und damit auch keine der Artikelmikroarchitektur, können sie in allgemeinen Artikelkonstituentenarchitekturbildern (vgl. WLWF-1 2010) und auch in allgemeinen Mikroarchitekturbildern berücksichtigt werden; bei diesen handelt es sich dann um entsprechend angereicherte Mikroarchitekturbilder. In Abb. 21 findet sich das allgemeine durch ein Artikelfenster angereicherte Mikroarchitekturbild von wa₂₉.

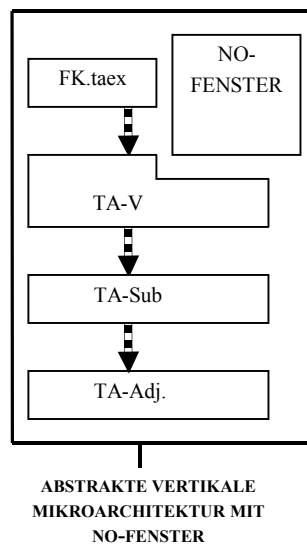


Abb. 21: Allgemeines durch ein NO-Artikelfenster angereichertes Mikroarchitekturbild für vertikal architektonisch ausgebaute zweifach komplexe Wörterbuchartikel wie wa₂₉ in Abb. 19; *Darstellungskonvention:* „**----->**“ bedeutet soviel wie *x ist oberhalb von y*

Das Artikelfenster ist in das allgemeine Mikroarchitekturbild so eingelagert, dass es rechts neben dem obersten Textblock, der vom teilartikelexternen Formkommentar gebildet wird, und rechts neben dem oberen Teil des Textblockes liegt, der vom Verb-Teilartikel gebildet wird, und damit analog zu der zweidimensionalen Situierung in wa₂₉

Damit sind die theoriebedingten Artikelformprobleme, die durch die Berücksichtigung von Fensterartikeln mit Wortfamilienfenster aufgetreten waren, nicht nur ad hoc gelöst, sondern so, dass alle in Zukunft auftretende Fensterartikel bereits bevor weiteres lexikographisches Facelifting dieser Art erfolgt, theoretisch beherrscht sind. Auch die Darstellungsmittel wurden angemessen erweitert.

Abschließend werden nun noch die Artikelfolgen mit Fenster behandelt, die sich im OALD⁸ 2010 finden und in Abb. 22 wiedergegeben sind.

<p>in·habit /ɪn'hæbɪt/ verb ~ sth (formal) to live in a particular place: <i>some of the rare species that inhabit the area</i></p> <p>in·hab·it·ant /ɪn'hæbɪtənt/ noun a person or an animal that lives in a particular place: <i>the oldest inhabitant of the village</i> ◊ <i>a town of 11 000 inhabitants</i></p>	<p>WORD FAMILY inhabit verb habitable adj. (≠ uninhabitable) inhabited adj. (≠ uninhabited) inhabitant noun habitation noun</p>
<p>accuse /ə'kju:z/ verb ~ sb (of sth) to say that sb has done sth wrong or is guilty of sth: <i>to accuse sb of murder/theft</i> ◊ <i>She accused him of lying.</i> ◊ <i>The government was accused of incompetence.</i> ◊ (formal) <i>They stand accused of crimes against humanity.</i> ▶ ac·cuser noun</p> <p>the ac·cused /ə'kju:zd/ noun (pl. the ac·cused) a person who is on trial for committing a crime: <i>The accused was found innocent.</i> ◊ <i>All the accused have pleaded guilty.</i> ↻ compare DEPENDANT</p>	<p>WORD FAMILY accuse verb accusation noun accusing adj. accusatory adj. accused noun</p>

Abb. 22: Wörterbuchartikel wa₃₂ u. wa₃₃ = Artikelfolge mit Fenster waf₁ und wa₃₄ u. wa₃₅ = Artikelfolge mit Fenster waf₂

Eine Artikelfolge mit Wortfamilienfenster besteht aus zwei unmittelbar aufeinander folgenden Wörterbuchartikel zu Lemmazeichen, die zur gleichen Wortfamilie gehören, und zwar Wörterbuchartikel, die für die Einpassung eines Fensters zu kurz sind, so dass das Fenster des vorangehenden Wörterbuchartikel in den folgenden Artikel hineinragt. In waf₂ verläuft der Spaltenumbruch durch den ersten Artikel. Damit er nicht auch durch das Fenster verläuft, steht das Artikelfenster nicht an der gleichen Stelle wie in waf₁.

Es ist klar, dass in beiden Fällen eine Art lexikographischer Notlösung vorliegt. Daher soll hier kein großer Beschreibungsaufwand betrieben werden. Es sollte aber klar sein, dass Fensterartikel von Artikelfolgen mit Fenster deutlich unterschieden werden müssen, weil zwei unterschiedliche Gegebenheiten der Wörterbuchform vorliegen.

Das Beispiel zeigt, dass offensichtlich ad-hoc geschaffene Elemente der Wörterbuchform, wie z.B. die Fensterartikel, zu weiteren ad-hoc-Lösungen führen. Die Wortfamilienfenster sind keine angemessene Lösung für das Vorhaben, in einem alphabetischen Wörterbuch Wortfamilien zu präsentieren, weil

das ohnehin nur dann geht, wenn man einen sehr engen (linguistisch altmodischen) Wortfamilienbegriff zugrunde legt und dazu einer skandalösen Selektionsphilosophie für die Wortfamilien folgt. Die englische Sprache hat solche Lexikographen nicht verdient.

2.3 Artikelformbedingte Benutzerprobleme

Während es im OALD⁸ 2010 und auch in zahlreichen anderen Lernerwörterbüchern, die bisher nur z.T. erwähnt wurden (vgl. Teil II), weil sie entweder keine komplexen Wörterbuchartikel oder keine besonderen Ausprägungen dieses Artikeltyps aufweisen, eine ganze Reihe von wörterbuchformbedingten Zugriffsproblemen für den Benutzer gibt sowie auch datendistributionsbedingte Benutzerprobleme, die im Teil II dieses Beitrages behandelt werden, treten artikelbedingte Benutzerprobleme für den kundigen Benutzer weder im OALD⁸ noch im OSACD² 2010 auf. Ein kundiger Benutzer des OALD⁸ ist dann allerdings nur einer, der nicht nur den Vorspanntext „Key to dictionary entries“ kennt, sondern auch die Seiten R 1 bis R 46 der Reference Section im Nachspann. Der sehr eigenwillige und z.T. benutzerunfreundliche Wörterbuchstil des OALD⁸ 2010, führt u.a. dazu, dass in der Reference Section z.B. die linguistische Erklärung, was ein transitives Verb ist, mit der Erklärung, wie es lexikographisch bearbeitet ist, Hand in Hand geht, und das gilt *mutatis mutandis* für alle lexikographisch bearbeiteten Eigenschaften der sprachlichen Ausdrücke. Ein deutscher Student z.B., der gewöhnt ist, dass die lexikographische Bearbeitung zusammenhängend in meistens vorspanninternen Benutzungshinweisen erklärt wird, wird daher anfänglich, wenn er beginnt, mit dem OALD⁸ 2010 zu arbeiten, durchaus artikelformbedingte Benutzerprobleme haben, z.B. wenn er fragt, ob *scowl* ein zählbares Substantiv ist, weil er dazu artikelintern keine Angabe der Zählbarkeit findet und nicht bemerkt hat, dass die bereits oben in 2.1 erwähnte Sonderregelung, die man natürlich angemessener als Nullangabe erklärt, in der Reference Section steht. Zwar kommt das Benutzerproblem des Studenten durch die Artikelform des Verb-Substantiv-Artikels zu *scowl* zustande, ist aber nicht von prinzipieller Natur, weil es bei Benutzern, die den Satz „Countable nouns are the most common type of noun. If they have only one meaning, or if all the meanings are countable, they are just marked noun“ (R 11) kennen, nicht auftritt. Solche Benutzerprobleme fallen nicht unter den in 1. charakterisierten Typ eines Benutzerproblems, weil sie nicht alleine durch die Wörterbuchform zustande kommen. Sie wären aber durch eine bessere Organisation der Metatexte vermeidbar, die im Teil II vorgeführt wird.

Im OALD⁸ 2010 gibt es häufiger sehr lange sowohl nichtkomplexe als auch komplexe Wörterbuchartikel. Es ist aus der Wörterbuchbenutzungsforschung bekannt, dass bei integrierten langen Wörterbuchartikeln mit vielen semantischen Subkommentaren und/oder Subkommentaren zur Form und Semantik die Auffindung der gesuchten Angaben mühsam ist (vgl. dazu u.a. Wiegand

1995). Hier haben die Oxford-Lexikographen eine *sehr gute benutzerfreundliche* Lösung zur Unterstützung des inneren Zugriffs gefunden (vgl. dazu Wiegand 2009b): Ab einer bestimmten Anzahl von Bedeutungen wird der semantische Kommentar oder der Kommentar zur Form und Semantik vertikalarchitektonisch ausgebaut: Jeder semantische Subkommentar (oder jeder Subkommentar zur Form und Semantik) ist dann als Textblock gestaltet, und der Polysemieangabe geht eine linksbündig gesetzte Bedeutungsauffindungsangabe voraus (ein so genannter *short cut*). Diese Regelung trägt dazu bei, dass die Benutzerprobleme, die durch die Artikellänge entstehen können, nicht auftreten können (vgl. Wiegand 2009b).

3. Schlussbemerkung

Die Frage, was für den Benutzer günstiger ist, ein komplexer Wörterbuchartikel, z.B. ein Substantiv-Verb-Adjektiv-Artikel, oder drei nichtkomplexe Wörterbuchartikel, z.B. ein Substantiv-, Adjektiv- und ein Verbartikel mit Wortklassenindices (wie z.B. in MEDAL² 2007, in LDCE⁵ 2009, im MWALED 2008 und im MWCD¹¹ 2003), ist nicht ohne aufwendige empirische Untersuchungen der Wörterbuchbenutzung zu entscheiden. Beide Wörterbuchstile lassen sich auch, sprachtheoretisch dann recht gut verteidigen, wenn man die entsprechenden Teile der Wörterbuchform angemessen bezeichnet und die Artikelstruktur deutlich erklärt. Dies ist im OALD⁸ 2010 allerdings nicht der Fall. Das zeigt u.a. der folgende Ausschnitt aus dem unbeholfenen Metatext „Key to dictionary entries“.

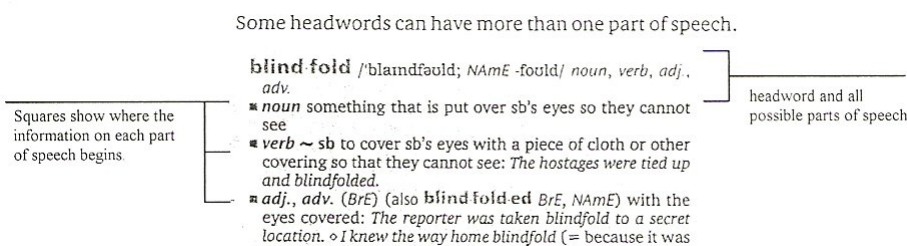


Abb. 23: Ausschnitt aus dem Vorspanntext „Key to dictionary entries“

Man muss es deutlich aussprechen: Der Satz „Some headwords can have more than one parts of speech“ ist ein Unsinnssatz. Ein headword ist eine Einheit eines Wörterbuchs. Diese kann weder zu einer Wortklasse noch zu mehreren gehören. Hier wird die Ebene des Wörterbuchs mit der des Wörterbuchgegenstandsbereichs verwechselt. Richtig könnte es heißen: „Haben zwei oder mehr Wörter die gleiche Form und gehören zu zwei oder zu mehreren Wortklassen, dann sind sie in *einem* Wörterbuchartikel bearbeitet. Mit dem Lemma (head-

word) wird dann die gleiche Form von zwei oder von mehreren Wörtern nur einmal genannt“.

Ausdrücklich sei Folgendes festgestellt: Aus dem vorausgehenden Abschnitt darf nicht geschlossen werden, wir wären der Meinung, dass in Sprachwörterbüchern vor allen Dingen linguistische Konzepte in den lexikographischen Texten einfließen sollten. Denn dies ist gerade nicht der Fall und ein Irrtum vieler Linguisten mit relativ schwerwiegenden Folgen für die benutzergerechte Gestaltung von Wörterbüchern. Es geht aber auch nicht an, dass in lexikographischen Metatexten linguistischer Unsinn verbreitet wird. Das schadet dem Ansehen der Lexikographie.

Anmerkungen

1. Eine Theorie der Wörterbuchform liegt seit 1989 mit der Grundlegung in Hausmann/Wiegand 1989 und in Wiegand 1989, 1989a und 1989b vor. Die Theorie wurde dann in mehreren Arbeiten, auch im Zusammenhang mit anderen Teilen einer Allgemeinen Theorie der Lexikographie ausgebaut. Größere und zentrale Arbeiten zur Theorie der Wörterbuchform sind u.a.: Bustos/Wiegand 2006; Gouws/Wiegand 2010, 2010a; Jesenšek/Wiegand 2009; Kammerer/Wiegand 1998; Meyer/Wiegand 2000; Wiegand 1990, 1991, 1992, 1994, 1995, 1996, 1996a, 1996b, 1996c, 1997, 1998, 1998a, 1998b, 1999, 2000, 2000a, 2000b, 2001, 2001a, 2002, 2002a, 2002b, 2002c, 2003, 2004, 2005, 2005a, 2005b, 2006, 2007, 2007a, 2007b, 2007c, 2007d, 2008, 2008a, 2009, 2009a, 2009b, 2010, 2010a, 2010b, 2010c; Wiegand/Fuentes Morán 2009; Wiegand/Smit 2010, 2010a.
2. Die Theorie der Wörterbuchform hat im Laufe der Jahre bereits mehrere Erweiterungen erfahren. Dies ist z.T. im Untertitel entsprechender Arbeiten erwähnt; vgl. z.B. Kammerer/Wiegand 1998; Wiegand 1991, 1997, 1998, 2005b.

Literatur

- Atkins, B.T. Sue und Michael Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: University Press.
- Bustos Plaza, Alberto und Herbert Ernst Wiegand. 2005/2006. Condensación textual lexicográfica: esbozo de una concepción integral. *Revista de Lexicografía* XII: 7-46.
- BW = *Brockhaus-Wahrig: Deutsches Wörterbuch in sechs Bänden*. 1980–1984. Hrsg. v. Gerhard Wahrig, Hildegard Krämer und Harald Zimmermann. 1. Bd.: A–BT 1980. 2. Bd.: BU–FZ 1981. 3. Bd.: G–JZ 1981. 4. Bd.: K–OZ 1982. 5. Bd.: P–STD 1983. 6. Bd.: STE–ZZ 1984. Wiesbaden: Brockhaus.
- CALD³ 2008 = *Cambridge Advanced Learner's Dictionary*. 2008³. Cambridge: Cambridge University Press.
- CD 2009 = *Collins English Dictionary*. Glasgow: HarperCollins.
- COBUILD 2001 = Sinclair, John (Hrsg.). 2001³. *Collins COBUILD English Dictionary for Advanced Learners*. Glasgow: HarperCollins.
- COBUILD 2009 = *Collins COBUILD Advanced Dictionary (of English)*. Glasgow: HarperCollins Publishers.

- COED 2009** = Soanes, Catherine und Angus Stevenson (Hrsg.). 2009¹¹. Neu bearb. *Concise Oxford English Dictionary*. Oxford: Oxford University Press.
- Duden-6 1990** = *Duden. Aussprachewörterbuch. Wörterbuch der deutschen Standardaussprache*. 3., völlig neu bearb. u. erw. Aufl. bearb. v. Max Mangold in Zusammenarbeit mit der Dudenredaktion. Mannheim [usw.]: Dudenverlag.
- Duden-3GW 1999** = Wissenschaftlicher Rat der Dudenredaktion. 1999. *Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden*. 3., völlig neu bearb. u. erw. Aufl. Bd. 1: *A-Bedi* und einem Geleitwort von Herbert Heckmann „Sprechen wir in der Zukunft noch Deutsch? [S. 9-13]; 2: *Bedr-Eink*; Bd. 3: *Einl-Geld*; Bd. 4: *Gele-Impr*; Bd. 5: *Impu-Leim*; Bd. 6: *Lein-Peko*; B. 7: *Pekt-Schi*; Bd. 8: *Schl-Tace*; Bd. 9: *Tach-Vida*; Bd. 10: *Vide-Zz*; mit Quellenverzeichnis zu den zitierten Belegen [S. 4725-4770]; einem Beitrag „Zur Neuregelung der deutschen Rechtschreibung“ v. Horst Sitta [S. 4748-4770]; einem Abriss „Die Geschichte der deutschen Sprache“ v. Jochen A. Bär [S. 4771-4782] und einer Auswahl „Wörter des Jahrhunderts“ aus „Der Sprachdienst“ [S. 4783-4900]. Mannheim [usw.]: Dudenverlag.
- Gouws, Rufus H. und Herbert Ernst Wiegand**. 2010. Macrostructures in Printed Dictionaries: An Overview. Typoscript. [Erscheint in HSK 5.4, 4. Teilband.]
- Gouws, Rufus H. und Herbert Ernst Wiegand**. 2010a. Addressing and Adressing Structures in Printed Dictionaries. Typoscript [Erscheint in HSK 5.4, 4. Teilband.]
- HAT 2009** = Odendal, F.F. und R.H. Gouws (Hrsg.). 2009⁵. *Handwoordeboek van die Afrikaanse Taal*. Kapstadt: Pearson Education South Africa.
- Hausmann, Franz Josef und Herbert Ernst Wiegand**. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. *HSK 5.1* [...] 1989: 328-360. [Wieder abgedruckt in: Hartmann, R.R.K. (Hrsg.). 2003. *Lexicography: Critical Concepts*. Vol. III: Lexicography, Meta-lexicography and Reference Science. London/New York 2003: 207-254].
- HSK 5.1** = Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie. Erster Teilband. 1989. Hrsg. v. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand und Ladislav Zgusta. Handbücher zur Sprach- und Kommunikationswissenschaft 5.1. Berlin/New York: De Gruyter.
- Jesenšek, Vida und Herbert Ernst Wiegand**. 2009. Zum Stand und den Perspektiven der allgemeinsprachlichen Lexikographie mit Deutsch und Slowenisch. *Lexikos* 19: 94-143.
- Kammerer, Matthias und Herbert Ernst Wiegand**. 1998. Über die textuelle Rahmenstruktur von Printwörterbüchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-238.
- LDCE⁵ 2009** = *Longman Dictionary of Contemporary English*. 2009⁵. Harlow: Longman Education.
- MEDAL² 2007** = **Rundell, M. (Hrsg.)**. 2007². *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan Education.
- Meyer, Meike und Herbert Ernst Wiegand**. 2000. Gemischt-semiintegrierte Mikrostrukturen für deutsch-spanische Printwörterbücher. Wiegand, Herbert Ernst (Hrsg.). 2000. *Studien zur zweisprachigen Lexikographie mit Deutsch V*: 87-171. Germanistische Linguistik 151-152. Hildesheim/New York: Georg Olms Verlag.
- MWALED 2008** = *Merriam-Webster's Advanced Learner's English Dictionary*. 2008. Springfield, Mass.: Merriam-Webster.
- MWCD¹¹ 2003** = *Merriam-Webster's Collegiate Dictionary*. 2003¹¹. Springfield, Mass.: Merriam-Webster.

- OALD**⁷ 2005 = *Oxford Advanced Learner's Dictionary of Current English*. (Edited by S. Wehmeier and A.S. Hornby.) Seventh Edition. Oxford: Oxford University Press.
- OALD**⁸ 2010 = *Oxford Advanced Learner's Dictionary of Current English*. (Edited by A.S. Hornby.) Eighth Edition. Oxford: Oxford University Press.
- OSACD**² 2010 = *Oxford South African Concise Dictionary*. (Edited by The Dictionary Unit for South African English.) Second Edition. Cape Town: Oxford University Press Southern Africa.
- POED**⁶ 2006 = Soanes, Catherine mit Sara Hawker und Julia Elliot. 2006⁶. *Paperback Oxford English Dictionary*. Oxford: Oxford University Press.
- Tarp, Sven**. 2010. Reflections on the Academic Status of Lexicography. *Lexikos* 20: 450-465.
- WDA** 1967 = *Wörterbuch der deutschen Aussprache*. [Hrsg. v. einem Kollektiv]. 2. Aufl. Leipzig: VEB Verlag Enzyklopädie.
- Wiegand, Herbert Ernst**. 1989. Aspekte der Makrostruktur im allgemeinen einsprachigen Wörterbuch: alphabetische Anordnungsformen und ihre Probleme. *HSK* 5.1 [...]: 371-409.
- Wiegand, Herbert Ernst**. 1989a. Der Begriff der Mikrostruktur: Geschichte, Probleme, Perspektiven. *HSK* 5.1: 409-462.
- Wiegand, Herbert Ernst**. 1989b. Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. *HSK* 5.1 [...]: 462-501.
- 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 [Auch in Wiegand, H.E. 2000. Bd.2: 951-1062].
- Wiegand, Herbert Ernst**. 1991. Über die Strukturen der Artikeltexte im Frühneuhochdeutschen Wörterbuch. Zugleich ein Versuch zur Weiterentwicklung einer Theorie lexikographischer Texte. Goebel, U. und O. Reichmann (Eds.) in collaboration with Barta, P.I. 1991. *Historical Lexicography of the German Language*. Vol. 2: 341-673. Studies in German Language and Literature 6. Lewiston/Queenston/Lampeter: Edwin Mellen Press.
- Wiegand, Herbert Ernst**. 1992. Elements of a Theory Towards a So-called Lexicographic Definition. *Lexicographica* 8: 175-289. [Auch in Wiegand 1999 [...]: 203-281].
- Wiegand, Herbert Ernst**. 1994. Deutsch-Turkmenisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Turkmenischen Staatlichen Magtymguly-Universität in Aschgabat. *Lexicographica* 10: 249-300.
- Wiegand, Herbert Ernst**. 1995. Lexikographische Texte in einsprachigen Lernerwörterbüchern. Kritische Überlegungen anlässlich des Erscheinens von Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Popp, H. (Hrsg.). *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag*: 463-499. München: Iudicium Verlag. [Auch in Wiegand, H.E. 2000. Bd. 2: 1128-1162].
- Wiegand, Herbert Ernst**. 1996. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. Wiegand, H.E. (Hrsg.). 1996. *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger lexikographischen Kolloquium*: 1-82. *Lexicographica*. Series Maior 70. Tübingen: Max Niemeyer Verlag.
- Wiegand, Herbert Ernst**. 1996a. Deutsch-Usbekisches Wörterbuch. Einblicke in die Wörterbucharbeit an der Staatlichen Usbekischen Weltsprachen-Universität in Taschkent. *Lexicographica* 12: 190-254.
- Wiegand, Herbert Ernst**. 1996b. Über die Mediostrukturen bei gedruckten Wörterbüchern. Zettersten, Arne und Viggo Hjørnager Pedersen, V. (Hrsg.). 1996. *Symposium on Lexicography VII*.

Proceedings of the Seventh Symposium on Lexicography May 5-6, 1994 at the University of Copenhagen: 11-43. *Lexicographica*. Series Maior 76. Tübingen: Max Niemeyer Verlag. [Auch in Wiegand, H.E. 2000. Bd. 2: 1163-1192].

- Wiegand, Herbert Ernst.** 1996c. A Theory of Lexicographic Texts. An Overview. *South African Journal of Linguistics* 14(4): 134-149.
- Wiegand, Herbert Ernst.** 1997. Printed Language Dictionaries and their Standardization: Notes on the Progress toward a General Theory of Lexicography. Hock, H.H. (Hrsg.). 1997. *Historical, Indo-European, and Lexicographical Studies. A Festschrift for Ladislav Zgusta on the Occasion of his 70th Birthday*: 319-380. *Trends in Linguistics. Studies and Monographs* 90. Berlin: Mouton de Gruyter: 319-380.
- Wiegand, Herbert Ernst.** 1998. Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband. Mit 159 Abbildungen im Text. Berlin/New York: Walter de Gruyter.
- Wiegand, Herbert Ernst.** 1998a. Altes und Neues zur Makrostruktur alphabetischer Printwörterbücher. Wiegand, H.E. (Hrsg.). *Wörterbücher in der Diskussion III. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 348-372. *Lexicographica*. Series Maior 84. Tübingen: Max Niemeyer.
- Wiegand, Herbert Ernst.** 1998b. Lexikographische Textverdichtung. Entwurf zu einer vollständigen Konzeption. Zettersten, Arne, Viggo Hjørnager Pedersen, V. und Jens Erik Mogensen (Hrsg.). *Symposium on Lexicography VIII. Proceedings on the Eighth Symposium on Lexicography, May 2-4, 1996 at the University of Copenhagen*: 1-35. *Lexicographica*. Series Maior 90. Tübingen: Max Niemeyer. [Auch in Wiegand, H.E. 2000, Bd. 2 [...], 1454-1489; spanische Übersetzung zusammen mit Alberto Bustos Plaza unter dem Titel: „Condensación textual lexicográfica: esbozo de una concepción integral. *Revista de Lexicografía* XII: 7-46].
- Wiegand, Herbert Ernst.** 1999. *Semantics and Lexicography. Selected Studies (1976-1996)*. Anntje Immen und Werner Wolski (Hrsg.). 1999. Supplementbände zum Internationalen Jahrbuch für Lexikographie 97. *Lexicographica*. Series Maior 97. Tübingen: Max Niemeyer.
- Wiegand, Herbert Ernst.** 1999a. Artikel einsprachiger Lernerwörterbücher, Textgestaltwahrnehmung und Suchbereichsstrukturen. Plädoyer für übersichtliche Printwörterbücher im Zeitalter der Neuen Medien: Skibitzki, Bernd und Barbara Wotjak (Hrsg.). 1999. *Linguistik und Deutsch als Fremdsprache. Festschrift für Gerhard Helbig zum 70. Geburtstag*: 259-281. Tübingen: Max Niemeyer.
- Wiegand, Herbert Ernst.** 2000. *Kleine Schriften. Eine Auswahl aus den Jahren 1970 bis 1999 in zwei Bänden. Band 1 1970-1988. Band 2 1988-1999*. Kammerer, Matthias und Werner Wolski (Hrsg.). Berlin/New York: Walter de Gruyter.
- Wiegand, Herbert Ernst.** 2000a. Über Suchbereiche, Suchzonen und ihre textuellen Strukturen in Printwörterbüchern. Ein Beitrag zur Theorie der Wörterbuchform. Wiegand, H.E. (Hrsg.). 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 233-301. *Lexicographica*. Series Maior 100. Tübingen: Max Niemeyer.
- Wiegand, Herbert Ernst.** 2000b. Adressierung in der ein- und zweisprachigen Lexikographie. Eine einführende Übersicht über die Forschungs- und Problemlage. *Lexikos* 10: 32-74.
- Wiegand, Herbert Ernst.** 2001. Über tabellarische Wörterverzeichnisse und deren Wörterbuchartikel. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica* 16: 212-234.
- Wiegand, Herbert Ernst.** 2001a. Sprachkontaktwörterbücher. Typen, Funktionen, Strukturen. Iglá, Birgit, Pavel Petkov und Herbert Ernst Wiegand (Hrsg.). 2001. *Theoretische und praktische Pro-*

bleme der Lexikographie. 1. Internationales Kolloquium zur Wörterbuchforschung am Institut Germanicum der St. Kliment Ohridski-Universität, Sofia, 7. bis 8. Juli 2000: 115-224. *Germanistische Linguistik* 161-162. Hildesheim/Zürich/New York: Georg Olms.

- Wiegand, Herbert Ernst.** 2001b: Was eigentlich sind Wörterbuchfunktionen? Kritische Anmerkungen zur neueren und neuesten Wörterbuchforschung. *Lexicographica* 17: 217-248.
- Wiegand, Herbert Ernst.** 2002. Über textuelle Strukturen der Wörterbuchartikel und Artikel-nischen im *de Gruyter Wörterbuch Deutsch als Fremdsprache*. Zugleich ein Beitrag zur Weiterentwicklung einer Theorie der Wörterbuchform. Wiegand, Herbert Ernst (Hrsg.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des de Gruyter Wörterbuchs Deutsch als Fremdsprache*: 497-595. *Lexicographica*. Series Maior 110. Tübingen: Max Niemeyer Verlag.
- Wiegand, Herbert Ernst.** 2002a. Zur Makrostruktur und zu den äußeren Zugriffsstrukturen im *de Gruyter Wörterbuch Deutsch als Fremdsprache*. Wiegand, Herbert Ernst (Hrsg.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des de Gruyter Wörterbuchs Deutsch als Fremdsprache*: 413-442. *Lexicographica*. Series Maior 110. Tübingen: Max Niemeyer Verlag.
- Wiegand, Herbert Ernst.** 2002b. Adressierung in zweisprachigen Printwörterbüchern. Wiegand, Herbert Ernst (Hrsg.). 2002. *Studien zur zweisprachigen Lexikographie mit Deutsch VIII*: 111-175. *Germanistische Linguistik* 166. Hildesheim/Zürich/New York: Georg Olms Verlag.
- Wiegand, Herbert Ernst.** 2002c. Altes und Neues zur Mediostruktur in Printwörterbüchern. *Lexicographica* 18: 168-252.
- Wiegand, Herbert Ernst.** 2003. Überlegungen zur Typologie von Wörterbuchartikeln in Printwörterbüchern. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica* 19: 169-313.
- Wiegand, Herbert Ernst.** 2004. Überlegungen zur Mediostruktur in Fachwörterbüchern. Auch am Beispiel des *Wörterbuchs zur Lexikographie und Wörterbuchforschung*. Brdar-Szabó, Rita und Elisabeth Knipf-Komlósi (Hrsg.). 2004. *Lexikalische Semantik, Phraseologie und Lexikographie. Abgründe und Brücken. Festgabe für Regina Hessky*: 339-365. *Duisburger Arbeiten zur Sprach- und Kulturwissenschaft* 57. Frankfurt a.M.: Peter Lang.
- Wiegand, Herbert Ernst.** 2005. Angaben, funktionale Angabezusätze, Angabetexte, Angabestrukturen, Strukturanzeiger, Kommentare und mehr. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica. International Annual for Lexicography* 21: 202-379.
- Wiegand, Herbert Ernst.** 2005a. Über die textuellen Strukturen im GWDS. Wiegand, Herbert Ernst (Hrsg.). 2005. *Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache II*. Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden. Print- und CD-ROM-Version. 2. Bd.: 295-335. *Lexicographica*. Series Maior 121. Tübingen: Max Niemeyer Verlag.
- Wiegand, Herbert Ernst.** 2005b. Über die Datenakzessivität in Printwörterbüchern. Einblicke in neuere Entwicklungen einer Theorie der Wörterbuchform. *Lexikos* 15: 196-230.
- Wiegand, Herbert Ernst.** 2006. Das Lern- und Konsultationswörterbuch. Ein neuer Fachwörterbuchtyp am Beispiel der *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK)*. *Lexikos* 16: 205-221.
- Wiegand, Herbert Ernst.** 2007. Aspectos escogidos de una teoría de las estructuras textuales de los diccionarios. Souto, Mar Campos, Eugenia Conde Noguero, José Ignacio Pérez Pascual und José Porto Dapena (Hrsg.). 2007. *Reflexiones sobre el diccionario*: 35-63. *Coruña Anexos de „Revista de Lexicografía“* 6. Coruña: Universidade da Coruña, Servizo de Publicacións.

- Wiegand, Herbert Ernst.** 2007a. Il riferimento indirizzato nei dizionari bilingui a stampa. *Studi Italiani di Linguistica Teorica e Applicata*: 7-87.
- Wiegand, Herbert Ernst.** 2007b. Sull'accessibilità ai dati nei dizionari a stampa. Sei considerazioni sugli ultimi sviluppi di una teoria sulla forma del dizionario. *Studi Italiani di Linguistica Teorica e Applicata*: 89-136.
- Wiegand, Herbert Ernst.** 2007c. Über Zugriffspfade in Printwörterbüchern. Ein Beitrag zur Schnittstelle von Benutzungshandlungen und Wörterbuchform. *Lexikos* 17: 180-211.
- Wiegand, Herbert Ernst.** 2007d. Neuere Aspekte einer Theorie und Typologie von Wörterbuchartikeln und ihre Praxisrelevanz. Gottlieb, Henrik und Jens Erik Mogensen (Hrsg.). 2007. *Dictionary Visions, Research and Practice. Selected Papers from the 12th International Symposium on Lexicography, Copenhagen 2004*: 183-200. Amsterdam/Philadelphia: John Benjamins.
- Wiegand, Herbert Ernst.** 2007e. Zur Adressierung in Printwörterbüchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 22: 187-261.
- Wiegand, Herbert Ernst.** 2008. Zugriffsstrukturen in Printwörterbüchern. Ein zusammenfassender Beitrag zu einem zentralen Ausschnitt einer Theorie der Wörterbuchform. *Lexicographica* 24: 209-315.
- Wiegand, Herbert Ernst.** 2008a. Wörterbuchregister. Grundlagen einer Theorie der Register in modernen Printwörterbüchern. *Lexikos* 18: 256-302.
- Wiegand, Herbert Ernst.** 2009. Nichtnatürlich über natürliche Sprache schreiben. Zu einigen formalen Aspekten von Wörterbuchartikeln. *Heidelberger Jahrbücher* 53.2009: 247-285.
- Wiegand, Herbert Ernst.** 2009a. Hybrid Text Constituent Structures of Dictionary Articles. A Contribution to the Expansion of the Theory of Textual Dictionary Structures. Nielsen, S. und S. Tarp (Hrsg.). 2009. *Lexicography in the 21st Century. In Honour of Henning Bergenholtz*: 23-41. Amsterdam/Philadelphia: John Benjamins.
- Wiegand, Herbert Ernst.** 2009b. Ausgewählte neuartige Komponenten der Wörterbuchform in deutschen und englischen einsprachigen Lernerwörterbüchern: Kritik und Vorschläge. *Lexicographica* 25: 169-211.
- Wiegand, Herbert Ernst.** 2010. Makro- und mikrostrukturelle Präsentationsprobleme bei Phrasemen in allgemeinen einsprachigen Wörterbüchern. Vorschläge für ihre Lösung. Durco, Peter (Hrsg.). 2010. *Feste Wortverbindungen und Lexikographie. Kolloquium zur Lexikographie und Wörterbuchforschung*: 161-180. *Lexicographica. Series Maior* 138. Berlin/New York: Walter de Gruyter.
- Wiegand, Herbert Ernst.** 2010a. Hybride textuelle Strukturen und hybride textuelle Einheiten. Ein Beitrag zur Theorie der Wörterbuchform. *Lexikos* 20: 495-539.
- Wiegand, Herbert Ernst.** 2010b. Semantik, Pragmatik und Wörterbuchform in einsprachigen Wörterbüchern. *Zeitschrift für germanistische Linguistik* 38: 405-441.
- Wiegand, Herbert Ernst.** 2010c. Zur Methodologie der Systematischen Wörterbuchforschung. Ausgewählte Untersuchungs- und Darstellungsmethoden für die Wörterbuchform. *Lexicographica* 26: 249-330.
- Wiegand, Herbert Ernst.** 2011. Komplexe Wörterbuchartikel. Ergänzende theoretische und methodologische Überlegungen. [Erscheint in *Lexicographica* 28. 2012].
- Wiegand, Herbert Ernst und Teresa Fuentes Morán.** 2009. Estructuras lexicográficas. Aspectos centrales de una teoría de la forma del diccionario. Granada: Ediciones Tragaçanto.
- Wiegand, Herbert Ernst und Ilse Feinauer.** 2010. Types of Dictionary Articles in Printed Dictionaries: An Overview. [Erscheint in HSK 5.4, 4. Teilbd.]

- Wiegand, Herbert Ernst und Maria Smit.** 2010. Microstructures in Printed Dictionaries. An Overview. [Erscheint in HSK 5.4, 4. Teilbd.].
- Wiegand, Herbert Ernst und Maria Smit.** 2010a. Mediostructures in Printed Dictionaries. An Overview. Typoscript. [Erscheint in HSK 5.4, 4. Teilbd.].
- WLWF-I 2010** = Wörterbuch zur Lexikographie und Wörterbuchforschung. Dictionary of Lexicography and Dictionary Research. Mit englischen Übersetzungen der Umtexte und Definitionen sowie Äquivalenten in neun Sprachen/With English Translations of the Outer Texts and Definitions as well as Equivalents in nine Languages. Hrsg. u. bearb. von Herbert Ernst Wiegand, Michael Beißwenger, Rufus H. Gouws, Matthias Kammerer, Angelika Storrer, Werner Wolski unter Mitarbeit von Ekaterina Budina-Koller (Russisch/Russian), Rute Costa (Portugiesisch/Portoguese), M^a Teresa Fuentes Morán (Spanisch/Spanish), Laura Giacomini (Italienisch/Italian), Rufus H. Gouws (Afrikaans; Englisch/English), Franz Josef Hausmann/Maria Hegner (Französisch/French), Regina Hessky/Zita Hollós (Ungarisch/Hungarian), Pavel Petkov (Bulgarisch/Bulgarian), Giovanni Rovere (Italienisch/Italian), Stefan J. Schierholz (Portugiesisch/Portuguese), Maria Smit (Englisch/English). I: Bd: Systematische Einführung/Vol. I: Systematic Introduction. A-C. Berlin/New York: Walter de Gruyter.

Categorising Example Sentences in Dictionaries for Research Purposes*

Lorna Hiles, Cape Town, South Africa (lorna@hiles.co.za)

Abstract: Examples in dictionaries come in many different forms. They may be sentences or phrases. They may be corpus-based or made up by the lexicographer. They may contain the lemma in its uninflected form or they may contain an inflection of the lemma. In some dictionaries the function of examples is to provide contextual support to the meaning of the headword, and in others the grammatical support that they provide is more important.

While there is literature on the usefulness of examples, and on whether examples should be corpus-based or not, there is very little on what makes one example more useful than another. I have set out to find out what sort of examples South African school users identify as most helpful. In this article, I look at whether examples in five South African school dictionaries do provide suitable contextual or grammatical support.

I have constructed a table to classify example sentences according to different criteria. I filled in this table with randomly selected words and their examples which have been taken from five different South African school dictionaries.

The goal of this research is to present characteristics of examples in a way that makes them easier to analyse and compare. This should help lexicographers in future dictionaries check whether they have written or selected the best possible examples for their users' needs.

Keywords: EXAMPLE SENTENCES, ILLUSTRATIVE EXAMPLES, SCHOOL DICTIONARIES, TABLE OF CATEGORIES

Opsomming: Kategorisering van voorbeeldsinne in woordeboeke vir navorsingsdoeleindes. Voorbeelde in woordeboeke verskyn in baie verskillende vorme. Hulle kan sinne of frases wees. Hulle kan korpusgebaseer of deur die leksikograaf geskep wees. Hulle kan die lemma in sy onverboë vorm bevat of hulle kan 'n verbuiging van die lemma bevat. In sommige woordeboeke is die funksie van voorbeelde om kontekstuele steun vir die betekenis van die trefwoord te verskaf, in ander is die grammatiese steun wat hulle verskaf, belangriker.

Terwyl daar literatuur oor die nuttigheid van voorbeelde bestaan, en of voorbeelde korpus-

* This article is a combination of two papers respectively presented at the Fourteenth International Conference of the African Association for Lexicography, organised by the Xhosa Department, University of the Western Cape, Bellville, South Africa, 6-8 July 2009 and the Fifteenth International Conference of the African Association for Lexicography, held at the University of Botswana, Gaborone, Botswana, 19-21 July 2010, both based on research for an M.Phil. Thesis *Examples in South African School Dictionaries: From Theory to Practice*, which was accepted by Stellenbosch University, Stellenbosch, South Africa, November 2009.

gebaseer moet wees of nie, bestaan daar baie min oor wat een voorbeeld nuttiger maak as 'n ander. Ek het begin om uit te vind watter soort voorbeelde Suid-Afrikaanse skoolgebruikers aanwys as die nuttigste. In hierdie artikel kyk ek of voorbeelde in vyf Suid-Afrikaanse skoolwoordeboeke wel geskikte kontekstuele of grammatiese steun verskaf.

Ek het 'n tabel opgestel om voorbeeldsinne volgens verskillende maatstawwe te klassifiseer. Ek het hierdie tabel voltooi met lukraak gekose woorde en hul voorbeelde wat uit vyf verskillende Suid-Afrikaanse skoolwoordeboeke geneem is.

Die doel van hierdie navorsing is om eienskappe van voorbeelde op 'n manier aan te bied wat dit makliker maak om hulle te ontleed en te vergelyk. Dit behoort leksikograwe in toekomstige woordeboeke te help om te kontroleer of hulle die bes moontlike voorbeelde geskryf of gekies het vir hul gebruikers se behoeftes.

Slutelwoorde: VOORBEELDSINNE, TOELIGTENDE VOORBEELDE, SKOOL-WOORDEBOEKE, TABEL VAN KATEGORIEË

While doing research into example sentences in school dictionaries I found a lot of theory about what examples are meant to achieve (proof of a word's use and existence in historical dictionaries, a complement to the definition, and an illustration of grammatical features of the word), but very little information on evaluating examples and what makes a good example.

I needed to analyse and evaluate examples in school dictionaries and to compare the examples in dictionaries with what the publisher set out to achieve with them, and with what learners wanted from example sentences in their dictionaries.

To this end, I created a table which I used to compare the features of examples from different dictionaries and extract related statistical data.

This made it easier to compare the dictionaries to each other and find patterns within the dictionaries. I was also able to compare what the dictionaries actually contain with what they were intended to contain, according to the introduction in the dictionary, the How to Use page, or discussions with the editor.

Before creating this table, I chose five South African school dictionaries: three monolingual and two bilingual.

The monolingual dictionaries are: South African Oxford School Dictionary (SAOSD) second edition, 2004, South African Oxford Secondary School Dictionary (SAOSSD) 2006, and Longman South African School Dictionary (Longman) 2007.

The bilingual dictionaries are *Oxford Afrikaans–Engels/English–Afrikaans Skoolwoordeboek/School Dictionary* (Tweetalige) 2006, and *Pharos Aanleerderswoordeboek vir Skole* (Pharos) third edition, 2006.

There are two Oxford school dictionaries because they are aimed at different users and they are quite different dictionaries, as the tables will show. The *South African Oxford School Dictionary* is for learners in Grades 4 to 10, while the *South African Oxford Secondary School Dictionary* is for learners in Grades 8 to 12.

So while there is an overlap of grades, the dictionaries are sufficiently different to warrant the inclusion of both.

The table contains fourteen columns of data, arranged in six sections. The first section is the example number, actual example sentence or phrase with the headword highlighted, and the part of speech. The second section indicates whether the example is a phrase or a sentence. The third section is for whether the example is a definition, whether it provides a context, or neither. The fourth section provides information about grammatical support: whether the example provides grammatical support, and if so, how. The fifth section shows whether the headword is simple or inflected in the example. The final section shows whether the example provides any other support, such as cultural information or usage restrictions not mentioned in the other columns.

		POS	Sentence	Phrase	Definition	Contextual support	No context	Grammatical support	Comments	Simple	Inflected	Additional support	Comments
	SAOSD												
1	a cake of soap	n		1			1	1	cake of	1			

Each column is added up, with totals provided for each dictionary, and for all the dictionaries together. These totals are used to generate statistics of the examples.

I selected 24 words randomly, keeping to the same words in each dictionary, where the dictionary contained the word as a lemma. I only chose nouns, verbs and adjectives, so that I could compare these part of speech categories. If a word was polysemous, I used all senses in that entry. For example, the verb **keep** has five senses in the SAOSD, four of which are exemplified. In the Longman dictionary there are seven senses of **keep**, six of which are exemplified. I used all of the senses in this study, rather than choosing one example from an entry, as in some cases the examples are meaningless on their own, but are useful when compared to other examples in that entry. I did not use examples from idioms or expressions. For example, in the Pharos dictionary, **keep** has eight senses followed by 17 fixed expressions. The expressions may contain valuable examples, but these are not used in this study. Where a lemma does not feature in the category table, it means either that the lemma is not treated in that dictionary or that it is not exemplified.

These 24 words generated a total of 146 examples across the five dictionaries. The results for all dictionaries are shown in the table below.

	POS	Sentence	Phrase	Definition	Contextual support	No context	Grammatical support	Comments	Simple	Inflected	Additional support	Comments
Total	146	114	32	3	98	45	62		103	43	60	
		78%	22%	2%	67%	31%	43%		71%	29%	41%	

The columns I was most interested in are the columns about the contents of the examples, and whether they assist the learner in any way (with contextual, grammatical or additional support).

The **Contextual support** section was the most difficult to keep consistent, since the allocation of values is determined by my intuition or rules. The basis for deciding whether an example contained contextual support was that if the headword could only be replaced by a synonym or near synonym in the sentence, or a word in the same lexical set, it was marked as providing contextual support. If the headword could be replaced by any other word of the same part of speech, then I marked it in the **No context** column. Other researchers could use their own policies for this. Consistency and being able to defend each decision are the keys to ensuring accurate data.

For example, in the hypothetical example *I love peaches* where **peaches** is the headword, I found there to be no clue, as the lemma could easily be replaced by any other noun: *I love soccer, I love animals*. In the sentence *I enjoy eating peaches* the options for replacing the headword are limited to the lexical set of food. In the sentence *Peaches are my favourite fruit* the options have been further narrowed down to the lexical set of fruit.

My main aim for looking for contextual support in the examples was if a learner does not understand the definition, will the example help? An example of how this has worked in this study is an example from SAOSD: *Defeats are offset by successes*. The definition (counterbalance or make up for something) contains difficult vocabulary and could be misinterpreted, but the example clearly shows what **offset** means and how it is used.

This support is especially valid in South African schools, where learners may not have the language skills and fluency for a monolingual dictionary aimed at first language English speakers. Their dictionaries need to give as much support as necessary.

As can be seen by the results table above, 67% of the examples (across the monolingual and bilingual dictionaries) examined do contain contextual support.

Grammatical support is another purpose of example sentences, specifically for the text production function of a dictionary. Most of the information in

the **Grammatical support** columns concerns collocations, and whether the example sentence provides suggestions for common collocations. For example, the sentence *The success of the product depends on marketing* shows that **depends** can collocate with **on**.

I have also looked at whether the example shows that verbs can be used transitively or intransitively. This is difficult to quantify because an example cannot show how a word cannot be used. I have indicated transitive/intransitive use where there are two or more examples for the same sense of a verb, and where the examples show different uses.

Out of 146 sentences examined 55 (38%) provide this type of grammatical support. It is difficult to draw conclusions and say whether this is significant, because many of the headwords do not need grammatical support of this kind.

The **Additional support** columns contain information that may be helpful to the learner, but cannot be described by the previous columns. These sentences may contain cultural data, which would include racial and gender information. This may be more useful in a bilingual dictionary where the cultures of the different languages are different. For example, in an English–Northern Sotho school dictionary, the lemma **stick** is exemplified by sentences about hitting children or animals, all from a Northern Sotho corpus. For example, *The teacher beats the children with a stick*. A reader is unlikely to find these sentences in a monolingual English school dictionary. This is considered cultural information.

Gender information would include whether gender roles are stereotyped or not, such as *His wife nagged him to paint the wall*. While it is difficult to avoid gender altogether, some dictionaries use strategies to avoid gender in example sentences, for example, by using "I", "we" or "you" as the subject of a sentence, instead of "he" or "she" or names. Another strategy is to use names that are not gender specific, such as Jo, Lee, or Terry. Gender information can also be neutral, such as *He painted a picture*.

Other information that fits in the **Additional support** section is whether the word is used figuratively or idiomatically. This is valuable support, especially for a second language user of a dictionary.

I have also included whether the sentence is a command as some teachers prefer school dictionaries to not be too instructive. If learners see their dictionary as a set of instructions, they may be discouraged from using it. The examples are supposed to get learners' attention and interest, without being didactic.

The information found in the **Additional support** section is also what makes an example, and therefore a dictionary, engaging to a learner. If a learner consulting a dictionary finds the examples interesting and varied, they are more likely to respond to the dictionary positively. In this study 41% of the examples contained additional support in some form. Again, these are difficult statistics to work with, since it is unlikely that every example can contain additional support. However, one must be aware when interpreting these results that support in the **Additional support** column may not be valuable as actual support in the dictionary.

I also compared the dictionaries according to various criteria, such as how many of the headwords are treated in each dictionary, how many senses are treated, and how many have examples. I also looked at the differences in nouns, verbs and adjectives, to see if there were patterns in their treatment.

An interesting result of these statistics is that in both of the bilingual dictionaries that I looked at, each entry contains example sentences, sometimes more than one, while the monolingual dictionaries did contain entries without examples. One can also see that while the monolingual dictionaries made use of example phrases in some instances, the bilingual dictionaries only used full sentences. These statistics also show that the monolingual dictionaries treated more words, with the SAOSSD and Longman containing the highest number of headwords treated for the 24 words.

The monolingual dictionaries and bilingual dictionaries yield different statistics, as shown in the table below. The bilingual dictionaries can be seen to provide more support across **Contextual** and **Additional support**.

	Sentence	Phrase	Definition	Contextual support	No context	Grammatical support	Comments	Simple	Inflected	Additional support	Comments
Monolingual	63%	37%	1%	54%	45%	44%		66%	34%	39%	
Bilingual	100%	0%	3%	86%	10%	41%		78%	22%	44%	

I analysed the category table further to find out how many examples provide support in more than one area (contextual, grammatical and additional). Pharos has the most with 21%, followed by Tweetalige (16%), SAOSD (14%), Longman (10%) and SAOSSD (4%). Again, one needs to be wary of using these statistics as some headwords may not require additional or grammatical support. However, it is interesting to note these numbers, as many examples could be changed to contain as much support as possible.

The Pharos dictionary performs well in all the categories showing that it provides the most support out of all the dictionaries. It is clear that the examples were chosen or written with this in mind. However, it must be noted that this is offset by the space that these supportive sentences take up. Without doing a word count comparison of the dictionaries, one can see that the Pharos examples are longer than those of the other dictionaries. Many of the Pharos examples contain two clauses or phrases, or even sentences, making them more complicated, and less accessible to learners of English (who do not speak Afrikaans).

The data supports the conclusion that phrases are less useful at providing support than sentences, and my recommendation is that unless there is a good reason for using a phrase in a particular instance, that a sentence is used instead. A phrase can be used with a sentence, but it is recommended that in school dictionaries, using phrases as the only way of illustrating a sense is avoided.

To conclude, I feel that this table and its corresponding statistics offer great value in analysing and comparing different dictionaries according to very specific criteria. For my research, in particular, I used it to determine whether learners' needs in example sentences were being met. I also used it to provide recommendations for future school dictionaries.

The table of categories can be used in different ways, depending on who is using it. One can populate the table with examples from one dictionary, and find areas where the dictionary is lacking sufficient support for the headwords. One can also use the table to make sure that the examples are balanced — for example, in terms of gender references or location, and that the examples are not weighted to a particular subject matter such as crime.

Another way this table can be used is to compare dictionaries according to different criteria — as I have used it, but with different information in the columns. One can determine whether there is a pattern in the structure of examples: how do they use tenses, are they written actively or passively, are the sentences simple or complex?

One can also use a similar table of categories for whole entries in a dictionary in order to find out what support each entry provides. Other forms of support include artwork, synonyms and antonyms, usage notes, etymology, cross references, and labels. These can be used in conjunction with word frequency data and a defining vocabulary to find out which words do need more support. Words that are frequent and in the defining vocabulary may not need as much support as more difficult, less common words.

Bibliography

- Bullon, S. et al. (Eds.).** 2007. *Longman South African School Dictionary*. First Edition. Harlow, Essex: Pearson Education Ltd.
- Du Plessis, M. (Ed.).** 2006. *Aanleederswoordeboek vir Skole/Learner's Dictionary for Schools*. Third Edition. Cape Town: Pharos.
- Justus, J.A. (Ed.).** 2004. *The South African Oxford School Dictionary*. Second Edition. Cape Town: Oxford University Press.
- Louw, P.A. et al. (Eds.).** 2006. *Oxford Afrikaans–Engels/English–Afrikaans Skoolwoordeboek/School Dictionary*. First Edition. Cape Town: Oxford University Press.
- Reynolds, M. et al. (Eds.).** 2006. *South African Oxford Secondary School Dictionary*. First Edition. Cape Town: Oxford University Press.

Problems of Usage Labelling in English Lexicography*

Lydia Namatende Sakwa, *Department of Languages, Faculty of Humanities
and Social Sciences, Uganda Martyrs University, Nkozi, Uganda*
(*lydiasakwa@yahoo.com*)

Abstract: Landau (1991: 217) stipulates that 'usage refers to any or all uses of language'. It is the study of good, correct, or standard uses of language as distinguished from bad, incorrect, and non-standard uses of language. Usage may also include the study of any limitations on the method of use, whether geographic, social or temporal. Basically it alerts users that certain terms should not be uncritically employed in communication. This article discusses the treatment of usage in English lexicography. It analyses the labelling practices in six monolingual English dictionaries namely: the *Oxford Advanced Learner's Dictionary* (OALD), the *Macmillan English Dictionary* (MED), the *Longman Dictionary of Contemporary English* (LDOCE), the *Cambridge International Dictionary of English* (CIDE), the *World Book Dictionary* (WBD) and the *New Shorter Oxford English Dictionary* (NSOED). Discrepancies in the contextual usage labelling in the dictionaries were established and are discussed.

Keywords: LEXICOGRAPHY, LINGUISTICS, USAGE, LABELLING, DICTIONARY, CORPUS/CORPORA, DESCRIPTION, PRESCRIPTION, PROSCRIPTION

Opsomming: Probleme van gebruiksetikettering in die Engelse leksikografie. Landau (1991: 217) stel dit dat "usage refers to any or all uses of language". Dit is die studie van die goeie, juiste of standaard gebruike van taal soos onderskei van die swak, onjuiste en nie-standaard gebruike van taal. Gebruik mag ook die studie van enige beperkinge op die gebruikswyse behels, hetsy geografies, sosiaal of temporeel. Dit waarsku gebruikers dat sekere terme nie onkrities in kommunikasie gebruik behoort te word nie. Hierdie artikel bespreek die behandeling van taalgebruikswessies in die Engelse leksikografie. Dit ontleed die etiketteringswerkwyse in ses eentalige Engelse woordeboeke naamlik: die *Oxford Advanced Learner's Dictionary* (OALD), die *Macmillan English Dictionary* (MED), die *Longman Dictionary of Contemporary English* (LDOCE), die *Cambridge International Dictionary of English* (CIDE), die *World Book Dictionary* (WBD) en die *New Shorter Oxford English Dictionary* (NSOED). Teenstrydighede in die kontekstuele gebruiksetikettering in die woordeboeke is vasgestel en word bespreek.

Sleutelwoorde: LEKSIKOGRAFIE, LINGUISTIEK, GEBRUIK, ETIKETTERING, WOORDEBOEK, KORPUS/KORPORA, BESKRYWING, PRESKRIPSIE, PROSKRIPSIE

* This article is a revised version of a paper presented at the Thirteenth International Conference of the African Association for Lexicography, organized by the Bureau of the *Woordeboek van die Afrikaanse Taal*, Stellenbosch, Republic of South Africa, 30 June–3 July 2008.

1. Introduction

Comments on usage are included in dictionaries as a guide to speakers of a language in their decisions on how to use words appropriately. They may be unsure, for example, whether **wireless** is an old-fashioned (and chiefly British) word now almost entirely replaced by **radio**, or whether to call a woman **petite**, **slim** or **slender** implies an approving attitude towards her in contrast to **skinny**, which suggests disapproval. To help learners with these difficulties, a number of labels are used in dictionaries to denote the stylistic values of words or the technical fields in which they are used; these are called 'usage labels'.

Although some readers may interpret all dictionary data as if they were 'authoritative guides' (Abecassis 2008: 1) and/or 'sources of knowledge' (Tarp and Gouws 2008: 236) on how to spell, pronounce, interpret meaning or check usage, the question surrounding the appropriateness of usage labelling as found in most dictionaries has raised concern among many linguists (Ptaszynski 2010: 411).

This article intends to show the dilemma surrounding usage labels, which makes them confusing and sometimes misleading from the point of view of second language learners. The focus will be on demonstrating the problems with regard to the labels assigned to words from the perspective of the second language speaker. The article therefore attempts to guide future lexicographers to make more informed usage labelling decisions by expounding the inadequacies of usage labelling in English lexicography.

2. Inconsistencies

Although usage labels are widely employed, their use is not at all consistent (Abecassis 2008: 3) and the same word will not represent the same label from one dictionary to the next. Very frequently, one dictionary labels one sense of a word as *slang* while the next labels the same sense *informal*. For instance, while MED uses the label *British English (BrE)*, *impolite* for the noun **arse** in the sense of 'the part of the body that you sit on', LDOCE uses the label *BrE*; CIDE *BrE & AustE*, *slightly taboo, slang*; the WBD *slang*; the OALD *BrE, taboo, slang*; and the NSOED uses the label *now coarse slang*.

It follows then that if dictionaries can agree that **bonk** is a verb that is spelled with 'o', and that one of its senses is 'to have sex with somebody', why is there discrepancy on the kinds of situations in which it may be used? If a noun like **coolie** in the sense of 'a worker in the Eastern countries with no special skills of training' is assigned the labels *old-fashioned taboo* (OALD), *offensive* (MED), *old-fashioned* (LDOCE), *dated* (WBD), and *now rare or obsolete* (NSOED), which usage label should users perceive as the correct one since all the dictionaries are in disagreement on the usage of the word?

These inconsistencies have been found to be prevalent in English dictionaries and, although lexicographers acknowledge the difficulty of labelling

words consistently (Crenn 1996), as a result of these inconsistencies, the authority of dictionaries in guiding users on contextual usage becomes questionable. Furthermore, these discrepancies can confuse users especially in the case of non-native speakers who may not have acquired a good command of the English language.

It has also been observed that lexicographers do not share the same view on the neutral register (the unmarked/unlabelled words) since in some instances words are labelled in some dictionaries while they are not labelled in others. For example, the adjective **antiquated** meaning 'old-fashioned and no longer suitable for modern conditions' is labelled *usually disapproving* in the OALD, *archaic* in the NSOED; while no label is assigned to it in the MED, LDOCE, CIDE or in the WBD; **anthropoid** is also labelled *technical* in the OALD while it is not labelled in the other five dictionaries. These discrepancies are very prevalent in English dictionaries.

It is remarkable how certain words that need to be labelled are not assigned usage labels in English dictionaries. For example, the word **cretin** meaning 'a very stupid person' is neither labelled by the NSOED nor by the WBD although the other three dictionaries label it; **dago** 'a very offensive word for a person from Italy, Spain or Portugal' is also neither labelled by the LDOCE nor by the CIDE; **slut** and **trollop**, which both mean 'a woman who has many sexual partners' are not labelled by the NSOED. Since usage labels are assigned to words to caution users about the potential danger of using some words, it may be a disservice to users to be deprived of such usage information for words like these.

3. Criteria for Usage Labelling

The question as to how lexicographers determine usage labelling is still an enigma. Ptaszynski (2010: 411) explains that 'Lexicographers have been searching in vain for an exhaustive and precise answer to the questions of which words to label in what kind of dictionaries and how to do it'. This position is reiterated by Crenn (1996) asserting that 'no particular monolingual dictionary is chosen as a basis for labelling by the lexicographers'. It is evident that usage labels are assigned to words according to the lexicographer's discretion. This probably accounts for the inconsistencies portrayed in section 2. Ptaszynski (2010: 411) explains that such problems 'stem from the lack of a firm theoretical basis for the application of diastematic information (i.e. information about restrictions on usage) in dictionaries'. It is no wonder then, that words transcribed the same way, from the same word class and with the same sense are labelled differently depending on the choice of dictionary one makes. The problem can affect both native and non-native speakers of English who, given a variety of labels for a single word, no doubt experience problems inferring the usage of such a word.

Furthermore, because there is no agreed-on criteria for usage labelling, the issue of subjectivity cannot be ruled out, for the interpretation of labels in other dictionaries and the corpus context is itself dependent on the individual lexicographer. Crenn (1996) suggests that 'although scholars [...] may point out the necessity of descriptive dictionaries, a completely objective system of register labelling has not yet been developed because lexicographers cannot feasibly let go of their subjective viewpoints'. Subjectivity is therefore another problem that may hamper good usage judgements in English lexicography.

Ptaszynski (2010) investigates the causes of the unsatisfactory theoretical treatment of information on usage, and outlines some general, theoretical guidelines for making adequate judgments concerning usage labelling. Research into a guide on usage labelling should impact positively in reducing some of the discrepancies in the usage labelling practices in lexicography. This is reiterated by Crenn (1996: 181) who asserts that although the issue of consistency in usage labelling might not be solvable in the near future, suggestions can be made to 'help lexicographers decide on register labels and to improve the consistency of register labelling in dictionaries'. Consistency has generally been linked to reliability of the dictionary (Swanepoel 2008: 216; Tomaszczyk 1988: 2), which according to Swanepoel (2008: 216) 'is increased if lexicographers adopt a set of clearly defined principles and adhere to them consistently'.

4. The Front Matter Information of Dictionaries

The front matter information of dictionaries was analyzed to show how practising lexicographers view and classify their comments on usage.

It was observed that few dictionaries provide justification for their usage labels. Questions that may arise from examining usage labels in English dictionaries can rarely be answered by the consultation of the front matter of the dictionaries. In fact, most dictionaries do not even explain what they mean by the terms 'usage labels' or 'usage labelling', taking it for granted that dictionary users know the meaning of the terms.

For instance, the LDOCE, the MED, and the CIDE explain the types of labels they select but there is no explanation of the terms 'usage labels' or 'usage labelling'. This leaves room for users not only to misinterpret what the lexicographers may mean but also to remain ignorant of the meaning of labelling, especially in the case of non-native speakers of English.

Lexicographers also have the tendency of assigning usage labels not listed in the front matter to words in the dictionary. This shows that inadequate information is given, which is evidence that the front matter notes in question are incomplete and therefore unreliable. Examples of usage labels assigned to words but not listed in the front matter are: *vulgar* in the WBD, *nonstandard* in the LDOCE and *not standard* in the CIDE.

Since some lexicographers do not explain or define the usage label terms in the front matter notes, it demonstrates their assumption that users of the

dictionaries will automatically understand what these labels mean. This assumption may be wrong, so that users may thus be misled in understanding the usage labels in question; this is substantiated by Hartmann and James (1998) who explain that the major hindrance to using dictionaries emanates from a lack of necessary skills to utilize them efficiently as well as the insufficient guidance offered by the dictionary preface to guide users to exploit its possibilities.

Crenn (1996: 49) implies that since there is great variation in the way dictionaries deal with usage, each dictionary should explain its usage labelling practices by explaining how it groups its usage labels and its explanations of these labels. This information is usually included in the dictionary front matter.

5. The Choice and Classification of Usage Labels

There seems to be little or no consensus on how to categorize the usage labels and as a result, the groups of labels as chosen by the lexicographers concerned vary from one dictionary to the next. There is also no uniformity in the choice or array of usage labels as presented in English dictionaries. The different dictionaries exhibit a diverse range of labels and categorize them differently depending on the lexicographer.

The LDOCE divides the usage labels into four categories namely: regional, foreign, attitudinal (formal, informal, humorous, approving) and style labels (biblical, dialect, law, literary, non technical, old-fashioned, old use, poetic, slang, spoken, taboo, technical and trademark).

The OALD does not divide its labels into broad groups, it simply presents the following list: approving, disapproving, figurative, formal, humorous, informal, ironic, literary, offensive, rare, slang, spoken, technical, AmE, BrE, dialect, old-fashioned, written, taboo, old use, saying and trademark.

The array of labels in the MED is divided into three categories namely: style and attitude labels (*formal, humorous, impolite, literary, offensive, old-fashioned, spoken, very formal, and very informal*), subject labels (for example *business, computing and journalism*) and regional labels (for example *Australian, Canadian and American*).

In the WBD there are labels like archaic, dialect, English speaking variation labels, foreign language labels, informal, obsolete, poetic, professional terms, slang, substandard, trademark, unfriendly use and figurative usage.

The NSOED divides its labels into four categories, namely: those showing restriction to geographical area, those referring to style or register, those indicating branch of knowledge or field of activity, and lastly, labels indicating frequency or extent of use.

The CIDE has the following array of labels: approving, dated, disapproving, not standard, old use, poetic, regional, slang, female, figurative, formal, humorous, informal, law, literary, male, medical, specialized, taboo and trademark.

It is evident that each dictionary makes its choice of the labels to be assigned to words. What is not clear is the criteria the lexicographers use in selecting certain labels over others.

Certain labels are not listed in some dictionaries, while they occur in others. For example, the OALD has labels mainly atypical to the other five dictionaries such as *ironic* and *saying*, the labels *very formal* and *very informal* are commonly used by the MED and rarely in the other five dictionaries. The WBD lists and explains the term *combining form* as a usage label but the other dictionaries do not have this as a usage label. The CIDE also employs *female* and *male* as labels, the two being alien to the other five dictionaries studied.

It is therefore apparent that there is no agreement on the number of dimensions in usage. It is evident that each dictionary makes its choice of the number, type and classification of usage labels to be assigned to words. Quirion (1995: 347) states that 'the reduction in the number of labels could lead to some consensus in labelling'.

6. Definitions

As already mentioned in Section 4, some of the English dictionaries do not try to explain the terms 'usage label' or 'usage labelling'; they simply embark on listing the labels assuming that users can discern the meaning by looking at the list of examples of usage labels or otherwise by looking at the dictionary articles, particularly at the entries assigned usage labels and then automatically comprehending what is meant by usage, which may not be the case. This has been established in three of the six dictionaries consulted for this study, namely the LDOCE, the MED and the CIDE.

Some of the definitions of the labels are imprecise, leaving room for questions. For example, the OALD defines the label *spoken* as indicating 'expressions used mainly in informal conversation'. This is a vague definition because of the use of the word *mainly*. Should it then be concluded that an adjective like **kaput** in the sense of 'not working properly' or an idiom like **no probs** with the meaning 'there is no problem' and labelled *spoken* can sometimes be used in say formal situations? This type of definition can therefore be misleading.

As a result of the inadequacies in the definitions accorded, some of the labels in English dictionaries seem to overlap and in the end, it appears as if there is a replication. In other words, labels, which are apparently synonymous, seem to be assigned to unconnected words. The inadequacy in the way these labels are defined can be misleading.

This can be illustrated by the way the labels *archaic* and *obsolete* are defined in the WBD; *archaic* is described as referring to 'a word that is no longer in general use. It may be found in special contexts, such as law or special styles of writing such as the Scriptures. Sometimes modern writers use archaic words to give their writings an old-fashioned flavor.' This definition has the same implications as the label *obsolete*, which is defined in the same dictionary as indica-

tion for 'a word only found in writings of an earlier time or in modern writings that imitate the style of earlier writing.' These definitions of the two apparently different labels need to be revised if the difference is to be realized, otherwise they can be misleading as the definitions do not clearly show the difference.

Additionally, when the difference between the way OALD and LDOCE define the labels *old use* and *old-fashioned* is analyzed, a discrepancy in the LDOCE's definition becomes evident. The LDOCE defines *old use* as indicating 'a word used in earlier centuries' and it defines *old-fashioned* as indicating 'a word that was used earlier in this century but would sound old-fashioned today'. On the other hand, the OALD defines *old use* as referring to 'expressions that are no longer in current use' while it defines *old-fashioned* as referring to 'expressions that are passing out of current use'. On analyzing the way the two dictionaries define the two different labels, it is evident that the OALD shows a clear distinction between the labels leaving no doubt that there is indeed no duplication of the labels. The LDOCE on the other hand does not evince this difference. Consequently, it could be argued that *old use* and *old-fashioned* have the same meaning and therefore are synonyms since even a word used in the earlier century (*old use*) would certainly sound old-fashioned today. This illustrates the deficiencies of some of the definitions of labels used in English dictionaries.

The CIDE does not even try to define the labels selected. It simply gives a list of the labels without even briefly explaining them.

Undefined or inadequately defined terms coupled with the fact that some label definitions vary from one dictionary to the other shows the inaptness of usage labelling in English lexicography.

7. Ambiguity

The editors do not spell out what specific areas are covered by labels like *informal*, *nonstandard*, *slang*, *vulgar*, *disparaging*, *offensive* ... To give a few examples: Does the label *literary* mark only the effect on the reader or listener since it is defined in the OALD as indicating 'language used mainly in literature and imaginative writing', although logically, the use of literary language depends also on the setting and the kind of communication intended?

Similarly, *nonstandard* seems to indicate only the kind of speaker who might use it since it is defined as 'not conforming to the speech or grammar of educated persons and often regarded as a marker of low social status'. However, can someone not adopt nonstandard words in certain settings to be witty/comical? Or perhaps in an attempt to satirize a situation?

Further still, a label like *slang* as explained by Crenn (1996: 186) 'is understood differently by each dictionary (and presumably) every user and unless dictionaries specify to which group a particular slang word corresponds ...', for example *slang military*, *slang law* ... then labels showing degree of formality (such as *very formal*, *informal*) could be more appropriate in achieving clarity.

In addition, one group may read one particular usage as *informal* or *non-standard* whereas another may use it as entirely natural for any conceivable circumstance. According to Landau (1991: 255-258), dictionaries cannot in good faith mark out some usages as *informal* without saying to whom they are informal. *Informal* may actually mean informal to those of the higher social classes, especially older, well-educated authors and professors in the humanities.

The combinations of usage labels assigned to words also sometimes pose as deterrents to clarity. It can be difficult to comprehend what exactly the lexicographer means by assigning two or more labels to a word. This problem may be illustrated by the following examples:

The noun **vale** meaning 'a valley' is assigned the label *now archaic or poetic* in the NSOED. Likewise, the noun **vagrant** meaning 'a person who has no job or home especially one who begs' is labelled *slightly dated or law* in the CIDE, the adjective **wee** meaning 'very small in size' is labelled *colloquial also chiefly Scottish* in the NSOED, the noun **wank** in the sense of 'an act of masturbation' is assigned the label *British and Australian taboo slang* in the CIDE, the adjective **woebegone** meaning 'looking very sad' is assigned the label *mainly literary* in the MED, and the noun **gloom** in the sense 'total darkness' is labelled *especially literary* in the LDOCE.

The use of conjunctions like 'or', 'and', and adverbs like 'slightly' and 'usually' can be puzzling. The entry **recuperate** is labelled *slightly formal* in the CIDE. The question then arises whether it should be used in formal situations. If it is slightly formal, then exactly in what situations it may be used? **Outlandish**, an adjective assigned the label *usually disapproving* in the OALD and **politicking** labelled *especially disapproving* in the CIDE are other examples.

The use of indications like 'mainly', 'especially', 'chiefly' in representing usage sometimes cause ambiguity as dictionary users may be uncertain as to what exactly the lexicographer means.

8. Dictionary Tone

The approach in a dictionary can be descriptive, prescriptive, or proscriptive (Tarp and Gouws 2008: 235; Van der Merwe 2008: 346). When it is descriptive, 'it avoids classifying occurring forms as either recommended or not permissible. It reflects the spectrum of actual language in use' (Tarp and Gouws 2008: 237). It shows the user 'what' by illustration, it says, 'it is like this' thereby leaving the responsibility to the user to decide whether a word is adequate for use in a particular situation. When it is prescriptive, 'lexicographers inform the user how to use language' (Tarp and Gouws 2008: 236); they say 'do it like this'. A prescriptive approach imposes the lexicographer's 'point of view on the dictionary and the target users of the dictionary' (Tarp and Gouws 2008: 236). They explain the term 'proscription' in lexicography which deviates from the English word 'proscribe' since it does not refer to the state of being forbidden.

A proscriptive approach informs 'the user not only about language use but also about the form recommended by the lexicographer' (Bergenholtz 2003: 13).

Landau (1989: 194) suggests that until the latter half of the 20th century, dictionaries felt duty-bound to prescribe 'correct usage'. However, there is claim that 'the attitude to lexicography has changed since the 1970s with dictionaries being increasingly descriptive rather than prescriptive' (Abecassis 2008: 2). This study exposes the contrary, with lexicographers being implicitly prescriptive as illustrated below.

The label *taboo* is defined in the OALD as 'a word that should not be used because it is very rude or offensive'. It is evident that the lexicographer is telling his users not to use this word because of the consequences that would follow. The lexicographer in this case uses a prescriptive tone. Examples of words labeled *taboo* are: a verb like **fuck** which means 'to have sex with somebody' as labeled in the LDOCE, the NSOED and CIDE. The noun **prick** meaning 'penis' is assigned the label *taboo* in the LDOCE, the CIDE, the OALD, and *coarse slang* which is synonymous with *taboo* in the NSOED, the noun **pussy** meaning 'the female sexual organs especially the vulva' is also labeled *taboo* in the OALD, the CIDE, the LDOCE and the NSOED.

Offensive is defined by the MED as 'extremely rude and likely to cause offense'. This is another typically prescriptive label where the lexicographer implicitly warns the dictionary user against using the word in question. Examples of nouns labeled accordingly are: **queen** to mean 'a male homosexual who behaves more like a woman than a man', this is found in the MED, **shit** 'an unpleasant person who treats other people badly' is another example from MED, **squaw** 'a native American woman' is an example from the NSOED.

The *standard/substandard* dichotomy implies a right and a wrong in a language. Abecassis (2008: 2) states that dictionaries 'still strongly emphasize the standard language'. According to Landau (1989: 175), 'variations based on register have no right or wrong; in certain situations, a word may be appropriate while in others, it may be inappropriate'. An example of a word labeled *substandard* is **ain't** to mean 'is it not?' as found in the WBD.

Other labels used in English dictionaries that are indicators of prescriptiveness are *vulgar* and *obscene*. *Vulgar* is prescriptive because it is assigned to a word (s) to show that it is generally uncouth, crude or coarse. The implication here is that such a word should never be used because it is basically in bad taste. *Obscene* as a label is prescriptive because it suggests that a word it has been assigned to is abominable and therefore should not be used. Prescriptiveness limits the use of a given word to a specific situation of communication and yet, in certain situations, a word may be adequate and inadequate in others.

Tarp and Gouws (2008) and Bergenholtz (2001, 2003 as cited in Tarp and Gouws 2008: 239) and Van der Merwe (2008: 345-346) explore the issue of dictionary tone/approach and make recommendations that are complemented by this study, to guide lexicographers in the area of dictionary tone.

9. Conclusion

There is no agreed-on criteria for making usage decisions in English lexicography. This is probably the reason for most of the problems concerning usage labelling in English dictionaries. It causes the inconsistencies which characterize usage labelling, and the problems, which emerge from such inconsistencies. As Stein (2002: 14) asserts, 'it is admittedly very difficult to make objective assessments on the social status of words but it seems ... that we need much more research in this area'. Leech and Nesi (1999: 295) recognize that dictionaries 'fall well short of perfection'; I contend that lexicographers, in their quest for perfection, can benefit invaluablely from research from lexicographers as well as different perspectives of users, like the second language user perspective provided by the current study.

References

- Abecassis, M. 2008. The Ideology of the Perfect Dictionary: How Efficient can a Dictionary be? *Lexikos* 18: 1-14.
- Barnhart, C. (Ed.). 1963. *The World Book Dictionary*. London: The World Book Inc.
- Bergenholtz, H. 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Brown, L. (Ed.). 1993. *The New Shorter Oxford English Dictionary on Historical Principles*. 2 volumes. Oxford: Clarendon Press.
- Crenn, T. 1996. *Register and Register Labelling in Dictionaries*. Ph.D. Thesis. Ontario: University of Ottawa, School of Translation and Interpretation.
- Hartmann, R.R.K. and James, G. 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Hornby, A.S. (Ed.). 1995. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.
- Landau, S.I. 1989. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Landau, S.I. 1991. Approaches to Meaning and their Uses in Lexicography. *Dictionaries*: 7: 261-9.
- Leech, G. and H. Nesi. 1999. Moving Towards Perfection: The Learners' (Electronic) Dictionary of the Future. Herbst, T. and K. Popp (Eds.). 1999. *The Perfect Learner's Dictionary*. Lexicographica Series Maior 95: 295-308. Tübingen: Max Niemeyer Verlag.
- Procter, P. (Ed.). 1995. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press.
- Ptaszynski, M.O. 2010. Theoretical Considerations for the Improvement of Usage Labelling in Dictionaries: A Combined Formal-Functional Approach. *International Journal of Lexicography* 23 (4): 411-442.
- Quirion, J. 1995. How Can Usage Labels Better Label Usage. Krawutschke, P. (Ed.). 1995. *Connections: Proceedings of the 36th Conference of the American Translators Association, November 8-12, Nashville, Tennessee*: 341-349. Medford (NJ): Information Today.
- Summers, D. (Ed.). 1987. *Longman Dictionary of Contemporary English*. Harlow: Longman.

- Swanepoel, P.** 2008. Towards a Framework for the Description and Evaluation of Dictionary Evaluation Criteria. *Lexikos* 18: 207-231.
- Stein, G.** 2002. *Better Words: Evaluating EFL Dictionaries*. Exeter: University of Exeter Press.
- Tarp, S. and R.H. Gouws.** 2008. A Lexicographic Approach to Language Policy and Recommendations for Future Dictionaries. *Lexikos* 18: 232-255.
- Tomaszczyk, J.** 1988. The Bilingual Dictionary under Review. Snell-Hornby, M. (Ed.). 1988. *ZurilEx '86 Proceedings. Papers Read at the EURALEX International Congress, University of Zurich, 9-14 September 1986*: 289-297. Tübingen: Francke Verlag.
- Van der Merwe, M.F.** 2008. Wine and Words: A Trilingual Wine Dictionary for South Africa. *Lexikos* 18: 337-348.

Dr. J.C.M.D. du Plessis

Redakteur van *Lexikos* 1997 tot 2010

Dr. Johan du Plessis het oor die afgelope veertien jaar sinoniem geword met *Lexikos*. Toe hy in 1997 die redakteurskap van *Lexikos* aanvaar het, het hy nog die veeleisende pos van Eindredakteur van die *Woordeboek van die Afrikaanse Taal* beklee. Skielik het hy nie meer kantoorure gehad nie: hy was min of meer permanent op kantoor.

In die voorwoord van die eerste uitgawe van *Lexikos* in 1991 skryf die redakteur, Pieter Hartevelt, dat "die gesprek tussen leksikograaf en taalkundige, asook dié tussen leksikograaf en leksikograaf, op 'n gereelde grondslag by kongresse en seminare behoort plaas te vind" maar dat dit nie altyd prakties haalbaar is "waar die deelnemers hulle oor so 'n uitgestrekte geografiese gebied soos Afrika bevind nie. Daarteenoor dien 'n vakpublikasie as 'n uitstekende medium en voertuig vir so 'n gesprek". *Lexikos* word dan ook in 1996 die ampelike mondstuk van AFRILEX — die leksikografiese vereniging van Afrika.

Johan du Plessis het met die afgelope 14 uitgawes van *Lexikos* op uitmuntende wyse daarin geslaag om nie alleen 'n voertuig te verskaf vir hierdie leksikografiese gesprek nie, maar ook verseker dat die gesprek op 'n wetenskaplike vlak plaasvind wat deur ander beny word. Nie net het hy meer en meer gespreksgenote betrek nie, maar hy het ook studente en leksikograwe opgelei om waardige gespreksgenote te word. Hier moet veral genoem word die talle leksikografiestudente van veral Gaboen en jong of onervare leksikograwe van die Nasionale Leksikografiese Eenhede wat deur hom begelei is om wetenskaplike artikels van gehalte oor die leksikografie te skryf.

Lexikos het sy ontstaan te danke aan die strewe van die Buro van die WAT om 'n groeipunt en stimulus vir leksikografiese denke en aktiwiteite te wees. Johan het as Eindredakteur van die WAT in 1997 die redakteurskap van *Lexikos* oorgeneem en na sy uittrede in 2000 vir nog tien jaar voortgegaan as redakteur tot in 2010 met *Lexikos* 20. Hy het vanaf 2001 tot 2010 steeds byna elke dag kantoor toe gekom en miskien 'n bietjie later as die ander personelede op kantoor gekom, maar as almal al huiswaarts gekeer het, was Johan steeds hier. Net Vrydae het hy homself 'n wegbreekkansie gegun.

Danksy die ywer van Johan du Plessis strek die reikwydte van *Lexikos* veel verder as Afrika en word dit in 23 lande versprei. Dit word internasionaal geag as een van die toonaangewendste akademiese joernale in die vakgebied linguïstiek. Die Amerikaanse Thomson Reuters, wat allerweë beskou word as die belangrikste internasionale keurder van kwaliteit in akademiese joernale, het *Lexikos* in ses van hulle indekse opgeneem, nl. *Arts and Humanities Citation Index*, *Current Contents/Arts & Humanities*, *Current Contents/Social and Behavioral Sciences*, *Journal Citation Reports/Social Sciences Edition*, *Social Sciences Citation*

Index, en *Social Scisearch*. Twee groot Nederlands-Europese uitgewers, naamlik Elsevier en Brill, lys *Lexikos* op hulle aanlynbibliografieë (SCOPUS en *Linguistic Bibliography Online*). As internasionaal geïndekseerde tydskrif, word 'n impakfaktor in die kategorie Linguistiek vir *Lexikos* bereken en met 'n impakfaktor van 0.667, is *Lexikos* nou 49ste in rang uit 92 internasionale tydskrifte in die Linguistiekkategorie, wat dit in die derde kwartiel plaas.

Ook *African Journals OnLine* (AJOL) het in 2009 verlot gevra om *Lexikos* aanlyn te huisves op hulle webwerf vir toonaangewende akademiese joernale uit Afrika.

Die wetenskaplike benadering, toewyding, sorgsaamheid en deeglikheid wat Johan du Plessis as redakteur kenmerk, het ongetwyfeld 'n deurslaggewende rol gespeel in die verwerwing van die statuur en aansien wat *Lexikos* vandag geniet.

Dr J.C.M.D. du Plessis Editor of *Lexikos* 1997 to 2010

Over the past 14 years, Dr Johan du Plessis has become synonymous with *Lexikos*. When he accepted editorship of *Lexikos* in 1997, he was still holding the demanding position of Final Editor at the *Woordeboek van die Afrikaanse Taal*. Suddenly he no longer had office hours: he was more or less permanently at the office.

In 1991 the editor of the first volume of *Lexikos*, Pieter Harteveld, stated in the Foreword that "meetings between lexicographer and linguist, as well as between lexicographer and lexicographer, should take place on a regular basis at congresses and seminars", but that it is not always possible in practice, especially "where participants are drawn from such a vast area as Africa. On the other hand a specialist publication serves as an excellent medium and vehicle for such discussions". In 1996 *Lexikos* therefore became the official mouthpiece of AFRILEX — the lexicographic association of Africa.

With the past 14 volumes of *Lexikos* Dr. Johan du Plessis succeeded admirably, not only in supplying a vehicle for this lexicographic discussion, but also in ensuring that the discussion took place on a scientific level that is envied by others. Not only did he involve more and more partners in the discussion, but he also tutored students and lexicographers to make them worthy partners. In this regard one should mention the many lexicography students, especially from Gabon, and young or inexperienced lexicographers from the National Dictionary Units who were guided to write scientific articles of quality.

When *Lexikos* was established, the aim of the Bureau of the WAT was that it should become a growth point and stimulus for lexicographical thought and activity. Johan took over the role as editor of *Lexikos* in 1997 when he was Final Editor of the WAT. After his retirement in 2000 he stayed on as editor for a further ten years — until *Lexikos* 20 in 2010. He continued to come in to the office nearly every day from 2001 to 2010. He may well have arrived somewhat later than the other personnel members but, when everyone else had already gone home, Johan was still here. Only on Fridays he allowed himself to relax and was not seen at the WAT.

Through the enthusiasm of Johan du Plessis the influence of *Lexikos* stretches much further than Africa. It is distributed to 23 countries, and it is regarded internationally as one of the most prestigious academic journals on the subject of linguistics. Thomson Reuters of the USA, which is widely regarded as the most important international judge of quality in academic journals, included *Lexikos* in six of their indices, namely *Arts and Humanities Citation Index*, *Current Contents/Arts & Humanities*, *Current Contents/Social and Behavioral Sciences*, *Journal Citation Reports/Social Sciences Edition*, *Social Sciences Citation*

Index, and *Social Scisearch*. Two large Dutch-European publishers, Elsevier and Brill, list *Lexikos* on their online bibliographies (SCOPUS and *Linguistic Bibliography Online*). As an internationally indexed journal an impact factor is calculated for *Lexikos* in the category Linguistics. With an impact factor of 0.667, *Lexikos* is now 49th out of 92 international journals in the category of Linguistics, which puts it in the third quartile.

In 2009 *African Journals OnLine* (AJOL) also asked permission to host *Lexikos* online on their website for prestigious academic journals from Africa.

The scientific approach, dedication, attentiveness and thoroughness that distinguishes Johan du Plessis as an editor, undoubtedly played a crucial role in gaining the stature and prestige that *Lexikos* enjoys today.

Quand la pratique lexicographique se modernise en RD Congo. Note sur *Nkòngamyakù Cilubà–Mfwàlànsa*, dictionnaire bilingue de NgoSemzara Kabuta*

Crispin Maalu-Bungi, *Faculté des Lettres et Sciences Humaines, Université de Kinshasa, République démocratique du Congo (crismaalu@yahoo.fr)*

Résumé: Le cilubà est l'une des quatre langues nationales congolaises parlée au centre-sud du pays. Elle a acquis le statut de langue écrite depuis la fin du 19^{ème} siècle et possède à ce jour plusieurs ouvrages lexicographiques dont le premier, daté de 1881, est un lexique ciluba-allemand réalisé par un explorateur allemand. *Nkòngamyakù Cilubà–Mfwàlànsa* est un dictionnaire bilingue élaboré par NgoSemzara Kabuta, professeur de linguistique et littérature africaines à l'Université de Gand, dans le triple but de compléter et combler les lacunes du dernier dictionnaire du genre publié en 1960 par un prêtre catholique, rendre compte de l'évolution de la langue et proposer la forme standard du cilubà tel qu'il doit être utilisé dans les communications institutionnelles, notamment dans l'enseignement. Par rapport aux ouvrages des prédécesseurs, missionnaires et agents territoriaux, ce nouvel outil de référence apporte plusieurs innovations, dont principalement: l'existence du dictionnaire sous formats électronique et support papier, l'augmentation du nombre d'entrées de 10 000 à 16 000, l'utilisation d'une orthographe simple, pratique et efficace susceptible de faciliter la lecture comme l'apprentissage de la langue, le nombre significatif de paramètres et de justifications sur chaque entrée pour en rendre la compréhension plus aisée, l'ajout d'un abrégé de grammaire, entièrement en cilubà, une manière de contribuer au développement, à la promotion et à l'enrichissement de cette langue à travers notamment la terminologie grammaticale. Malgré les erreurs dues à son caractère novateur, cet ouvrage lexicographique dont la valeur est, à divers égards, incontestable, constitue une réponse à l'appel lancé naguère aux dictionnaristes africains, singulièrement bantu, en faveur de travaux plus complets.

Mots clés: DICTIONNAIRE DE LANGUE, DICTIONNAIRE BILINGUE, DICTIONNAIRE PLURILINGUE, DICTIONNAIRE SOUS FORMATS ÉLECTRONIQUE ET SUPPORT PAPIER, ENTRÉE DICTIONNAIRIQUE, MÉTHODE AFRILEX, ABRÉGÉ DE GRAMMAIRE, PARAMÈTRES ET JUSTIFICATIONS

Abstract: When lexicographic practice is being modernized in the DR Congo. Note on *Nkòngamyakù Cilubà–Mfwàlànsa*, Bilingual Dictionary of Ngo-

* NgoSemzara Kabuta. *Nkòngamyakù Cilubà–Mfwàlànsa*. 2008, XI + 365 pp. ISBN 2-930393-07-06. Gand: Recall.

Semzara Kabuta. Ciluba is one of the four national Congolese languages spoken in the central south of the country. Since the end of the 19th century, it has acquired the status of a written language and up to the present possesses several lexicographic works of which the first, dating from 1881, is a Ciluba–German vocabulary compiled by a German explorer. *Nkòngamyakù Cilubà–Mfwàlànsa* is a bilingual dictionary developed by NgoSemzara Kabuta, professor of African linguistics and literature at the University of Ghent, with the triple aim of complementing and filling the gaps of the last dictionary in this genre published in 1960 by a Catholic priest, giving an account of the evolution of the language and proposing the standard form for Ciluba so that it could be used in institutional communication, especially in education. Compared with the works of predecessors, missionaries and regional officials, this new reference tool brings several innovations, mainly: the provision of a dictionary in electronic format and paper, the increase of the number of entries from 10 000 to 16 000, the use of a simple, practical and effective orthography likely to help with reading when learning the language, the significant number of parameters and justifications for each entry to make comprehension easier, the addition of a minigrammar, wholly in Ciluba, a way of contributing to the development, promotion and enriching of this language, particularly through grammatical terminology. In spite of the errors owing to its innovative character, this lexicographic work whose merit is, in various respects, unquestionable, is an answer to the call recently made to African, especially Bantu lexicographers for more comprehensive works.

Keywords: LANGUAGE DICTIONARY, BILINGUAL DICTIONARY, MULTILINGUAL DICTIONARY, DICTIONARY IN ELECTRONIC FORMAT AND PAPER, DICTIONARY ENTRY, AFRILEX METHOD, MINIGRAMMAR, PARAMETERS AND JUSTIFICATIONS

La langue lubà-kàsai (L31) ou plus simplement le cilubà, comme le nomment ses locuteurs, figure aujourd'hui parmi les langues congolaises qui ont le plus bénéficié de l'attention des chercheurs.¹ *Nkòngamyakù Cilubà–Mfwàlànsa* (Dictionnaire Cilubà–Français) vient donc compléter la liste d'ouvrages lexicographiques de cette langue dont le premier, daté de 1888, est un lexique cilubà–allemand rédigé par l'explorateur H. von Wissmann,² arrivé au Kàsai en octobre 1881. À travers cet outil de référence, l'auteur, J. Kabuta, alias *Ngo Semzara*³ se propose de répondre à l'appel lancé entre autres par J.G. Kiango une dizaine d'années auparavant, face aux lacunes que présentent la plupart des dictionnaires des langues africaines qui, selon K.A. Mairo (1990: 10), "were meant to serve as expeditionary guides for the European explorers (or) were produced to be used by the missionaries in learning the African languages for the purpose of evangelization". À ce sujet en effet, J.G. Kiango (2000: 5) écrit:

For a very obvious reason, these dictionaries were not aimed at being complete records of the languages concerned and up to this time, no such records exist. This is one of the facts which the present Bantu lexicographers should understand. Therefore, the tendency of the current Bantu lexicographers to perpetuate the tradition of compiling monolingual and bilingual dictionaries for foreign learners as their main dictionary projects, will not push the Bantu lexicography one step further from where the pioneer lexicographers have left. It is high time now

the contemporary Bantu lexicographers embarked on compiling complete and comprehensive records of the Bantu languages.

Nkòngamyakù Cilubà–Mfwàlànsa est un dictionnaire de langue, plus précisément un dictionnaire bilingue qui comprend trois parties: une introduction, une *wordlist* et une esquisse grammaticale.

L'introduction (*kàadyosha*) est constituée de quatre points précédés d'une note où l'auteur présente les objectifs poursuivis dans l'élaboration de ce dictionnaire, à savoir compléter et combler les lacunes du dernier du genre confectionné en 1960 par A. De Clercq et E. Willems et qui n'avait que plus ou moins 10 000 entrées alors que le sien en a 16 000, ce qui en définitive, lui permet de rendre compte à la fois de l'évolution de la langue et de la manière dont elle se parle actuellement; proposer la forme standard du cilubà tel qu'il doit être utilisé dans l'enseignement comme dans d'autres communications institutionnelles. Le premier point (*Tumanyinu nè tukèèpeshilu*) présente les symboles et les abréviations employés alors que le deuxième (*Ngaashilu wa kanungu*) donne, à partir de quelques exemples, les éléments constitutifs des différents articles du dictionnaire notamment la vedette ou l'entrée (*mwakù wà kumvwija*), la catégorie grammaticale (*cifuku*), la définition (*dicincika*) ou l'équivalent (*ngandamwinu*), l'exemple (*cileejilu*), le renvoi (*mufunkumu*), l'idiotisme, la parémie (*lusùmwinù*), etc. Le troisième point est une présentation de ce dictionnaire sous sa forme électronique tel qu'il se trouve sur deux sites dont celui du CIYEM, sigle lubà de *Cikèbulwidi cyà pa Yètù Myakulu* (Centre de recherche sur les langues nationales), créé en 2003 et dont l'auteur retrace la genèse ainsi que les objectifs. Il précise ensuite que cet ouvrage lexicographique dont environ 10 000 mots viennent du dictionnaire Tshiluba–Français de De Clercq et Willems, diffère des autres sur les points suivants:

- il est le résultat de recherches fouillées, basées sur la méthode d'Afrilex;
- la compréhension quasi-totale des mots est rendue possible par le nombre significatif d'informations qui l'accompagnent, ce en quoi l'auteur rencontre la préoccupation du tanzanien J.G. Kiango (2000: 6) qui plaide en faveur de dictionnaires plus complets: "The West-African lexicographers just as the Bantu lexicographers have to embark on the projects that would produce dictionaries that are complete records of the languages concerned";
- la méthode utilisée facilite la création de nouveaux termes et leur normalisation.

Dans le quatrième et dernier point de l'introduction, l'auteur indique la procédure à suivre pour consulter le dictionnaire électronique compte tenu des différences qu'il présente avec le dictionnaire imprimé. Pour terminer, il invite les lecteurs à lui faire des critiques constructives et des propositions concrètes en vue d'améliorer la qualité de ce dictionnaire qui, de par son caractère novateur, ne peut être exempt d'erreurs.

Alors que la deuxième partie du dictionnaire, de loin la plus importante (pp. 1-344), présente l'ensemble des articles ou la *wordlist* plus exactement, la troisième et dernière partie est une esquisse grammaticale du cilubà entièrement rédigée dans cette langue avec ça et là des traductions en français de la métalangue et de quelques énoncés pour une meilleure compréhension de la matière traitée. L'auteur y examine plusieurs questions entre autres l'alphabet (*lùfwàbeetà*), les règles d'orthographe (*mibikù yà mfùndilu*) et de prononciation (*nshìndumwìnu*), les substitutifs (*mìpinganyì*), les catégories grammaticales et formes variables (*cifùku nè mbidi mishintuluki*), les déterminants spéciaux (*misunguluji yà pabwàyi*), les possessifs (*bwenacintu*), les présentatifs (*tulubudi*), les morphèmes verbaux (*mifwèmà yà lwakù*), les paradigmes et tiroirs (*ndeejilu nè nsànzù*), les tiroirs de kwikwala (*nsanzu yà lwakù kwikala*), les extensions verbales (*misampu*), les suffixes de dérivation (*mikwàmishi*), les variantes combinatoires (*bifwànu byeyemenyi*), les formes invariables (*mbidi mikashintuluki*), etc.

Comparé aux travaux des prédécesseurs, essentiellement œuvre de missionnaires et d'agents coloniaux,⁴ *Nkòngamyakù Cilubà-Mfwàlànza* est un ouvrage de référence qui s'impose, à plus d'un titre. Aboutissement de recherches manifestement fouillées, ce dictionnaire témoigne de l'érudition de son auteur ainsi que de sa détermination à mettre à la disposition du public un travail de qualité. En effet, j'ai noté pour ma part les innovations significatives ci-après, outre celles signalées par lui-même:⁵

- (1) l'utilisation de l'orthographe des langues congolaises recommandée par le Premier Séminaire National des Linguistes du Zaïre en 1974 qui, du fait de sa simplicité, de son caractère pratique et de son efficacité, est susceptible de faciliter la lecture comme l'apprentissage de la langue par les locuteurs natifs et non-natifs. En ce qui concerne le cilubà plus précisément, il s'agit de:
 - l'adoption des semi-voyelles *y* et *w* à la place de *i* et *u* qui met un terme à la mauvaise représentation graphique de ces sons et évite les risques de confusion avec ces deux dernières voyelles dans une séquence de voyelles différentes comme par exemple dans les mots *kwina* (plonger) et *kuina* (les plonger), *kwiba* (dérober) et *kuiba* (les voler), *twàkulè* (parlons) et *tùakùlè* (parlons-en), etc.;
 - la notation systématique des tons et de la quantité vocalique, celle-ci par le redoublement du signe, suivant en cela la recommandation de l'Institut Africain International;⁶
 - la transcription de l'affriquée palatale sourde par le signe *c* en remplacement du trigraphe *tsh* en vertu du principe de base de toute orthographe à vocation fonctionnelle qui veut que tout son distinctif soit représenté par un seul signe;⁷
- (2) la pertinence et le nombre relativement élevé d'informations fournies sur chaque entrée qui confirment la nature dictionnaire de l'ouvrage et

traduisent le souci qu'a l'auteur de satisfaire au maximum les besoins des utilisateurs;

- (3) l'intégration d'anthroponymes, de noms d'entités administratives du Kàsaï, d'emprunts faits aussi bien aux langues européennes qu'africaines et de néologismes techniques et scientifiques forgés par l'auteur lui-même, principalement dans le domaine linguistique;
- (4) l'abrégé de grammaire lubà sur une quinzaine de pages, un événement dans l'histoire centenaire de cette langue nationale qui a acquis le statut de "langue écrite" à la fin du 19^{ème} siècle et celui de "langue littéraire" au début du 20^{ème}.⁸ Cet essai, une première du genre dans cette langue, administre la preuve de la capacité qu'ont les langues africaines d'exprimer, à l'instar d'autres langues du monde, les réalités scientifiques et techniques modernes. À ce titre et au regard de l'impératif pédagogique de la transmission des connaissances en langues nationales dans le cadre formel et institutionnel qu'est l'enseignement, cet ouvrage arrive à point. En effet, il se présente globalement comme une contribution significative à l'effort de développement, de promotion et d'enrichissement de la langue lubà-kàsaï, l'une des quatre langues nationales congolaises. C'est pourquoi son auteur est en droit d'être félicité, voire admiré quand on sait par exemple que né au Katanga de parents kasaiens, il a eu le kiswahili comme langue première et que, selon son propre témoignage, il n'a séjourné au Kàsaï qu'à de rares occasions jusqu'à l'âge de quinze ans où le destin lui donna la Belgique comme nouvelle patrie.⁹ Au sujet de sa pratique du cilubà et du kiswahili, il note ce qui suit en rapport avec le climat d'insécurité et de haine ethnique nourrie par les Katangais contre les Kasaiens, à l'accession du Congo à l'indépendance en 1960 (Kabuta 2009: 119):

Nos parents, au moins, savaient d'où ils étaient venus et ils y retournaient de temps en temps. Nous, les enfants, nous étions nés au Katanga. Le Katanga, c'était notre pays! A part l'Europe que j'avais visitée deux ans auparavant, je n'en connaissais pas d'autre. Je ne pouvais pas imaginer avoir un autre pays, en dépit des nombreux événements qui nous rappelaient que nous étions originaires du nord: la famille proche ou lointaine qui venait du Kasaayi en visite, les parents qui y allaient en voyage, ou les commerçants balubà qui circulaient entre les 2 provinces et rapportaient des produits qu'on trouvait seulement là-bas (...) Qui plus est, le Kasaayi nous avait toujours paru comme étant une contrée arriérée. Aucun d'entre nous n'aurait eu l'idée d'aller y vivre. Nous, les Katangais, nous nous différencions de ces gens un peu rustres qui venaient du village. Nous, nous parlions le kiswahili, tandis qu'eux, ils parlaient le kikasayi. Nous étions un peu gênés d'apprendre qu'ils étaient nos frères ou nos sœurs.

Ceci dit, je voudrais, dans les lignes qui suivent, formuler quelques observations en guise de contribution à l'amélioration des éditions futures de cet

ouvrage de référence et répondre ainsi au vœu émis par l'auteur lui-même (Kabuta 2009: IX, XI):¹⁰

Quand bien-même ce dictionnaire serait le premier parmi les ouvrages lexicographiques en langues africaines auxquels vous seriez habitués, il faut savoir qu'il ressemble à un enfant qui fait ses premiers pas et comme tel, il ne peut être exempt d'erreurs. Un proverbe de chez-nous ne dit-il pas que "c'est en soufflant à plusieurs sur le feu qu'on fait cuire le haricot?" Voilà pourquoi nous attendons vos critiques et suggestions pour nous permettre d'améliorer ce dictionnaire dont pourront se servir les écrivains, les traducteurs, les élèves, les chercheurs, les créateurs de nouveaux termes bref, tous ceux qui s'intéressent à la langue lubà-kàsaï. Certains mots repris dans ce dictionnaire sont notre propre œuvre parce que nous voulions tout exprimer dans cette langue. Les lecteurs sont donc invités à nous communiquer leurs avis là-dessus, à travers le Ciyem afin que celui-ci puisse retenir les termes qui conviennent (...) Sachez donc qu'un travail de ce genre ne peut manquer d'erreurs ni de fautes. Les Anciens n'ont-ils pas dit: "Le forgeron forge au bord de la route pour que les passants lui montrent les erreurs qu'il pourrait commettre" (...) Nous publions donc ce travail commencé bien longtemps avant dans l'état où il est. Aussi, vous nous feriez plaisir en nous faisant part des erreurs et en nous communiquant les mots qui en seraient absents.

Ma première remarque porte sur des cas d'incohérence qui, tout bien considéré, nuisent à l'harmonie de la forme générale du dictionnaire. Commençons par le titre même de *Nkôngamyakù Cilubà-Mfwàlànsa*¹¹ qui indique naturellement qu'il s'agit d'un dictionnaire bilingue mais à l'analyse, il s'avère que nous avons plutôt affaire à un dictionnaire quasi plurilingue, plus exactement quadrilingue, dans la mesure où, en lieu et place d'une seule langue cible, en l'occurrence le français, certaines entrées en nombre relativement élevé, ont des équivalents en kiswahili, en kiluba et parfois même en latin, s'agissant des noms scientifiques d'animaux ou de plantes. Qu'est-ce qui justifie le choix du kiswahili et du kiluba en plus du français annoncé dans le titre de l'ouvrage? Si, comme je le suppose, l'un des objectifs poursuivis dans cette démarche est pédagogique et vise l'apprentissage des langues concernées par les utilisateurs de ce dictionnaire, le choix du kiluba me paraît injustifié, le kikongo ou peut-être mieux le lingala qui tend de plus en plus à devenir la deuxième langue des villes de Kananga, Tshikapa, Ilebo et Mbuji-Mayi ou même une langue africaine transfrontalière auraient, à mon avis, davantage servi cette cause.

D'autres cas d'incohérence ou de manque de systématisation concernent notamment:

- les paramètres et les justifications qui accompagnent les différentes entrées. En effet, l'auteur fait, sans nul doute, œuvre utile en recourant à des idiotismes, des phraséologismes, des proverbes et des exemples d'emploi pour faire mieux comprendre chaque vedette auprès des lecteurs. Cependant, il faut regretter que certaines de ces informations soient traduites en français et d'autres pas, ce qui fait preuve de manque de

rigueur et ne permet guère d'atteindre, dans l'ensemble, l'objectif recherché. Des exemples de ce type sont nombreux, j'en relèverai seulement quelques-uns liés aux vedettes qui suivent et que je donne simplement à titre illustratif: *bidya* (p. 20), *-cibukà* (p. 46), *cidibudibu* (p. 47), *ciyòdyò* (p. 71), *lubombu* (p. 161), *lukàsu* (p. 163), *-lòmba* (p. 158), *lufu* (p. 162), *kasuuyi* (p. 129), *meeji*, *mènu*, *mèsu* (p. 188), *matanga* (p. 183), etc. De même, en ce qui concerne les proverbes et d'autres locutions de cette nature à double signification, référentielle et connotée, la règle devrait être la même pour tous les cas, à savoir donner le sens littéral d'abord et le sens figuré ensuite, ce qui faciliterait leur meilleure compréhension. Cette règle n'est malheureusement pas toujours observée! Ainsi par ex. si ce principe s'applique aux entrées *kabùlukù*, *kabundi* (p. 119), *katèndè* (p. 134), *-lela* (p. 155), *-lenga* (p. 157), *mukànyà* (p. 204), *mu-* (p. 203), *nshìngù* (p. 2554), *mukàndu* (p. 204), *mukishi* (p. 205), en revanche, elle n'est guère observée dans le cas des adresses ci-après: *mukàjì* (p. 204), *kantu* (p. 128), *kanza* (p. 129), *nshìndì* (p. 254), *mudimu* (p. 201), *mutumba* (p. 222), *mukalenge* (p. 204), *-nanga* (p. 233), *ngooyi-mwàna* (p. 243), *-sèla* (295), *-songa* (p. 303), etc.

- l'utilisation d'une orthographe différente pour les mêmes mots comme dans *nsoya* (p. 255)/*nsooya* (p. 227) ou les mêmes sons distinctifs, singulièrement pour l'affriquée palatale *c* suivie d'une voyelle: *ca* (pp. 43, 190)/*cya* (pp. 72-74, 221), *cyâwù* (p. 74), *mucààwùdì* (p. 200), *ce* (pp. 43, 131)/*cye* (pp. 74-76), *co* (pp. 71, 72)/*cyo* (pp. 7, 235), *cu* (pp. 20, 235)/*cyu* (pp. 47, 76-77), etc.
- l'intégration sélective, dans un dictionnaire de langue, de noms de personnes et d'entités politico-administratives. Pour prendre un exemple, si *Cilenge* qui est un chef-lieu de district y figure, il devrait en être de même de Kabinda, Tshimbulu, Mwene-Ditu, Luebo et Ilebo, cité portuaire importante sur la rivière Kasai. Un tel choix ne se justifie guère, il donne au contraire une image tronquée des domaines anthroponymique et toponymique lubà qui devraient plutôt faire l'objet de dictionnaires spécialisés.

S'agissant toujours de l'orthographe, notons que la transcription des nasales vélaire *ŋ* et palatale en position initiale par le redoublement de la consonne **n** me semble peu pratique.¹² Le maintien de l'orthographe traditionnelle aurait été préférable pour deux raisons: en cilubà la nasale vélaire *ŋ* transcrite par le signe **ng** fonctionne comme un phonème au contraire du signe semblable en lingala ou en kikongo où il représente le complexe nasal **n + g**. Il n'est donc pas indiqué de la noter avec le redoublement de la consonne **n**, ni en initiale, ni en intervocalique. Quant à la palatale, elle devrait être notée **ny**, suivant la coutume en vigueur, notamment en position initiale où elle est toujours nasalisée et par conséquent neutralisée, ce qui n'est pas le cas en position intervocalique, sauf dans les formes locatives telles que *pannyimà*, *kunnyimà*, *munnyimà* et dans

les termes de parenté de classe 1 comme *baannyìsu* (vos pères), *baannyoku* (vos mères), *bannyìnkà* (nos grands-parents). En plus, le redoublement pose des problèmes pratiques en cas de notation de l'indice prédicatif comme dans *nngoma* (c'est un tambour) à la place de *nnngoma*, qui, sans conteste, a l'air rébarbatif.

Dans la 3ème partie du dictionnaire consacrée à l'esquisse grammaticale, l'auteur énonce quelques règles orthographiques relatives à la réalisation des voyelles et à la neutralisation de la longueur vocalique. En ce qui concerne la réalisation des voyelles, il y a entre autres la règle de coalescence ou la fusion des voyelles **u** et **o** dont la résultante devrait, à mon avis, s'écrire **oo** et non **wo** qui semble être d'ordre plutôt morphologique alors que l'orthographe des langues congolaises, par définition phonologique, est censée représenter les sons et non les sens. Ainsi donc, la voyelle longue **oo** devrait être adoptée à la place de **wo** dans les contextes ci-après:

- les substantifs des classes 3 et 11 dont le thème nominal commence par la voyelle **o**, les formes pronominales correspondantes ainsi que les adverbes. Ex.: *mu-ole* > *moole* <> *mwole*, *mu-onji* > *monji* <> *mwonji*, *mu-omu* > *mômu* <> *mwômu*, *mu-osa* > *môsà* <> *mwôsà*, *mu-oyo* > *mooyo* <> *mwoyo* (p. 232); *lu-onji* > *lonji* <> *lwonji*, *lu-ondapu* > *londapu* <> *lwondapu*, *lu-olu* > *lôlu* <> *lwôlu*, *lu-onsu* > *lônsu* <> *lwônsu*, *ku-oku* > *kôku* <> *kwôku*, *ku-onsu* > *kônsu* <> *kwônsu* (pp. 150-151); *mu-omu* > *mômu* <> *mwômu*, *mu-omumwe* > *moomùmwe* <> *mwomùmwe* (p. 231); etc.
- les infinitifs dont le radical a comme voyelle initiale **o**. Ex.: *ku-olola* > *koolola* <> *kwolola*, *ku-onda* > *konda* <> *kwonda*, *ku-ona* > *koona* <> *kwona*, *ku-ousha* > *koousha* <> *kwousha*, etc. (pp. 266, 267).¹³

En ce qui concerne la neutralisation de la longueur vocalique, il y a lieu d'y ajouter deux autres cas, en plus des trois déjà mentionnés par l'auteur, soit:

- lorsque la voyelle porte un *ton* complexe qui n'affecte que les voyelles longues. Ex.: *baàna* > *bâna* (p. 15), *boòbo* > *bôbo* (p. 23), *maàyi* > *mâyi* (p. 184), *meème* > *même* (p. 188), etc.
- le connectif. Ex.: *maalu à(à) mwakulu* (p.177), *kashèètè kà(à) kabàdi* (p. 132), *màzuwà à(à) mâyi*, *bânà bà(à) Musàwu* (p. 11), etc.

L'examen rapide de la *wordlist* fait ressortir quelques insuffisances qu'il convient de signaler. Elles concernent notamment la définition ou le sens donné à certaines entrées qui, par endroits, est soit incomplet, soit inexact. La liste qui suit, non exhaustive est une indication à l'auteur, selon son propre vœu. Il s'agit de:

- *bulubà* (p. 31), le mot désigne également la variante du *cilubà* parlée au Kasai Oriental, aussi appelée *cyena mbùji-mâyi*, opposé à *bupèmbà*, *cyena luluwà* ou *cyena kanàngà* (Maalu-Bungi 1991: 185; Kayembe 1987: 10);

- *bààbèbè, bààbènù, bààbèndè et bààbètù* (p. 12), les deux premiers termes n'existent pas en cilubà comme le dernier et n'ont pas les sens respectifs que l'auteur leur attribue à savoir 'ta mère', 'votre mère', 'sa mère' et 'notre mère' dont les vrais équivalents, dans la variante luluwà du cilubà sont *nyoku, baannyoku, nyinândi et baabààbà*, considérés aujourd'hui, par rapport au standard, comme dialectaux. Par contre *bààbànyì et bààbètù*, suivis ou non de *bààbà/bààb'èè*, sont des expressions exclamatives traduisant l'étonnement, la douleur ou le regret. Pour sa part, *bààbèndè* est une forme possessive qui signifie 'à autrui', sous entendu 'personnes humaines', hommes ou femmes;
- *cibungu* (p. 47), son autre sens est 'gros livre'. Ex.: *Cibungu cyà mèyi nè mikàndu*, code;
- *cicò* (p. 47), le deuxième sens est celui de 'refus d'obtempérer';
- *cìdìdì* (p. 48), le mot désigne originellement 'mensonge' et par élargissement sémantique 'politique', du fait du comportement des premiers hommes politiques du pays dits '*beena cìdìdì*', donc des menteurs, vendeurs d'illusions;
- *cilelelu* (p. 55), ses autres sens sont: 'maternité, au sens de lieu où les femmes accouchent', 'naissance'. *Ku cilelelu*, à la naissance; *disàmà dyà ku cilelelu*, maladie congénitale; *dîná dyà ku cilelelu/dyà munda*, nom de naissance, singulièrement des enfants spéciaux (*Ngalamulùme*, né après trois filles successives; *Ngalula/Cyàndà*, née après trois garçons successifs; *Mbùùyi/Cibwàbwà, Kânkù/Kabangà*, donnés aux 1er et 2ème jumeaux; *Citùùkà*, prématuré; etc.);
- *cilùmbààyi* (p. 56), ne s'emploie que pour les hommes, donc le traduire par 'héroïne' est inapproprié;
- *cilùmbu* (p. 56), désigne 'devineresse'. Ex.: *Misenga cilùmbu*, Misenga la devineresse;
- *cipààpa* (p. 63), désigne 'action de donner sans compter', d'où le nom de *cipààpaayi*, absent dans la *wordlist* et employé toujours avec une pointe d'ironie pour désigner 'personne trop généreuse';
- *cipimu* (p. 64), autre sens 'pesée', 'balance' en langage médical. *Kwela ku cipimu*, peser; *kuya nè mwâna ku cipimu*, amener l'enfant à la pesée;
- *cisemuna* (p. 66): 'gros véhicule', 'camion', le sens de 'moyen de transport' est inconnu;
- *citeetela* (p. 69), s'emploie également comme glossonyme hétéronyme de la langue des Otetela;
- *citù* (p. 69), son sens général est 'gaffe', 'erreur grave'. *Kudyà citù*, commettre une gaffe, une erreur grave qui fait scandale;

- *cyàbôyà* (p.72), désigne également 'personne très poilue', particulièrement au torse pour les hommes et dans la partie inférieure du ventre pour les femmes;
- *cyambula* (p. 73), 'véhicule de transport de personnes ou de biens'. *Cyambula babèèdi*, ambulance; *cyambula bafwè*, corbillard; *cyambula makonde*, train marchandise; etc.;
- *cyòmbelà*, absent du dictionnaire, désigne 'tambourinaire';
- *-cyèlewa* (p. 43): 1er sens: marquer du retard en parlant de moyens de transport de masse (train, avion, bus, etc); 2ème sens: rater ce moyen pour arrivée tardive au lieu de départ; 3ème sens (argot): coiffer la sainte Cathérine; *Cicyèlèwe*, fille ayant dépassé l'âge de mariage;
- *cyanza* (p.74), le sens référentiel de *kwela cyanza mukana* est 'manger', celui de 'provoquer' est un sens figuré. Le mot s'emploie aussi, en langue de spécialité, pour 'écriture'. *Cyanza cîmpè*, belle écriture; *kwikala nè cyanza*, avoir la main (littéralement), se dit d'une personne dont tout ce qu'il plante pousse bien;
- *Dibòko* (p. 79), 2ème sens: 'direction', 'voie'. (*Dibòkù*) *dyà balùme*, *dyà bakàji*, à droite, à gauche;
- *Dikàya* (p. 79), 2ème sens: 'action de faire des cérémonies de protection des champs contre les voleurs et dont les conséquences sont généralement graves' (mort, handicap physique du coupable, etc.); 3ème sens: 'action de proférer des insultes indécentes ayant pour cible la mère de l'interlocuteur';
- *kayanda* (p. 135), juron d'origine cyokwe, utilisé comme défi en souvenir des souffrances endurées au rite d'initiation dit précisément 'kayanda', dont la circoncision. *Kayanda kàànyì*, *cyèna ndyà bidyà byèbè*, je jure (par la souffrance endurée au *kayanda*) que je ne mangerai point ta nourriture; *Ne udi mulùme kayanda kèèbè*, montre de quoi tu es capable, si tu es un homme; etc. *Kayanda* est également un anthroponyme;
- *mâyi* (p.184), autre sens, 'sperme' (*mây' à balùme*);
- *Malandji/Malandi* (p. 180), déformation lubà de *Malange*, ville d'origine des Imbangala dits 'Bimbadi', porteurs angolais de la suite des explorateurs allemands Podge et von Wissman. Il désigne le premier poste d'Etat fondé au Kasai en 1884. Site historique situé à environ 12 km du chef — lieu de la province du Kasai Occidental, il est devenu un lieu de pèlerinage des chrétiens catholiques. *Kananga Malandji wa nshinga*, nom de louange de la ville de Kananga, signifie 'Kananga aux fils électriques innombrables';

- *mbondya* (p. 186), *kwela mbondya muulu (ne ...)*, se livrer à l'acte sexuel (avec ...), sens généralement péjoratif; *kutàkisha mbondya*, être porté à l'acte sexuel;
- *mfùnda* (p. 189), sans doute un néologisme mais *mufùndi*, déjà en usage (Maalu-Bungi 2004: 23; Beya 2006; et d'autres) conviendrait certainement mieux. De la même racine *-fund-a*, on a aussi *bufùndi*, 'art littéraire';
- *Mjùngu* (p. 190), pl. de *mujùngu*, ensemble enroulé d'entrailles de chèvre fait d'estomac et d'intestins qui se mange cuit ou grillé, très prisé au Kasai;
- *mpàla* (p. 194), *mufwè-mpàla*, laid de figure;
- *mpeelù* (p. 195), du français 'appel'. L'expression *kwela mpeelù*, faire l'appel s'oppose à *kwitaba mpeelù*, répondre à l'appel;
- *mufùndu*, absent de la *wordlist* signifie 'trace', 'empreinte'. *Mufùndù wà nyòka*, trace de serpent; *mufùndù wà màshinyi*, trace de véhicule, *mufùndù wà dikalù*, trace de vélo. Par néologie sémantique, il désigne 'écrit' (Kabuta 2008: VIII), 'lettre', d'où le terme *mifùndu milenga*, littérature écrite (Maalu-Bungi 2004: 17);
- *mukààyà* (p. 203), autre sens: *kudyèla/kwela mukààyà*, pratiquer l'abstinence sexuelle;
- *mukalenge* (p. 204), le 1er sens est 'chef', dont l'élargissement sémantique a donné le sens de 'homme de race blanche', synonyme de *mutòòke* (blanc, européen);
- *mupongu* (p. 215), *kupyà mupongu*, 1er sens: attraper la sorcellerie par l'acte magique d'un ancien sorcier, le plus souvent une nourriture qui change de nature une fois dans la bouche pour être consommée; 2ème sens: être émerveillé par un acte jugé extraordinaire;
- *mushìku* (p. 218), *mwenamushìku*, personne à l'injure facile;
- *mwakulu* (p. 226), le terme désigne, en langue de spécialité, 'grammaire';
- *mwanya* (p. 228), la 2ème définition de 'vide laissé par une dent extraite' donnée par l'auteur correspond plutôt à *mubòla*, non repris dans la liste des mots ou *mulàpù* (p. 200);
- *mwânyà/mwanyà* (p. 228), légume à vertu thérapeutique donnée à la femme stérile pour lui permettre de concevoir. *Dìlòmbolombò mmwanyà wà kupà nkumba wàdya*, *nkumba wàbwela mu nzùbu wàlela mwànà*, le *dìlòmbolombò* est le légume à donner à la femme stérile, qu'elle le mange, couche avec son mari et conçoit;
- *nanga* (p. 234), faire du mal à quelqu'un pour le corriger à la suite d'un acte répréhensible;

- *njila* (p. 244), les significations suivantes doivent être ajoutées:
 - *njila munène, njila wa màshinyi*, grande route, route pour automobiles, opposé à sentier, piste;
 - permission, autorisation. *Kulòmba njila*, demander l'autorisation; *ku-pèèsha njila*, autoriser, donner, accorder l'autorisation; *mukàndà wà njila*, ordre de mission, autorisation de sortie;
- *nsàlà* (p. 252), 1er sens: 'plume d'oiseau'; 2ème sens: 'distinction', qui a donné *Nsàlà wa bufundi*, Prix littéraire (Maalu-Bungi 2004: 5; Beya 2006: 7);
- *nnyoku* (p. 250), l'équivalent français est 'ta mère' et non 'ma mère' (*mààmù*), son synonyme est '*mamwèbè*';
- *punga* (p. 284), réunir (*kupunga makàndà*, réunir ses forces, se liguer);
- *sàtànà* (p. 293), 2ème sens: 'tentateur', plus exactement le séducteur qui par son langage ou son comportement (gestes, regards) veut amener l'autre à commettre l'impudicité (Maalu-Bungi et Kapudi 1992: 267).¹⁴

La grammaire abrégée du cilubà, une sorte d'appendice qui, en termes d'enrichissement des langues africaines, ouvre certainement la voie à l'instrumentalisation de cette langue. Et si, comme d'aucuns le savent, la modernisation des langues africaines passe inmanquablement par leur dotation de terminologies des domaines prioritaires du savoir scientifique et technique, la création de nouveaux termes doit être basée sur les principes et méthodes en vigueur dans cette discipline relativement récente. C'est ce que Kabuta s'est proposé de faire, avec plus ou moins de bonheur, dans le souci à la fois de corriger, compléter et améliorer le travail réalisé voilà plus d'un demi-siècle par les missionnaires, généralement sans formation linguistique adéquate.¹⁵ Toutefois, il faut regretter que certaines règles n'aient pas été respectées, comme par ex. le principe d'ordre de toute recherche terminologique à savoir un concept, un terme et la prise en compte des usages existants. Commençons par la monosémie, règle fondamentale en terminologie pour relever que certaines notions sont désignées par au moins deux termes différents à l'exemple de:

- alphabet: *lùfwàbeetà* (p. 163), *bukwà màleetà* (p. 29);¹⁶
- ton: *dîyi* (pp. 347, 348), sens du reste non repris dans le dictionnaire, *lutèndu* (p. 348) et *cikùmà* (p. 349);
- source: *mfùki* (p. 360), *mpokolo* (p. 196);
- classe nominale/genre: *cifùku* (pp. 49, 348), *mulongo* (p. 210).

S'agissant de la règle qui recommande de tenir compte des usages existants, il y a, à mon avis, des termes dont le remplacement n'est pas justifié, puisqu'ils

fonctionnent sans ambiguïté, tels par exemple:

- les noms des disciplines scientifiques, rendus traditionnellement par *maalu a*, suivi du nom du domaine en question, comme l'auteur le fait aussi lui-même pour 'linguistique' ou 'science du langage' et 'sciences religieuses' (*maalu à mwakulu*, *maalu à Nzambi*, p. 177) sans pourtant généraliser le principe. Par conséquent, à la p. III, il serait préférable d'avoir les termes suivants:
 - hydrologie: *maalu à bukwa misùlu* <> *bushikuluji*;
 - sociologie: *maalu à nsòmbelu* <> *bushikulankònga*;
 - géographie: *maalu à buloba*, déjà en usage depuis près d'un siècle <> *bushikulaba*;
 - botanique: *maalu à bukwa micì* <> *bushikulamicì*;
 - économie: *maalu à bubanji* <> *bubanji*;etc.

Il en est de même de plusieurs autres termes dont le remplacement ne se justifie guère. Il s'agit notamment de:

- *dìleetà dijikùke* <> *kajùkùke* (voyelle);
- *dìleetà dicìntàmàne* <> *kacintàmàne* (consonne);
- *mwaku wà pa mwandà / mwada pa wôwo*, utilisé du reste ailleurs par l'auteur lui-même (p. 227) <> *lwakù*;
- *muleeji ngiikàdilù* <> *mufileedi* (adjectif qualificatif);
- *mubidi wà mwakù* <> *dishinà dyà mwakù* (thème nominal);
- *cibàsa, dibaya* <> *cyâta* (tableau);
- *kashipeelu* <> *mubilu* (apostrophe);
- *mutù wà cibìkidilu* <> *mufileedi* (préfixe nominal);
- *bulongeshi* <> *bundongesha* (enseignement);
- *kàsai* <> *kàsaayi*, variante abusivement généralisée. Le standard en usage devrait être maintenu car la quantité vocalique n'est pas représentative de l'ensemble des parlers de cette langue au Kasai Oriental et encore moins au Kasai Occidental. Du reste, cette longueur vocalique disparaît automatiquement et reflète donc la réalité du terme quand il s'agit de désigner en français les ressortissants de cette contrée géographique: *kasaiens* et non *kasaiiens*. Perpétuer cette façon d'écrire pêche contre le principe de toute standardisation à savoir qu'on ne peut laisser chacun écrire comme il parle, ce qui serait une manière de légitimer la représentation par écrit de chaque dialecte ou de chaque parler, très diversifiés en ce qui concerne cette langue au Kasai Occidental comme au Kasai Oriental;

- *buumùngàngà* <> *bungàngà*;
- *mufwànu* <> *mufwàna* (synonyme), utilisé dans tous les manuels de ciluba, de la 1ère à la 6ème année primaire, dans chaque leçon. Cette habitude ne devrait pas être brisée.

D'autres néologismes posent problème quant à leur aptitude sémantique et leur pertinence onomasiologique, tels:

- *bijaanu*: critique, le mot en usage est *dilondolola*, *dikonkonona* (Kabongo et Bilolo 1994: 206; Maalu-Bungi 2004: 16), ce mot ayant un sens généralement négatif;
- *kanungu*: article de dictionnaire, à distinguer de *mwakù*, entrée ou dette;
- *kaabukùlù kaakula*: tradition orale, le meilleur terme est *kaabukùlù*, par essence oral en Afrique, *kaakula* (parlé) devenant de ce fait tautologique;
- *mifundu myakula nè mifunda*: documents oraux et écrits, le meilleur terme doit est *maalu mamba nè mifundu* (littéralement, 'faits dits', 'racontés' et 'écrits', ce dernier terme étant le substantif du verbe *kufunda* 'écrire');
- *mfündamyakù*: lexicographie, le meilleur terme doit être *difùndà* (*dyà*) *nkòngamyakù* (élaboration de dictionnaires) et *maalu à difùnda nkòngamyakù* (discipline scientifique);
- *bumfùnda*: littérature écrite, le terme en usage est plutôt *mifundu milenga* ('écrits agréables', littéralement).

Ici s'achève cette évaluation critique de *Nkòngamyakù Cilubà-Mfwàlansa*, vivement souhaitée par son auteur. Sa valeur reste globalement indéniable en dépit des remarques qui viennent d'être formulées et dont le but est de contribuer, tant soit peu, à l'amélioration de ses futures éditions.

Notes

1. La 2ème édition de l'*Atlas linguistique de la RD Congo*, sous presse au Centre Régional de Recherche et de Documentation sur les Traditions Orales et pour le Développement des Langues Africaines (CERDOTOLA) à Yaoundé, donne les chiffres suivants arrêtés en 2008: 103 ouvrages et articles, 146 thèses, mémoires et travaux de fin d'études.
2. Il s'agit de Von Wissmann, H., 1888, *Im Innern Afrikas: die Erforschung des Kassai während der Jahre 1883, 1884 und 1885*. Verf. H. von Wissmann, L. Wolf, C. von François, H. Müller, Leipzig.
3. Pseudonyme composé de *Ngo*, abréviation de *ngèlènji*, longues plumes qui poussent à la queue de l'engoulevent, vers la fin du mois d'août et qui présagent le retour de la saison des pluies. *Ngèlènji* se postpose habituellement à *lubuta/kabuta* (diminutif) et donne, en termes de

nom d'éloge, *lubuta/kabuta ngèlènji*, engoulevant à la longue queue. Quant à *Semzara*, c'est un anthroponyme tanzanien que Kabuta s'est donné en souvenir du swahili, sa langue d'enfance.

4. Outre le lexique publié dans le livre de l'explorateur allemand H. von Wissmann, nous avons comme ouvrages lexicographiques de cette langue, entre autres:
Vocabulaire français-luba, 68 p. dans De Clercq, A., 1897, *Grammaire de la langue des Bena Lulua*, Bruxelles; Vocabulaire luba-français, 164 p. et Vocabulaire français-luba, 200 p. dans De Clercq, A., 1903, *Grammaire de la langue luba*, Louvain; De Clercq, A., 1910, *Petit vocabulaire Français-Tshiluba*, Hemptinne; Morriison, W. M., 1910, *Grammar and Dictionary of the Baluba-Lulua Language as Spoken in the Upper Kasai and Congo Basin*, New York, 417 p.; De Clercq, A., 1914, *Dictionnaire Français-Luba, Luba-Français*, Bruxelles; Gabriel, F., 1921, *Dictionnaire Français-Tshiluba*, Bruxelles, 474 p.; Kinnon, A.C., 1929, *Dictionnaire français-buluba et buluba-français*, Luebo, 208 p.; De Clercq, A., 1937, *Dictionnaire Luba-Français, Français-Luba*, Léopoldville, 598 p. (2ème édition); De Clercq, A., 1960, *Dictionnaire Tshiluba-Français*, Nouvelle édition revue et augmentée par E. Willems, Léopoldville, 392 p.; Willems, E., 1986, *Dictionnaire Français-Luba*, Éd. de l'Archidiocèse, Kananga; Willems, E., 2006, *Dictionnaire Français-Tshiluba*, Éd. de l'Archidiocèse, Kananga, 4ème éd. Revue, corrigée et actualisée avec la collaboration de Fr. Pamba Kamba Kamba et M. Lutumbu Kapata, 413 p.; Nzongola, K., 1965, *Dictionnaire des synonymes Tshiluba*, Improka, Kananga.
5. Cf. p. IX.
6. *Orthographe pratique des langues africaines, Nouvelle impression de l'édition de 1930 publiée avec la collaboration de l'Agence de Coopération Culturelle et Technique*, Institut International des Langues et Civilisations Africaines, Paris, 1980, 23 p.
7. Quoique utilisé bien longtemps auparavant par les spécialistes du cilubà tels que Coupeuz, De Rop, Stappers, Maalu-Bungi, Mufuta, etc. et vulgarisé à travers les manuels de cette langue élaborés par le Centre de linguistique théorique et appliquée "Celta", ce signe est encore l'objet de contestation dans les milieux ecclésiastiques catholiques et protestants. Il apparaît cependant dans certains anthroponymes kasaiens depuis les années soixante-dix. Cf. Stappers 1962, De Rop 1963, Coupeuz 1953, Maalu-Bungi 1974, 1991, 2004, Maalu-Bungi et Kaputi 1992, et Mufuta 1968.
8. Bien que la première école du Kasaï ait ouvert ses portes en 1893 à Mikalayi, quelques manuels et livres religieux avaient déjà été imprimés après la *Grammaire de la langue des Bena Lulua* d'A. De Clercq en 1897, notamment *Malu a mu mikanda minsanto ya Mvidi Mukulu*, Bruxelles, 1898, *Mukanda wa kubadisha bana*, Bruxelles, 1898, *Miaku misunguluke*, Bruxelles, 1900, *Mukanda wa kusambila*, Bruxelles, 1905, etc. Quant à la littérature écrite, elle débute par des poèmes ressortissant à la *creative writing* publiés dans Nkuruse en 1914, année de création de ce périodique catholique.
9. Après avoir séjourné dans ce pays en 1958 à l'occasion de l'Exposition Universelle comme membre de la Chorale d'enfants de Kamina "Les Troubadours du Roi Baudouin", puis en 1960 en vacances dans une famille flamande, il y retournera définitivement en 1962 pour y vivre jusqu'à ce jour où, à l'issue de ses études universitaires et son doctorat, il enseigne la linguistique et la littérature africaines à l'université de Gand depuis 1999.
10. *Nànsha mùdì nkòngamyakù ewu mwikàle kumpàlà kwà nkòngamyakù yà myakulu yà mu Afrikà inùdì mwà kwikala biibidilàngàne naayi, ùcìdì pèndè mwàna mutekète, nè mapanga àà bupyankùnde.*

Kādi kabambilepù nè: "Nkùndè yà bāngì ibobèle nè matè?" Ke bwalu kaayi tūdi batèkèmene bijaanu nè meeji mapà byà kunùdì bwà kumòna mwà kwakaja cyàmudimu eci, cidì mwà kwambulwisha bafūndi, bakūdimunyi, balongi, bakèbuludi, bafūki bàà myakù nè bakwàbò bônso bàdì cilubà cilengeleela; imwè myakù idì mu nkòngamyakù ewu mmifūka àmu bwalu bìdì bikèngèla kwamba maalu ônsò mu cilubà. Kādi babadi bàdì àmu mwà kufila yàbò nngènyi ku Ciyem bwà kumònacì mwà kusungula idì myakàne (...) Kādi nūmanyè nè: mu mudimu wà bwena ewu, kamwèna mwà kupangika bilèma nè mapanga makwàbò nānsha. Kabambilepù nè: 'Nsenda ùtu ùfudila ku mwelelu kwà njila, bwà bàdì bàpita bāmulejè bilèma bidīye wènsa? Ke kaayi tūdi twimīna bwà kupātula mukānda ewu mubanga kale, bu mùdīwu mpīndyewu. Nānku nudi mwà kutūsāmba pèènu pa kutūleeja bilèma nè kutūtūmīna myakù inūdi kanūyi basangānemū".

11. Ce mot, emprunt du français 'français', n'est malheureusement pas repris dans la *wordlist* du dictionnaire.
12. Relevé dans les mots tels que *nngaji*, *mngoma*, *nnyaanà*, *nnyama*, *nnyuunyu*, etc. (pp. 39, 61, 64, 249-251).
13. Seuls les milieux religieux catholiques et protestants continuent à utiliser cette graphie, abandonnée depuis par tous les linguistes spécialistes du cilubà.
14. À cette liste tirée d'un corpus à adresser à l'auteur, on peut ajouter plusieurs mots absents de la *wordlist*, notamment *cindānda* (guidare), *mbeesa* (internat, pensionnat), *dibālasà* (maison gigantesque), *kālèbefū* (cœur de bœuf), *kaludikilu* (règle), *cimbalakata* (vieux véhicule), *mājīnà* (homonyme), *dikaadikà* (wagon), *kàtèkemenyi* (cathécumène), *kāleelèki* (commis, employé de l'État), *dibungū* (port), *mùshikāngondò* (retraité), *mūcyūngà ntentū* (convoyeur), *cisèngà* (nourriture de base faite de farine de maïs), *mùlòdfū* (chômeur), *mbàlakà* (célibataire), *àlūmeetà* (allumette), *kālānde* (canal d'eau), *kākòkòlā* (coca-cola), *dikèlèmbà / dinkidimbà* (ristourne), *ngandanjiika* (cacahuète rouge), *būnkonde* (papaye), *ntukutuku* (motocyclette), *cinkālafū* (goyave), *mùshèèshè* (policier), *jāntè* (jante), *toolòshū* (torche), *kandida* (candidat), *foto* (photo), *avòka* (avocat), *kucyòka* (être fatigué), *mājāngì* (trahison), *mbānda* (beau-frère, mari de la belle-sœur, rival dont on se partage la concubine); *būdì bwēnde* (épidémie, sida); *kalòwu* (petite calebasse), *kushipa kalòwu* (faire un scandale), *mūnwà-māàyì* (buveur d'eau, diabétique); *nsaasa* (prêtre), *lāikè* (laïc); *jūrnālè* (journal), *ààbè / lāàbè* (abbé); etc.
15. La grammaire lubà s'enseigne depuis les années 1930 à travers six manuels allant de la 1ère à la 6ème année primaire sous le titre de *Mukanda wa Tshiluba. Kalasa ka kumpala*. 1, *kalasa kibidì* 2, *kalasa kisatu* 3, *kalasa kinayi* 4, *kalasa kitanu* 5, *kalasa kisambombo* 6. (Manuel de ciluba. 1ère année, 2ème année, 3ème année, 4ème année, 5ème année, 6ème année). La dernière édition de ces livres date de 1995. Une abondante terminologie grammaticale y est développée quoique laissant parfois à désirer eu égard au développement de la linguistique africaine moderne.
16. L'on trouve déjà, pour ce concept, le terme *alfābe*, un emprunt du français (Maalu-Bungi 2004: 15).

Bibliographie

- Beya Ngindu, B. 2006. *Bafundi betu ba ciluba*. Sudbury: Glopro.
 Coupeuz, A. 1953. *Etudes sur la langue luba*. Tervuren: MRAC.
 De Rop, A. 1963. *Introduction à la linguistique bantoue congolaise*. Bruxelles: Mimosa.

- Kabongo-Kanundowi, E. et Bilolo-Mubabinge.** 1994. *Conception bantu de l'autorité suivie de Baluba: Bumfumu ne Bulongolodi*. Munich/Kinshasa: Publications Universitaires Africaines.
- Kabuta, J.** 2009. *J'ai été Troubadour du Roi Baudouin*. Bruxelles: Dialogue des Peuples.
- Kayembe, N.** 1987. *Akula tshiluba tshimpe*. Kananga: Liproka.
- Kazadi, N., B. Kempf et H. Mutombo.** 1980. *Kubala nè kufinda cilubà 2*. Kinshasa: Mont-Noir/Nathan Afrique.
- Kiango, J.G.** 2000. *Bantu Lexicography: A Critical Survey of the Principles and Process of Constructing Dictionary Entries*. Tokyo: ILCAA, Tokyo University of Foreign Studies.
- Maalu-Bungi.** 1974. *Contes populaires du Kasai*. Kinshasa-Lubumbashi: Mont-Noir.
- Maalu-Bungi et K. Kapudi.** 1992. Toponymes et anthroponymes dans les versions luba des Écritures Saintes. Müller, E.W. et A.M. Brandstetter (Éds.). 1992. *Forschungen in Zaire: in Memoriam Erika Sulzmann (7.1.1911–17.6.1989)*: 259-268. Hamburg/Münster: LIT.
- Maalu-Bungi, C.** 2004. *Mwakulu wa ciluba. Leelu ne makeelela*. Sudbury: Glopro.
- Maalu-Bungi, L.L.** 1991. Langues zairoises et standardisation. Le cas du ciluba. Cyffer, N. et al. (Éds.). 1991. *Language Standardization in Africa*: 183-189. Hamburg: Helmut Buske.
- Mairo, K.A.** 1990. Historical Background, with Special Reference to Western Africa. Hartmann, R.R.K. (Éd.). 1990. *Lexicography in Africa: Progress Reports from the Dictionary Research Centre Workshop at Exeter, 24–25 March 1989*: 8-18. Exeter: University of Exeter Press.
- Mufuta, P.** 1968. *Le chant kasala des Luba*. Paris: Julliard.
- Stappers, L.** 1962. *Textes luba. Contes d'animaux*. Tervuren: MRAC.

Regard critique sur *On Est Ensemble: 852 mots pour comprendre le français du Gabon**

Pierre Ondo-Mébiame (*pierreondo@yahoo.fr*)

et

Guy-Modeste Ekwa Ebanéga (*guymodeste_e@yahoo.fr*)

Département des Sciences du Langage, Université Omar Bongo, Libreville,
Gabon

Résumé: L'ouvrage intitulé *On Est Ensemble: 852 mots pour comprendre le français du Gabon* (L. Ditougou, 2009) constitue une contribution importante dans l'histoire du français parlé au Gabon. Cette œuvre présente cependant quelques insuffisances sur le plan métalexographique. Celles-ci portent sur les pré-textes, la nomenclature (macrostructure, microstructure, structure d'accès, structure d'adressage et médiostucture) et les post-textes. Cet article vise donc à relever ces insuffisances et, concomitamment, à proposer les voies qui pourraient aider à les améliorer.

Mots clés: DICTIONNAIRE, LEXICOGRAPHIE, PRE-TEXTES, NOMENCLATURE, MACROSTRUCTURE, MICROSTRUCTURE, STRUCTURE D'ACCÈS, STRUCTURE D'ADRESSAGE, MEDIOSTRUCTURE, POST-TEXTES

Abstract: *A Critical Look at On Est Ensemble: 852 mots pour comprendre le français du Gabon.* The book *On Est Ensemble: 852 mots pour comprendre le français du Gabon* (L. Ditougou, 2009) is an important contribution to the history of French spoken in Gabon. However, this work presents some shortcomings from a metalexographic point of view. These relate to the front matter text, the central list (the macrostructure, the microstructure, the access structure, the addressing structure and the mediostucture) and the back matter text. This article aims to address these shortcomings, and, at the same time, to propose ways that could help improve them.

Keywords: DICTIONARY, LEXICOGRAPHY, FRONT MATTER TEXTS, CENTRAL LIST, MACROSTRUCTURE, MICROSTRUCTURE, ACCESS STRUCTURE, ADDRESSING STRUCTURE, MEDIOSTRUCTURE, BACK MATTER TEXTS

1. Définition des notions

Les aspects qui n'ont pas fait l'objet d'un traitement suffisant dans l'œuvre de Ditougou sont: les pré-textes, la nomenclature, la macrostructure, la micro-

* Lucien Ditougou. *On Est Ensemble: 852 mots pour comprendre le français du Gabon*. Préface par Guy Rossatanga-Rignault. 2009, 154 pp. ISBN 978-2-912776-85-3. Hommes et sociétés. Libreville: Éditions Raponda Walker.

structure, la structure d'accès, la structure d'adressage, la médiostucture et les post-textes. Ces termes se définissent comme suit:

(a) Pré-textes

Les pré-textes constituent la première des trois principales composantes d'un dictionnaire. On y trouve généralement un texte obligatoire: le guide aux usagers dans lequel le lexicographe donne la ligne éditoriale de son dictionnaire (explicitation de la structure du dictionnaire) et un ensemble de données pour une utilisation optimale de ce dernier (où et comment accéder aux données?). On peut également y trouver la préface et l'avant-propos.

(b) Nomenclature

Deuxième composante du dictionnaire, la nomenclature consiste en des séries d'articles et chaque série comprend une variété d'articles fonctionnant comme des textes autonomes (cf. Gouws 1999: 41). La nomenclature est également appelée « liste centrale » en ce sens qu'elle constitue la principale composante du dictionnaire. C'est à l'intérieur de celle-ci qu'a lieu le traitement lexicographique.

(c) Macrostructure

La macrostructure est la principale composante de la nomenclature. Elle regroupe l'ensemble des items lexicaux retenus comme lemmes¹ du dictionnaire. Généralement présentée en ordre alphabétique, la macrostructure peut, dans certains cas, avoir une organisation thématique, sémantique, par nidification² ou encore par nichification.³ Il faut noter qu'il peut exister dans un dictionnaire, deux types de macrostructures:

- (i) une macrostructure primaire, qui est constituée de tous les lemmes;
- (ii) une macrostructure secondaire, qui regroupe tous les sous-lemmes.

(d) Microstructure

La microstructure est la somme totale, ordonnée de manière linéaire, des types d'informations apparaissant à la suite du lemme (Hausmann et Wiegand 1989: 340). Elle constitue la composante structurale de la nomenclature. Un élément macrostructurel combiné avec son traitement microstructurel constitue un article de dictionnaire⁴ (Gouws 2001: 122).

(e) Structure d'accès

La structure d'accès est l'itinéraire suivi par un usager pour atteindre un signe-lemme spécifique ou une catégorie de données (Gouws 1999: 42). Une distinc-

tion est faite entre la structure d'accès externe et la structure d'accès interne.

La structure d'accès externe est l'itinéraire de recherche guidant l'utilisateur vers le signe-lemme approprié. Elle lui permet donc à partir des entrées portées sur la couverture, d'accéder au signe-lemme.

La structure d'accès interne est l'itinéraire de recherche à l'intérieur de l'article guidant l'utilisateur vers une donnée ou entrée spécifique (Gouws 2001: 123).

(f) Structure d'adressage

La structure d'adressage est la manière dont une forme et l'information relative à cette forme sont coordonnées. Une distinction est faite entre structure d'adressage lemmatique et structure d'adressage non lemmatique.

Lorsqu'une entrée particulière est adressée au lemme, on parle de procédé d'adressage lemmatique. Tandis que tout procédé d'adressage à l'intérieur des éléments microstructurels est du domaine de la structure d'adressage non lemmatique (Hausmann et Wiegand 1989: 328-329).

(g) Médiostructure

La médiostructure est le système de renvois qui guide l'utilisateur à partir d'une position de référence vers une adresse. Une distinction est faite entre renvoi interne et renvoi externe.

Dans un renvoi interne, le lexicographe guide l'utilisateur vers une autre donnée à l'intérieur de l'article traité afin d'y trouver des informations supplémentaires. Par contre, dans un renvoi externe, l'utilisateur est orienté vers un autre article, ou alors, vers un autre texte du dictionnaire (pré-textes et post-textes) où il est susceptible de trouver un complément d'informations (Gouws 1999: 43).

(h) Post-textes

Les post-textes constituent la dernière des trois principales sections du dictionnaire. C'est dans celle-ci que les rédacteurs du dictionnaire indiquent les données additives telles que les abréviations, les symboles, les unités de mesure et les noms géographiques. Les post-textes peuvent également contenir les textes optionnels tels que des éléments de grammaire ou des notes sur la prononciation.

2. Présentation de l'ouvrage

On Est Ensemble, que nous abrègerons OEE, est un ouvrage produit pour illustrer la particularité du français parlé au Gabon, un pays caractérisé par une situation diglossique (français et langues africaines) dans les centres urbains, et

par son hétérogénéité linguistique. Ainsi, comme le montre Rossatanga dans la préface de l'ouvrage, se nourrissant des apports externes (langues européennes et africaines) et du génie des différentes langues nationales, ce parler particulier du français, tout en étant "souché" sur le français dit de France, est une langue propre aux Gabonais et à ceux qui habitent au Gabon. C'est ce qui fait son intérêt.

Il s'agit d'une œuvre lexicographique monolingue (cf. français) de Lucien Ditougou, enseignant-chercheur au Département de Littératures africaines de l'Université Omar Bongo. Comme nombre d'autres avant lui, il a une méconnaissance de la métalexigraphie⁵ et de la lexicographie⁶ en tant que disciplines scientifiques.

Par la publication de OEE, Ditougou se démarque donc de son domaine de prédilection, à savoir: la littérature. Cela justifie le peu de considération métalexigraphique de son œuvre. Ndinga-Koumba-Binza (2006: 298-299) nous montre, au demeurant que s'il est vrai que la pratique des dictionnaires préexiste à la métalexigraphie (Gouws 1999: 15-17 et 23-34; et Mihindou 2001: 7), il est aussi vrai que les pionniers dans la production dictionnaire n'ont pas toujours la formation théorique adéquate. Nous devons donc comprendre que la profession et la formation de l'auteur aient influencé la macrostructure et la microstructure de son ouvrage. Celui-ci a bénéficié du concours de la Fondation Raponda Walker dans la collecte des données.

Il peut être subdivisé en trois parties:

- (1) les pré-textes;
- (2) la nomenclature; et
- (3) les post-textes.

3. Les pré-textes de l'ouvrage

Les pré-textes de OEE sont organisés comme suit:

- les quatre premières pages (y compris la page de couverture) comportent: le nom de l'auteur, un dessin illustrant une exécution du parler français du Gabon, le titre de l'ouvrage suivi de "852 mots pour comprendre le français du Gabon", l'auteur de l'illustration de couverture (Lybek), le nom de l'éditeur (Éditions Raponda Walker), son adresse et le lieu de l'édition (Libreville, Gabon);
- les pages 5 et 6 elles, comprennent: la préface de Guy-Rossatanga-Rignault, le Président de la Fondation Raponda Walker. Ce dernier situe l'ouvrage dans son contexte sociolinguistique. Le titre "On Est Ensemble", reprend-il à Ditougou, est une "expression amicale pour s'encourager ou se soutenir. Les amis ou connaissances qui se le disent réaffirment simplement leur attachement, leur solidarité et leur complicité. On peut

aussi dire "On est ensemble" pour signifier qu'on ne se quitte pas des yeux". La Fondation Raouba Walker, découvre-t-on dans cette préface, a en effet, bien voulu "être ensemble" avec l'auteur pour offrir aux lecteurs 750 gabonismes, des mots du français parlé au Gabon. Le français, indique le préfacier, est l'unique langue officielle du Gabon; ce statut lui est conféré par la Constitution de 1991, en son article 2. Les gabonismes que répertorie l'auteur, Rossatanga dit d'eux qu'ils sont du registre du *franbonais*;

- les pages 7 et 8 enfin, rendent compte de l'avant-propos de l'auteur. Il prévient que "la langue n'est pas la même partout et en tout temps". Au Gabon, où c'est par le biais exclusif du français que les diverses cultures entrent en communication, cette langue n'est pas à l'abri des influences de ces cultures. Il se crée donc, fort de cela, un «français carrefour», et c'est celui-là qui est le plus parlé, et qui a donné lieu à son ouvrage. C'est dans un élan, dit-il documentaire qu'il a répertorié 750 mots particularisant le français parlé au Gabon. Quand il l'a pu, il a signalé le lieu de forge ou décrit l'histoire de ces mots. Il signale aussi que l'immense majorité des gabonismes qu'il propose est secrétée par les jeunes. Il prévient aussi qu'il a "rarement considéré les mots ou expressions du français courant"; quand cela s'est trouvé, c'est que ceux-ci "ont vu leur sens s'enrichir, varier partiellement ou totalement". Il reconnaît enfin que ce travail d'une grande actualité qu'il met entre les mains des lecteurs n'est pas exempt de faiblesses "que des recherches ultérieures pourraient corriger";
- la page 9 elle, et enfin, présente les signes et les abréviations. L'avantage de cette présentation à cet endroit-là est que l'utilisateur peut disposer de ces signes et abréviations avant le parcours de la nomenclature.

Les reproches que l'on pourrait formuler à l'adresse de l'auteur au sujet des pré-textes sont de deux ordres:

- nous relevons d'abord l'imprécision relative au nombre de lemmes ou entrées contenues dans cet ouvrage. Les premières pages des pré-textes (couverture et page de garde), annoncent 852 mots ou entrées, alors que les autres pages (préface et avant-propos), indiquent plutôt 750 mots et expressions qui particularisent le français parlé au Gabon. L'auteur et son préfacier laissent en rade 102 entrées, sans la moindre explication sur leur absence dans le décompte;
- l'absence d'un guide de l'utilisateur dans cette partie-là du texte. C'est aussi, sans aucun doute, l'un des points faibles relevé par certains auteurs,⁷ sur les dictionnaires qui ont porté sur les langues gabonaises, et qui ont été produits par des missionnaires et des administrateurs coloniaux. Le guide de l'utilisateur est un texte obligatoire dans tout dictionnaire. Il con-

tient les indications relatives à l'utilisation du dictionnaire et permet l'accès à toutes les données du dictionnaire (Ekwa Ebanéga 2007: 112). Dans le cas d'espèce, l'utilisateur est obligé de deviner la structure des données incluses dans l'ouvrage.

4. La nomenclature de l'ouvrage

Les suggestions portant sur la nomenclature de OEE concernent les cinq composantes suivantes: la macrostructure, la microstructure, la structure d'accès, la structure d'adressage et la médiestructure.

4.1 La macrostructure de l'ouvrage

4.1.1 La base du dictionnaire

Les ouvrages lexicographiques peuvent être constitués à partir de trois sources différentes (Wiegand 1984: 15):

- (a) les sources primaires (2 types);
- (b) les sources secondaires (3 types); et
- (c) les sources tertiaires (3 types).

L'ouvrage de Ditougou connaît ces trois types de sources, mais seules les sources primaires et les sources secondaires ont fait l'objet d'une revue par nous.

(a) Sources primaires

L'avant-propos de l'ouvrage indique que le texte est basé sur des documents. Ce sont ces derniers qui ont permis à l'auteur de répertorier les "sept cent cinquante mots et expressions qui particularisent le français parlé au Gabon". Cependant, l'auteur n'indique pas la nature des documents qu'il a utilisés. Le même avant-propos prévient aussi que l'auteur a conduit sa recherche à l'aide d'un questionnaire, mais ne mentionne ni le lieu où il a effectué la recherche, ni le nombre de ses informateurs, ni les questions qui ont guidé cette recherche. Les techniques qu'il a utilisées pour collecter ses données sont ignorées.

Ce qu'indique seulement cet avant-propos, ce sont les remerciements adressés aux membres de la Fondation Raponda Walker pour leur contribution dans la collecte des gabonismes.

Nous ne trouvons une mention des sources primaires qu'en parcourant la préface de l'ouvrage. Nous y découvrons que "c'est [...] en toute conviction que j'invite l'auteur à découvrir ce lexique agrémenté de dessin du célèbre caricaturiste gabonais du journal *L'Union*, Lybek, et dans lequel chacun finira, n'en

doutons pas, par trouver un intérêt, si minime et si futile puisse-t-il être pour certains". Ces mots du préfacier produisent la preuve que l'auteur a choisi le journal *L'Union* comme source primaire, au cours de l'élaboration de son ouvrage.

Le parcours régulier de ce quotidien fait effectivement découvrir la somme importante d'expressions du français parlé au Gabon, ainsi que celle des dessins du célèbre caricaturiste gabonais Lybek. Ce journal a donc, à n'en point douter, constitué une source primaire de OEE.

(b) Sources secondaires

Les sources secondaires ne sont pas indiquées dans les pré-textes de l'ouvrage. Nous nous contentons encore de l'allusion de l'auteur faite dans l'avant-propos, au sujet de l'utilisation de documents. Nous imaginons simplement que l'ensemble des dictionnaires consultés pour l'élaboration de ce travail (dixit l'auteur) pourraient faire partie de ces documents que nous classerions parmi les sources secondaires.

4.1.2 Les usagers-cibles

Les usagers-cibles ne sont pas explicitement indiqués dans les pré-textes de l'ouvrage. Cependant, en parcourant minutieusement l'avant-propos de l'auteur, nous pouvons les déterminer. L'auteur affirme en l'occurrence: "À l'observation, il existe de nombreux lieux de sécrétion du français parlé au Gabon. Ainsi, en dehors des milieux protocolaires et orthodoxes comme ceux de l'école, du lycée et de l'université, et de la haute administration ou de la diplomatie, où est exigé un usage moliéresque de français, il subit ailleurs les caprices de ses diverses sources de fabrication." Ce propos nous invite à penser que l'ouvrage est conçu pour satisfaire les besoins des écoliers, des lycéens, des étudiants, des enseignants, des cadres de la haute administration et des diplomates, pour les renseigner sur les diverses modifications que connaît le français en dehors de leurs milieux que l'auteur décrits comme protocolaires et orthodoxes.

4.1.3 Lemmatisation des items lexicaux

En amont du dictionnaire (cf. Avant-propos), l'auteur signale que l'immense majorité de ce qu'il appelle des gabonismes, ses 852 mots pour comprendre le français du Gabon, est produite par les jeunes qui, comme dans la plupart des pays africains, sont les plus nombreux pour pouvoir les servir sur la langue (l'organe) de tous; pour les introduire donc dans le français.

Les mots (ou lemmes) sont présentés dans l'ordre alphabétique (de A à Z). Mais dans le parcours, l'on relève l'absence des items commençant par la lettre X. Ainsi, parce que l'auteur a manqué d'indiquer, dans les pré-textes, la

raison qui motive cette absence, l'utilisateur pourrait se demander si celle-ci est due à une insuffisance au niveau de la collecte des mots du français du Gabon ou à la non-existence des termes commençant par la lettre X dans ce français.

De manière plus générale, nous avons observé que l'auteur n'a pas mentionné les critères à partir desquels les items lexicaux ont été lemmatisés dans son ouvrage. Cela amène à se demander (a) quelles sont les stratégies de lemmatisation qu'il a retenues, (b) si ce sont concomitamment les formes du singulier et du pluriel ou uniquement les formes du singulier qui ont été retenues, ou enfin (c) si c'est le radical ou le mot qui fait l'objet des entrées; ce dernier point est surtout valable pour les langues africaines, cf. tradition du mot versus tradition du thème.

Le choix de ces stratégies de lemmatisation a manqué d'être indiqué dans les pré-textes de l'ouvrage. On trouve ce genre d'informations dans les prétextes du dictionnaire, notamment dans le guide de l'utilisateur.

4.1.4 Lemmatisation des formes du singulier et du pluriel des noms

À l'observation, l'auteur a choisi d'inclure dans son ouvrage les deux formes. Cette stratégie de lemmatisation n'est pas inconnue. Elle est similaire à celle utilisée par Kriel (cité par Gouws et Prinsloo 2005: 75). Celui-ci y a déjà eu recours dans la livraison qu'il a faite du *Pukuntšu*.

Prenons par exemple les cas de *chose* (mot singulier, numéro 237) et *choses de mon corps* (groupe de mots dont le mot déterminé est pluriel, numéro 239). En amont, aucune connaissance préalable de la langue n'est demandée; l'auteur suppose que les usagers connaissent les mots utilisés ainsi que l'alphabet. Cette stratégie ne pose aucun problème pour les gabonismes puisqu'ils sont inclus en tenant compte de la stratégie de lemmatisation.

4.1.5 Lemmatisation des noms composés

L'ouvrage connaît des noms composés. Béjoint (1999: 82) distingue les noms composés de types "XY", "X Y" ou "X-Y". L'auteur présente simultanément les deux derniers types. Il distingue ainsi des noms composés de type "X Y" (*Tabac congo*, numéro 760) et "X-Y" (*S'en-fout-la-mort*, numéro 741).

Le traitement des items lexicaux composés n'est donc pas appliqué assurément dans l'ouvrage. Cela amène à se demander comment les noms composés sont inclus dans cet ouvrage. S'agit-il de lemmes ou de sous-lemmes? S'ils avaient été utilisés comme des sous-lemmes, le texte aurait réalisé une importante économie d'espace. Il n'en est rien, malheureusement. Tous les noms composés sont traités dans l'ouvrage comme lemmes. Dans l'exemple qui suit, l'article modifié du lemme **Boul** est présenté:

Boul: N.m. Diminutif de boulevard.

1. ~ **bess**. Diminutif de boulevard Bessieux, un des espaces animés de Libreville. *On s'est pris un jus au Boul Bess avant de nous séparer.*

2. ~ **mich**. France ou paris. Désignations abrégées ou diminutifs de Boulevard Saint Michel à Paris. *Il a obtenu son visa pour le Boul Mich et c'est ce soir qu'il prend l'avion.*

Dans les exemples mentionnés ci-dessus, l'article modifié du lemme **Boul** est ordonné horizontalement. Il est considéré comme le lemme principal. Il donne accès à des lemmes ordonnés horizontalement, ou des lemmes partiels (ou sous-lemmes) consécutifs à l'application du processus de condensation textuelle qui a conduit à l'omission d'une partie du lemme, à savoir: ~ **bess** et ~ **mich**.

Ces sous-lemmes sont présentés sur une nouvelle ligne et sont suivis de paraphrases de sens et d'exemples. Cette façon d'ordonner les lemmes aurait contribué à l'économie de l'espace, plutôt que d'avoir:

149- **Boul bess:** N.m. Diminutif de boulevard Bessieux, un des espaces animés de Libreville. *On s'est pris un jus au Boul Bess avant de nous séparer.*

150- **Boul ou Boul mich:** N.m. France ou Paris. Désignations abrégées ou diminutifs de Boulevard Saint Michel à Paris. *Il a obtenu son visa pour le Boul Mich et c'est ce soir qu'il prend l'avion.*

4.1.6 Lemmatisation des groupes de mots

L'auteur a également choisi de traiter les groupes de mots comme lemmes. Certains dictionnaires tels que le *South African Dictionary Junior Secondary* et le *South African Oxford Dictionary* leur donnent un statut sous-lemmatique. Les exemples de ce type de traitement sont fournis par *Faire papa et maman* (numéro 362), *Vin d'en bas* (numéro 829) et *Vin d'en haut* (numéro 830). Ces groupes de mots sont traités dans l'ouvrage comme lemmes.

Ils auraient pu, en fait, être sous-lemmatisés respectivement sous **Faire** et **Vin** comme dans les deux dictionnaires que nous évoquons, ci-dessus. Nous voyons en effet que le groupe de mots *grandfather clock* est sous-lemmatisé sous le lemme **grandfather**. Nous proposons, ci-dessous, le traitement qui aurait pu être fait des groupes de mots relevés. Cela à partir des lemmes **Faire** et **Vin**. Les sous-lemmes sont précédés des tildes.

Faire: V. 1. Réaliser par son travail, son action.
a) ~ **le bord de mer:** faire le trottoir.
b) ~ **la route:** faire le transport de personne et de biens par voie routière.
c) ~ **le blanc:** se comporter comme un blanc, adopter des mœurs de blanc.
d) ~ **les médicaments:** envoûter, ensorceler.
2. Avoir des rapports sexuels. ~ **papa et maman.**

Vin: N.m. Toute boisson alcoolisée.
1. ~ **d'en bas.** Vin récolté sur le palmier battu.
2. ~ **d'en haut.** Vin récolté sur le palmier encore debout.

Dans les exemples mentionnés ci-dessus, les articles modifiés des entrées **Faire** et **Vin** sont considérés comme des lemmes principaux. Ils donnent accès à d'autres types d'entrées qui sont des sous-lemmes.

4.1.7 Lemmatisation des abréviations

L'auteur a également choisi d'inclure des abréviations comme lemmes. Svensén (1993: 234) montre qu'il y a deux possibilités d'inclusion des abréviations dans un dictionnaire. La première voie, et la plus répandue, consiste à les répertorier séparément dans des appendices; et la seconde consiste à les répertorier alphabétiquement dans une liste principale.

Une troisième voie existe, qui fait que les abréviations soient traitées dans le dictionnaire comme des lemmes principaux. C'est celle qu'a choisie l'auteur. Mais la présentation qu'il a retenue peut être améliorée.

Il peut en effet être avantageux pour les usagers d'inclure les formes déclinaison des abréviations suivies de celles-ci. Les abréviations seraient ainsi précédées d'un symbole les spécifiant. Cette stratégie permet d'éviter la redondance que l'on observe dans la présentation des lemmes **Moutouki** (numéro 593) et **MTK** (594). On pourrait ainsi présenter cette entrée comme suit:

Moutouki – MTK [...]

Le lemme **Moutouki** serait ainsi suivi d'un tiret (–) indiquant un autre gabonisme du même mot qui est son abréviation.

4.1.8 Lemmatisation des phrases

Les phrases sont, elles aussi, incluses comme lemme dans l'ouvrage de Ditougou. Le reproche qui peut être fait ici, c'est que certaines phrases ne soient pas

sous-lemmatisées sous des lemmes principaux. L'article modifié du lemme **Chaque jour que Dieu fait** (numéro 218) peut être présenté comme suit:

Jour: N.m. Clarté, lumière du soleil permettant de voir des objets. **Chaque ~ que Dieu fait:** Jour après jour.

Dans cette présentation, le signe-lemme **Jour** est traité comme lemme principal et donne accès au sous-lemme **Chaque ~ jour que Dieu fait**; plutôt que de constituer deux lemmes distincts (cf. **Jour**, numéro 462).

4.2 La microstructure de l'ouvrage

4.2.1 Les types d'entrées

L'examen d'un article de dictionnaire exige que l'on prenne en compte les entrées⁸ et les données de celui-ci. Nous en montrons quelques aspects dans l'ouvrage de Ditougou. Considérons l'article suivant du lemme **Arranger** (numéro 38):

Arranger: V. 1. Conclure ou régler une affaire. *Tonton est parti arranger le mariage de Jacqueline.*
2. Donner des coups à quelqu'un jusqu'à le défigurer.

L'article de ce signe-lemme nous montre que *conclure ou régler une affaire* est une entrée; c'est la paraphrase de sens indiquant un des sens polysémiques du mot *arranger*. *Tonton est parti arranger le mariage de Jacqueline* est également une entrée. C'est le contexte d'utilisation du mot. La lettre V qui indique la partie du discours à laquelle appartient le mot, est aussi une entrée.

Considérons aussi l'article du signe-lemme **Assurer** (numéro 43), ci-dessous:

Assurer: V. 1. Obtenir, décrocher. *Il a assuré son passage en licence.*
2. Donner satisfaction, combler. *C'est le gars de Marie. Il assure.*

Il nous y est donné de voir que:

- la virgule (cf. ,) située entre *obtenir* et *décrocher* est une entrée. Elle constitue un élément textuel marquant une relation spécifique dans le paradigme de synonymes de cet article du lemme;
- la définition (cf. Conclure ou régler une affaire), l'exemple (cf. *Tonton est parti arranger le mariage de Jacqueline*), et la lettre (cf. V) indiquant la partie

du discours, sont des items que l'on retrouve dans l'ouvrage pour fournir des informations à l'utilisateur, au sujet du mot *arranger*.

Au titre donc des entrées, l'auteur utilise des marqueurs structurels pour spécifier un item ou rendre compte des différents types de données contenues dans l'article de dictionnaire. Il recourt, en l'occurrence à deux types de marqueurs structurels:

- (a) les marqueurs structurels typographiques; et
- (b) les marqueurs structurels non typographiques.

En ce qui concerne les marqueurs structurels typographiques, l'auteur utilise:

- le *gras* pour indiquer la représentation du signe-lemme;
- la police de caractères *roman normal* pour exposer les définitions et les synonymes;
- l'*italique* pour exposer les exemples;
- la *majuscule* pour marquer la première lettre du signe-lemme et les abréviations.

Ainsi, l'utilisateur qui recherche les entrées portant sur les exemples peut accéder rapidement au tiroir d'article où cette catégorie de données est traitée en cherchant tout simplement l'utilisation des italiques.

Concernant les marqueurs structurels non typographiques, nous avons ainsi observé que:

- l'auteur utilise des chiffres mis en gras (1, 2, ... 852). Ceux-ci précèdent les signes-lemmes pour indiquer le numéro d'ordre du lemme traité;
- d'autres chiffres (1, 2, ...) contenus dans un cercle noir sont utilisés comme entrées pour indiquer les différents sens du mot;
- les parenthèses servent respectivement à indiquer: le genre du mot (cf. **Bonami(e)**, numéro 126); la variante orthographique ou de prononciation (cf. **A tare zame! (A tiri zame!)**, numéro 3, **Lbv (prononcé elbève)**, numéro 394), ou un mot (ou groupe de mots) accessoire(s) et complémentaire(s) (cf. **cellulaire (Le)**, numéro 211; **Ensemble (on est)**, numéro 355; **Même pipe même tabac (c'est)**, numéro 562);
- le signe (,) suivant immédiatement un signe-lemme annonce l'accord féminin d'un mot (cf. **Coudou, te**, numéro 269);
- le signe (:) placé immédiatement après le signe-lemme annonce le traitement microstructurel du mot (cf. **A mana:**, numéro 1);
- le signe (!) aide à exprimer une interjection (cf. **A tare zame!**, numéro 3);

- les signes (*) et (↔) indiquent les renvois au(x) mot(s) utilisé(s) dans l'ouvrage comme signes-lemmes.

Au sujet de ces deux derniers signes, nous pensons qu'un seul signe aurait dû être utilisé pour ce type de renvoi.

4.2.2 Les données sur l'orthographe

Comme nous l'indiquons, ci-dessus, l'ouvrage comporte des mots qui ont des variantes orthographiques. Voici quelques articles:

A tare zame! (ou A tiri zame!): Exclamation ou juron fang qui signifie « mon Dieu! »
Abeau ou Abo: Apost. Abréviation de Beau-frère. *Je compte sur toi, Abo, pour mon billet de retour.*
Akiééh! (ou Ekiééh!): Expres. En langue fang, interjection marquant l'étonnement. *Akiééh! Il a osé te dire une telle méchanceté.*
Baude ou bode: N.f. 1. Jolie fille. *Pas mal, la baude qui passe.* 2. Petite amie, amante. *Je te présente Fanny. C'est ma baude.*
Riengo (ou oriengo): N.m. Du verbe ipunu « urienga », signifiant se déhancher ou faire rouler sa croupe (en faisant l'amour); danse inventée par un handicapé pendant la célébration de la fête nationale du 17 août 2003 et devenue depuis lors très célèbre.

Dans les exemples repris, ci-dessus, l'auteur utilise alternativement la conjonction de coordination *ou* seule, et concomitamment la même conjonction de coordination et des parenthèses, pour rendre compte des variantes orthographiques. Il aurait dû ne retenir qu'un seul des deux procédés.

4.2.3 Les données sur la prononciation

Pendant la confection du dictionnaire, l'une des importantes décisions que le lexicographe doit prendre consiste à décider de la transcription (phonétique) de tous les signes-lemmes, ou de celle de ceux qui posent des problèmes de prononciation. Nous avons relevé que l'auteur a choisi d'assortir un lemme d'une spécification pour indiquer qu'il a une prononciation particulière:

Lbv (prononcé elbève): abréviation de Libreville, à l'origine dans le langage de l'aviation. *Je passerai mes vacances à Lbv.*

Le traitement de l'article du signe-lemme **Lbv** montre que l'auteur a choisi la transcription orthographique au détriment de la transcription phonétique. Ce-

pendant le choix de la transcription orthographique n'est pas expliqué dans les pré-textes de l'ouvrage.

Ce que l'on remarque c'est que la transcription orthographique pourrait combler les besoins des usagers qui ont une bonne connaissance de l'emploi des gabonismes, mais ne pas satisfaire ceux qui ne disposent pas de cette connaissance. Pour éviter ce travers, la solution aurait dû être le recours à la transcription basée sur l'Alphabet Phonétique Internationale (API). Lombard (1990: 110) lie ce recours au fait que l'API soit largement utilisée et Louw (2004: 118) pour l'accréditer, dispose que la standardisation universellement répandue de l'API induit l'utilisation des mêmes symboles dans tous les dictionnaires, pour dénoter les mêmes sons. L'auteur aurait donc mieux fait d'utiliser la transcription basée sur l'API.

4.2.4 Les données sur les parties du discours

Dans OEE, les parties du discours telles que le nom, le verbe, l'adjectif ou l'adverbe sont respectivement signalés dans l'ouvrage sous la forme d'abréviations (cf. N. m. ou f., V., Adj., Adv.). Le texte souffre ainsi du traitement des sous-catégories des parties du discours, notamment pour le verbe. L'auteur a représenté cette partie du discours par la seule abréviation V., mais a manqué d'indiquer les sous-catégories aidant à mieux caractériser cette classe comme *verbe transitif* ou *verbe intransitif*.

Gouws (1999: 223) distingue deux types d'indications des parties du discours, à savoir:

- (a) la classification lexicale primaire; et
- (b) la classification lexicale secondaire.

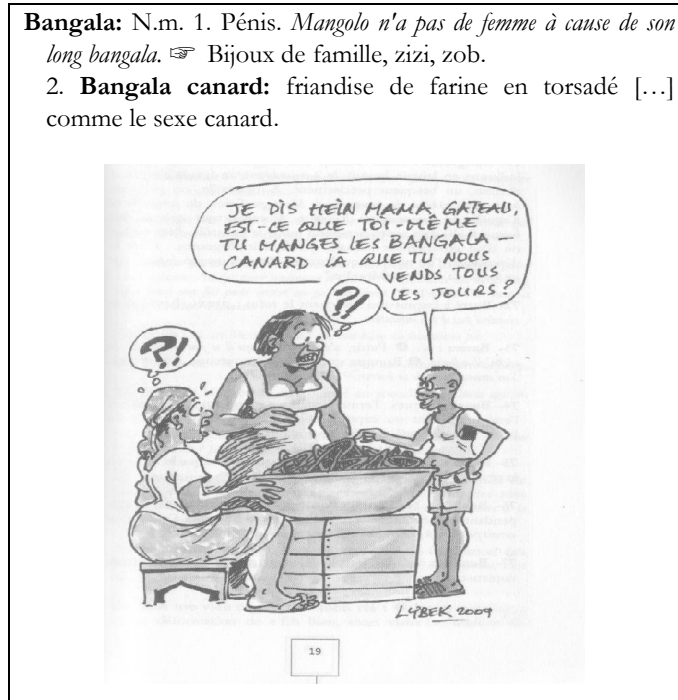
La classification lexicale primaire se définit comme étant la première classification adressée au lemme. La classification lexicale secondaire elle, s'entend comme celle qui est adressée à la première classification et qui devrait être lue, à l'unisson, avec elle.

L'auteur aurait dû utiliser ce principe de classification pour les verbes. Ainsi on aurait comme classification primaire la partie du discours « verbe » (V.) et comme classification secondaire « verbe transitif » ou « verbe intransitif » (V. trans. ou V. intrans.). Il a d'ailleurs manqué de décrire les abréviations des items représentant les parties du discours dans les pré-textes de l'ouvrage.

4.2.5 Les données sémantiques

L'auteur de OEE inclut des illustrations dans son traitement microstructurel. Nous avons relevé quelques carences dans celui-ci. Considérons l'article du lemme **Bangala** (numéro 69).

Bangala: N.m. 1. Pénis. *Mangolo n'a pas de femme à cause de son long bangala.* ☞ Bijoux de famille, zizi, zob.
2. **Bangala canard:** friandise de farine en torsadé [...] comme le sexe canard.



Ce que l'on observe, c'est que l'illustration est précédée de la définition du signe-lemme. Le reproche que l'on peut faire à l'auteur est qu'il réserve un nombre trop important et trop d'espaces aux illustrations. Il aurait dû réduire la taille de celles-ci pour gagner ainsi en nombre de pages (plus de 20 pages). En outre, le but des illustrations est de compléter l'information quand la compréhension d'un mot pose problème; or, l'illustration ci-dessus n'est pas claire et l'usager a du mal à saisir le sens du mot *bangala*, à partir d'elle.

4.3 La structure d'accès de l'ouvrage

Nous rappelons, pour mémoire, que la structure d'accès est l'itinéraire que suit un usager pour atteindre (dans le dictionnaire) un signe-lemme spécifique, ou une catégorie de données. Hausmann et Wiegand (1989: 337) en distingue deux types,⁹ à savoir:

- (a) la structure d'accès externe; et
- (b) la structure d'accès interne.

Dans l'ouvrage de Ditougou, la structure d'accès externe est constituée par: les entrées situées sur la couverture, dont le titre de l'ouvrage, suivi de 852 mots pour comprendre le français du Gabon; le nom de l'auteur, un dessin illustrant

une exécution du parler français du Gabon; la déclinaison de l'éditeur (Raponda Walker).

La structure d'accès interne elle, se caractérise par l'utilisation des conventions lexicographiques qui guident l'utilisateur à travers le chemin de sa recherche. Ces conventions comportent deux types de marqueurs structurels:

- (a) les marqueurs structurels typographiques; et
- (b) les marqueurs structurels non typographiques.

Les marqueurs structurels ont été exposés dans la section précédente consacrée aux entrées. Cependant, il est bon de faire observer que le texte souffre du mauvais traitement de ceux-ci. Considérons les exemples des articles des lemmes suivants:

Bangando: N.m. Tenant son origine du terme ngando (caïman en langue ipunu), le bangando est un bandit de grand chemin, un braqueur précisément. A Libreville, on parle de *bangando d'Akébé*, de *bangando de Montagne-Sainte*, de *bangando de Kinguélé*, de *bangando de Rio*, etc. Comme tout caïman, un bangando désigne toute personne qui a le contrôle d'un espace ou d'un groupement humain qui lui fait allégeance. C'est le dépositaire d'un pouvoir quelconque. *Kuta a beaucoup de pia, c'est un bangando.*

Barré! Expression marquant le refus; jamais. *Barré! Je ne viendrai pas à ton invitation.*

Bendelle: N.f. Jeune amante coquette. *C'est Fanny. C'est ma bendelle. Go*, bode*, petite*.*

Ce que l'on observe dans les illustrations, ci-dessus, et à travers l'ensemble du texte, c'est que les exemples sont présentés en italique. Or, les exemples des articles de signes-lemmes proposés ici, montrent une double utilisation de l'italique pour les exemples et les renvois. L'auteur aurait gagné en utilisant un seul principe de présentation spécifique à chaque donnée de l'ouvrage. Ceci aurait permis à l'utilisateur d'avoir un accès rapide aux données du dictionnaire.

4.4 La médiostructure de l'ouvrage

Les lexicographes utilisent les renvois pour signifier aux usagers qu'il existe dans l'ouvrage une autre entrée qui contient des données additionnelles du lemme recherché. Pour ce qui est de OEE, l'auteur utilise le marqueur de renvoi ↗ et l'entrée indiquant l'adresse de référence. Cependant, nous notons un certain nombre de manquements au sujet du traitement des renvois faits dans le dictionnaire. Ceux-ci portent sur:

- (1) des renvois non existants;
- (2) un renvoi qui ne guide pas l'utilisateur;
- (3) la non-utilisation de renvois; et
- (4) l'utilisation d'un renvoi pour éviter le traitement du signe-lemme.

4.4.1 Renvois non existants

L'une des erreurs que les lexicographes commettent souvent est de fournir à l'utilisateur une entrée de renvoi qui se réfère à une adresse-renvoi qui n'existe pas. C'est le cas avec l'exemple de l'article du signe-lemme suivant:

<p>Balles: N. Argent. ☞ kolo*, pya*, tchoko*.</p> <p>Bangala: N.m. 1. Pénis. <i>Mangolo n'a pas de femme à cause de son long bangala.</i> ☞ Bijoux de famille, zizi, zob [...]</p>
--

Dans les articles des lemmes **Balles** et **Bangala**, les renvois à *tchoko*, *zizi* et *zob* sont non existants puisque les mots ne sont pas traités comme des signes-lemmes dans l'ouvrage.

4.4.2 Renvoi qui ne guide pas l'utilisateur

L'ouvrage connaît un type de renvoi qui ne guide pas l'utilisateur par rapport à l'information à en tirer. C'est ce qu'on observe dans l'article du signe-lemme **Cafette** (numéro 189):

<p>Cafette: [...] petit restaurant rapide bon marché. ☞ Dos-tourné.</p> <p>Dos-tourné: [...] petit restaurant rapide bon marché; restauroute.</p>

Dans cet article du signe-lemme **Cafette**, le synonyme *restauroute* amène à se demander si celui-ci n'est lié qu'à **Dos-tourné**, et non à **Cafette**. En outre, aucun signe de renvoi n'est utilisé ni avec **Dos-tourné**, ni avec **Cafette**. La relation entre **Cafette** et **Dos-tourné** n'est donc pas claire. L'utilisateur ne peut pas déterminer la relation qui sous-tend ces deux mots.

4.4.3 Non-utilisation de renvois

Un autre type de lacune que connaît l'ouvrage en matière de renvoi, est la non-utilisation de ceux-ci où ils sont nécessaires. Ainsi, certains renvois ne sont pas

utilisés lorsqu'une adresse de référence existe. Il en est ainsi de l'article du signe-lemme **Faire**:

Faire: V. 1. Avoir des rapports sexuels [...]

Dans l'article ci-dessus, le terme **Faire** est traité comme un signe-lemme. Ici, l'usager obtient des informations à partir du traitement du signe-lemme en question, notamment, la partie du discours et la définition. Dans l'article du signe-lemme **Faire papa et maman**, on trouve quasiment la même définition:

Faire papa et maman: Expres. Avoir des relations sexuelles.
Forme contractée de « faire comme papa et maman ont fait pour me mettre au monde ».

Malheureusement, aucun renvoi n'est mentionné, ni à partir du signe-lemme **Faire**, ni à partir de celui **Faire papa et maman**. L'auteur aurait gagné à mentionner un renvoi à la suite de l'article du signe-lemme **Faire papa et maman** pour se référer à l'article du signe-lemme **Faire**.

4.4.4 Utilisation de renvoi pour éviter le traitement du signe-lemme

Ce problème apparaît quand le lexicographe ne donne pas le traitement complet du signe-lemme. Il utilise dans ce cas le renvoi pour éviter le traitement complet du signe-lemme en question. C'est ce que nous observons dans les articles des signes-lemmes **Afro** (numéro 14) et **Peigne afro** (numéro 642):

Afro: N. Coiffure masculine ou féminine touffue sur l'ensemble du crâne. ⇨ Peigne afro.
Peigne afro: N.m. 1. Genre de peigne approprié à la chevelure dure et drue de certains Africains et dont le manche présente un poing [...]

En dehors des définitions de ces deux signes-lemmes, aucun exemple n'est proposé pour les illustrer. Un renvoi est cependant fait à **Peigne afro**. Dans l'article du signe-lemme **Peigne afro**, une autre définition est donnée à **Afro**, à savoir: chevelure dure et drue de certains Africains, qui souffre de ne pas disposer d'exemple de soutien.

5. Les post-textes de l'ouvrage

Les post-textes qui constituent la troisième composante du dictionnaire ne sont pas du tout fournis dans l'ouvrage de Ditougou. Ils comprennent uniquement les publications de la Fondation Raponda Walker. C'est à se demander si celles-ci sont d'un apport quelconque à ce texte.

C'est en fait en leur sein que l'on devrait trouver les signes et abréviations contenus dans les pré-textes. C'est seulement après eux qu'auraient dû figurer les publications de la Fondation Raponda Walker.

6. Conclusion

Nous observons que OEE est un ouvrage de référence qui répond à certains critères scientifiques. Tout est donc à l'honneur de son auteur, qui n'a rien, au départ, d'un lexicographe. Il peut, en effet, être décomposé en trois moments, à savoir: les pré-textes, la nomenclature et les post-textes, et comporte des sources qui constituent une base de dictionnaire. Il satisfait ainsi, et entre autres, certaines stratégies de lemmatisation, avec l'inclusion de noms, noms composés, groupes de mots, phrases, abréviations, etc.

Cependant, des améliorations portant sur les trois aspects de sa structure auraient donné une meilleure qualité à sa tenue.

Ainsi, sur le plan des pré-textes, l'auteur a semé une confusion dans l'esprit des lecteurs au sujet du nombre de lemmes que comporte l'ouvrage. Dans le nombre que son préfacier et lui y indiquent, ils laissent en rade cent deux lemmes, sans la moindre explication sur leur absence dans le décompte. Le texte est également dépourvu d'un guide de l'utilisateur, ce qui fait que le lecteur est obligé de deviner la structure des données incluses dans l'ouvrage.

Sur le plan de la nomenclature, des insuffisances ont été relevées aux niveaux de la macrostructure, de la microstructure, de la structure d'accès et de la médiestructure.

Au sujet de la macrostructure, nous nous sommes appesantis sur les sources qui ont aidé à la composition de l'ouvrage, sur les usagers-cibles et sur la lemmatisation des items.

L'auteur a manqué d'indiquer les sources primaires et secondaires qui lui ont permis de répertorier les 852 mots et expressions qui particularisent le français parlé au Gabon. Il a également manqué d'indiquer les usagers-cibles à qui son texte est destiné.

Pour la lemmatisation des items, nous avons relevé l'absence de ceux commençant par la lettre X, mais l'auteur n'indique pas la raison qui motive cette absence. Il n'a pas indiqué non plus, les stratégies de lemmatisation qu'il a retenues pour les différentes entrées. Nous avons ensuite parcouru les différents lemmes qui constituent l'ouvrage, en montrant les différentes limites que comporte leur présentation.

Au sujet de la microstructure, l'auteur inclut des données telles que les marqueurs structurels typographiques et non typographiques, les définitions, les exemples, les synonymes et les illustrations. Nous pensons qu'il aurait dû retenir une seule entrée qui représenterait une seule donnée du dictionnaire. Il aurait également dû réduire la taille de ses illustrations, puis les rendre plus claires, pour que l'utilisateur puisse bien comprendre le sens du mot parcouru.

Au sujet de la structure d'accès, nous avons relevé une double utilisation

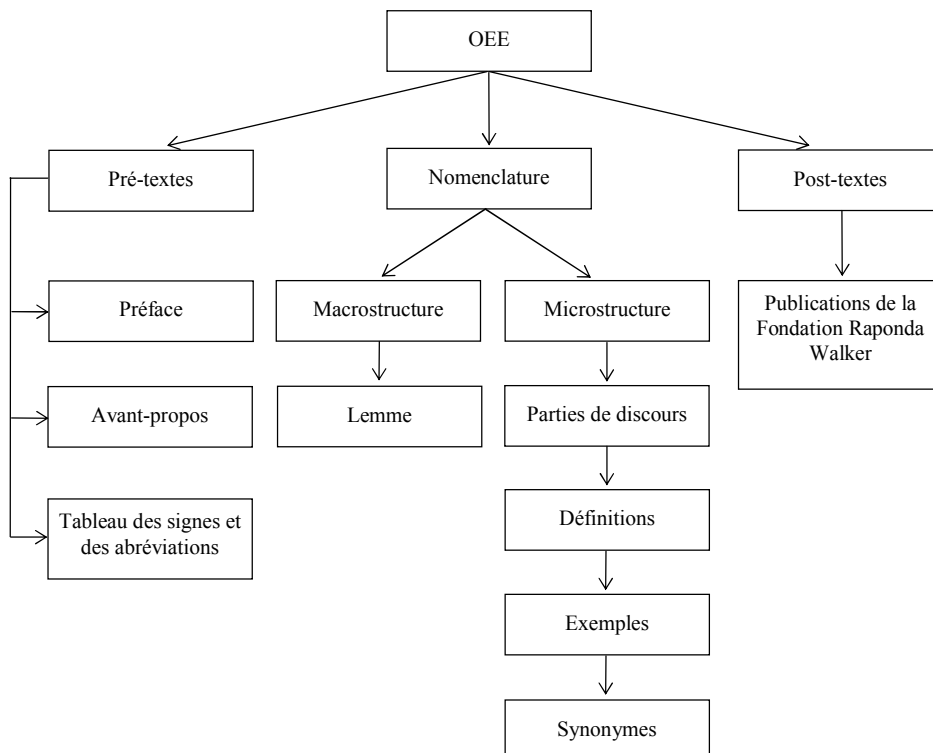
de l'italique pour les exemples et les renvois. L'auteur aurait gagné à utiliser un seul principe de représentation spécifique à chaque donné.

Au sujet de la médiostucture enfin, un certain nombre de manquements ont été relevés, qui concernent le traitement des renvois faits dans l'ouvrage. Ceux-ci ont porté sur des renvois signalés mais qui ne sont nullement traités dans le texte; un renvoi qui ne guide pas l'utilisateur; la non-utilisation de renvois lorsqu'il se doit et l'utilisation d'un renvoi pour éviter le traitement d'un signe-lemme.

Sur le plan des post-textes, nous avons observé que ceux-ci ne comportent que l'inventaire des publications de la Fondation Raponda Walker qui ne sont en fait, d'aucune utilité à cet ouvrage. C'est en leur sein que devraient se trouver les signes et abréviations contenus dans les pré-textes.

Les améliorations que nous venons de suggérer, nous le rappelons, en auraient probablement ajouté à la qualité du texte, pour le mettre en phase avec les principes scientifiques qui sous-tendent la confection d'un dictionnaire. Mais leur absence n'entame en rien la tenue de cette proposition riche et intéressante pour l'utilisateur. C'est l'occasion pour nous de rendre un hommage mérité à son auteur et de l'encourager à ne point se détourner de *On Est Ensemble* car, il y a encore fort à faire, le français étant en perpétuelle mutation au Gabon.

Voici, pour clore notre propos, la vue synoptique de l'ouvrage, son ossature:



Notes

1. Le lemme est l'unité lexicale (généralement présenté en gras) qui sert d'unité de traitement à un article du dictionnaire. Dans certaines littératures lexicographiques francophones, le lemme est appelé « entrée », alors que dans la littérature anglo-saxonne, il est dénommé « headword ».
2. Comme par construction d'un nid. La nidification implique un regroupement des lemmes qui escamotent l'ordre alphabétique strict pour faire ressortir les relations morphosémantiques entre les mots (cf. Hausmann et Wiegand 1989: 336; Gouws 2001: 79-80)
3. Comme par construction d'une niche. La nichification implique un groupement alphabétique strict des lemmes qui sont plus ou moins liés sémantiquement (Gouws 2001: 125).
4. C'est tout le traitement lexicographique présenté pour un lemme donné. Il part du signe-lemme jusqu'à la fin des données lexicographiques traitées à la suite de celui-ci. Un article de dictionnaire renferme une variété de données lexicographiques et contient une macrostructure, une microstructure, une structure d'adressage, une médiostructure, une structure d'accès et une structure de distribution des données.
5. Discours théorique sur la production des ouvrages de référence lexicographique.
6. Science de la production des ouvrages de référence lexicographiques: dictionnaires, encyclopédies, lexiques ...
7. Ondo-Mébiame (2005), Tomba Moussavou (2007) et Ekwa Ebanéga (2007).
8. Toutes données lexicographiques présentées après le lemme, contrairement à la considération habituelle qui en fait l'unité graphique mise en vedette et qui fait l'objet du traitement.
9. Voir la caractérisation de ces structures d'accès dans la définition des notions, en 1-, supra.

Bibliographie

- Béjoint, H.** 1999. Compound Nouns in Learners' Dictionaries. Herbst, T. and K. Popp (Éds.). 1999. *The Perfect Learner's Dictionary (?)*: 81-99. Tübingen: Max Niemeyer.
- Chambers-Macmillan.** 1996. *Chambers-Macmillan South African Dictionary Junior Secondary*. Manzini: Macmillan Boleswa.
- Ekwa Ebanéga, G.-M.** 2007. *A Microstructural Programme for Dictionaries in Fang*. Thèse de doctorat non publiée. Stellenbosch: Université de Stellenbosch.
- Gouws, R.H.** 1999. *A Theoretically Motivated Model for the Lexicographic Processes of the National Lexicography Units*. Rapport de recherche soumis au Pan South African Language Board.
- Gouws, R.H.** 2001. Formation lexicographique: approches et thèmes. Emejulu, J.D. (Éd.). 2001. *Éléments de lexicographie gabonaise. Tome I*: 95-134. New York: Jimacs-Hillman.
- Gouws, R.H. et D.J. Prinsloo.** 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PReSS.
- Hausmann, F.J. et H.E. Wiegand.** 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. (Éds.). 1989-1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*: 328-360. Berlin: Walter de Gruyter.
- Hawkins, J.M.** 1996. *The South African Oxford School Dictionary*. Cape Town: Oxford University Press.

- Kriel, T.J.** 1983. *Pukuntšu/Dictionary*. Pretoria: J.L. van Schaik.
- Lombard, F.J.** 1990. *'n Metaleksikografiese fundering van Afrikaanse skoolwoordeboeke*. Thèse de doctorat non publiée. Stellenbosch: Université de Stellenbosch.
- Louw, P.A.** 2004. *Criteria for a Multifunctional, Monolingual Dictionary in Junior Secondary Education*. Thèse de doctorat non publiée. Stellenbosch: Université de Stellenbosch.
- Mihindou, G.-R.** 2001. Apports des missionnaires à la lexicographie gabonaise: dictionnaires bilingues fang-français/français-fang; français-yipounou/yipounou-français; français-mpongwé. Emejulu, J.D. (Éd.). 2001. *Éléments de lexicographie gabonaise. Tome I: 7-37*. New York: Jimacs-Hillman.
- Ndinga-Koumba-Binza, H.S.** 2006. *Lexique Pove-Français/Français-Pove*, Mickala Manfoumbi: Seconde Note de Lecture. *Lexikos* 16: 293-308.
- Ondo-Mébiame, P.** 2005. De la révision du dictionnaire de Samuel Galley. *Lexikos* 15: 151-163.
- Svensén, B.** 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Traduit du suédois par J. Sykes et K. Schofield. Oxford/New York: Oxford University Press.
- Tomba Moussavou, F.** 2007. *Metalexigraphic Criteria for a Compilation of a Descriptive Monolingual Dictionary in Yipunu*. Thèse de doctorat non publiée. Stellenbosch: Université de Stellenbosch.
- Wiegand, H.E.** 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Éd.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983: 13-30*. Tübingen: Max Niemeyer.

Monika Bielińska. *Lexikographische Metatexte. Eine Untersuchung nichtintegrierter Außentexte in einsprachigen Wörterbüchern des Deutschen als Fremdsprache.* 2010, 375 pp. Danziger Beiträge zur Germanistik 32. ISBN 978-3-631-60500-4 (Hardback). Frankfurt am Main: Peter Lang. Price: €61.50.

This book, which is based on the doctoral dissertation of Monika Bielińska, investigates the outer texts of several monolingual learners' dictionaries for German as a foreign language. She maintains that German learners' dictionaries constitute a fairly young addition to German lexicography, the first one only having been published in 1993 by Langenscheidt (1993). This is in stark contrast to countries such as France and England, where the publication of learners' dictionaries has long been established.

Bielińska selected thirteen dictionaries which she considers to be representative of the genre of monolingual learners' dictionaries for German as a foreign language. These include different editions of the same dictionary, and smaller dictionaries as well as more comprehensive ones. Her main aim with the investigation is to contribute to "dictionary care" with the idea that her comments could help lexicographers to plan outer texts more carefully in order to fulfil users' needs to a greater extent. The primary users of dictionaries for German as a foreign language are, of course, learners of German. The secondary users are lexicographers and metalexicographers who research the properties of dictionaries as a field of study.

Bielińska draws up lists of criteria for learners' dictionaries which she uses in her analysis of the outer texts of the German dictionaries in question. In her analysis, she not only looks at the outer matter of these dictionaries, but also takes into account comments made in dictionary reviews and analyses, and the vast metalexicographical literature as a whole. The "catalogue of questions" (p. 308) she draws up reflects the expectations of lexicographical researchers.

Some of the aspects which Bielińska investigates are (a) the target group (or the addressees) and the extent to which differentiation was made between beginners and advanced learners; (b) the needs of the particular users; (c) the principles according to which the dictionary was conceptualised (which have to do, among others, with explanations and guidelines pertaining to the layout of the dictionary and its outer texts, the way in which the presentation of data in the word list is dealt with in the outer texts, the extent to which the selection of lemmas was justified in the outer texts, and the way in which the outer texts inform the user about finding and interpreting the data types given in the dictionary articles); and (d) the user-friendliness of and the functions of the outer texts with regard to situations of dictionary use.

The contents and form of lexicographical outer texts could have a profound influence on the successful use of dictionaries, even though research has shown that most dictionary users never consult the outer texts. For example, Kipfer (1987: 45) in a study of American students, established that "the only explanatory matter students felt a need to refer to was the pronunciation key"

(p. 60, note 36). Bielińska (p. 60, note 36) refers to Svartvik (1999: 283) who mentions that it is "most atypical of the run-of-the-mill user" to read user's guides. Hartmann (2000: 387) also points out that 70,9% of users admit to "manage without front-matter notes altogether", and that "information contained in the appendices of dictionaries is rarely consulted" (p. 60, note 36).

According to Bielińska (p. 53), there are different types of outer texts. The first type, which should be placed in the front matter, should help users to successfully and effectively consult the dictionary. These outer texts should include the list of contents, the introduction, the users' guide (which should contain sample articles explaining the structures and contents of the different article types), the abbreviations used in the articles to denote different item types, the system used in the alphabetical ordering, and a list of phonetic signs.

The second type of outer texts should support users in their language acquisition and language use, including theoretical information on language. These outer texts could inter alia deal with information about inflections and the formation of tenses on the one, and semantic relations, lists of idioms, and empty formulae on the other hand. In addition, monolingual learners' dictionaries for foreign users should give information on the geography, culture and history of the language area. Such texts could either be in the front matter or in the back matter, or they could appear as insertions into the word list throughout the dictionary. With regard to this second type of outer texts, Bielińska refers to Gouws (2004), who distinguishes between outer texts with *knowledge-orientated functions* (e.g. texts giving encyclopedic non-language information as well as theoretical information on language) and those with *communication-orientated functions* (e.g. giving information on how to actively deal with the language). This distinction would also be applicable to monolingual learners' dictionaries for foreign learners. Bielińska (p. 54) expands the communicated-orientated functions to include outer texts of the first type, namely, texts which enhance *the effective use of the dictionary*. Outer texts dealing with the layout and use of a particular dictionary have a *metatextual* relation with the word list, whereas other outer texts, such as the ones giving geographical information, can be considered as being of a *paratextual* nature. Using Bergenholtz and Tarp's (2005) classification, Bielińska (p. 55) identifies metatexts as *non-integrated outer texts* and paratexts as *integrated outer texts*.

Integrated outer texts are for example diagrams, tables and illustrations which could range from lists of irregular verbs to lists of measurements and weights, lists of numbers, and the names of the days and months, and many more. The function of these lists is to enhance systematic expansion of the user's vocabulary within the given contexts, because the context disappears in an alphabetical ordering system. These integrated outer texts are not the main focus of Bielińska's work.

The greater part of Bielińska's book is dedicated to her analysis of the non-integrated outer texts in the thirteen selected German monolingual dictionaries for German as a foreign language. She evaluates three aspects of the *contents* of

these texts. Firstly, she looks at the *conceptual background* of the metatexts. Secondly, she determines whether the metatexts give an appropriate description of the different parts of the dictionary as a whole, and of the structures of the dictionary articles in the word lists (p. 90). An explanation of the structures of dictionary articles is usually done by means of sample articles. Thirdly, the *meta-textual treatment* of the individual types of lexicographical data in the word lists is examined. Bielińska draws up a "catalogue of questions", some of which are only relevant to users, some only to researchers of dictionaries, and some relevant to both these groups. This catalogue of questions will briefly be discussed in the rest of this review, in order to whet the reader's appetite for this book.

Incorporated in the issue of the *conceptual background* of dictionary metatexts is the question whether these texts clearly mention the envisaged target group (or group of addressees). Bielińska (p. 115-116) reveals that most of the dictionaries are quite vague on who their addressees are ("all learners of German"), and that the smaller dictionaries are more precise than the more comprehensive ones (p. 290). Especially advanced learners are usually not mentioned as addressees, because of the vagueness regarding the identity of the users. It is, of course, of crucial importance that users will know whether a dictionary is targeted at them before they buy it.

Another conceptual issue is whether the metatexts mention *the functions* of a particular dictionary. The selected dictionaries are not totally inadequate with regard to mentioning the dictionary functions in the metatexts, although some of them only state the functions implicitly which entails that one has to look for them in different metatexts as well as in the advertising text on the cover, within one and the same dictionary. Users might want to know the functions of the dictionaries before they buy them, so it is important that they can find *inter alia* information whether the dictionary is meant for passive or active language use.

Information on the *dictionary basis* also resorts under the conceptual background of dictionaries. Bielińska poses questions about whether text corpuses, other dictionaries, literature on grammar and other linguistic works were used in the planning and compilation phases, and if so, which ones, and whether they are mentioned in the metatexts of the analysed dictionaries. An interesting phenomenon is that many of the dictionaries claim on their covers that they were "completely reworked" or "totally newly developed" — a claim which almost always proves to be incorrect. Bielińska (p. 138) admits that the issue of the dictionary basis is not interesting to the average user, even though researchers of dictionaries would definitely want this information. Of crucial importance to the user, however, is the question whether a particular dictionary is based on a previous one which the user may already possess. This implies that giving false information in this regard will confuse users and could waste their money.

Another aspect which Bielińska investigates is whether criteria for the outer selection of lemmas and for the selection of data types, as well as the illustrations, were justified in the metatexts (p. 139-146). Even though she admits

that average users may not be interested in these criteria, dictionary researchers may need them. The smaller dictionaries mostly do not give any information on the dictionary basis in their metatexts, whereas the more comprehensive ones do give some information, although in a very general and hardly informative way.

Regarding the mentioning of the *text compound* in the particular dictionaries, Bielińska looks at issues such as whether the metatexts present the reader with information on and search paths to all the direct constituents of the dictionary, and whether the macrostructural ordering and problems of finding certain lemmas were mentioned. In addition, she tries to determine whether the dictionaries unambiguously stated the number of lemmas they contain. Whereas the first cluster of questions could be given a positive evaluation, Bielińska states that most of the dictionaries gave a less than satisfying report on the number of lemmas included in the word list. Many of the dictionaries contain far fewer lemmas than they claim to have, because they would give vague counts which include not only the lemmas themselves, but also the examples and the idiomatic expressions. This makes it impossible to ascertain the exact number of actual lemmas, which in its turn makes it impossible for users to compare the range of different dictionaries with each other.

With regard to the information and the help metatexts give on the structures of dictionary articles, Bielińska (p. 170) detects two theoretical principles: (a) a lexicographer can decide to discuss as many item types as possible, or (b) he/she can decide to discuss only the ones most frequently needed by users. Bielińska is of the opinion that the dictionaries which only discussed a smaller number of sample articles treated these with greater care. For example, in spite of the existence of ample metalexicographical literature on the treatment of sample articles, the latest edition of the Langenscheidt dictionary (2008) did not make use of this, which resulted in inadequate sample articles. The presentation of more precise and more useful sample articles will not necessarily take up more space in outer texts; on the contrary: Bielińska suggests that one simply has to select the sample articles more carefully and more purposefully (p. 170).

In her investigation of the treatment of spelling in the metatexts of the analysed dictionaries, Bielińska (p. 178-179) concludes that only the more comprehensive dictionaries give sufficient information. Pronunciation is mostly treated sufficiently in all the metatexts (p. 186-187).

With regard to expositions in the metatexts of the grammatical information given in the dictionary articles, Bielińska (p. 299) finds that, apart from the *Wörterbuch Deutsch als Fremdsprache* (2000) from the De Gruyter publishing house, which mentions one textbook on grammar, none of the dictionaries mention any grammatical textbooks they used in the planning and compilation of their dictionaries. Her conclusions on the treatment of parts of speech (p. 211-212), inflection (p. 220), morphology (p. 232-233), and syntax (p. 243-244) are summed up on page 305, where she claims that the more comprehensive dictionaries give more useful and informative details on these gram-

matical aspects in their metatexts than the smaller ones. This does not mean, however, that the larger dictionaries are complete in all respects.

Research has shown that the meaning of the lemma is the item type most often looked up in dictionaries (e.g. Ripfel and Wiegand 1988: 512, Ripfel 1989: 192; Köster and Neubauer 1994: 225). According to Bielińska (p. 244), it is therefore very important that the meaning descriptions are understandable to someone who is only learning a language: the descriptive language should be comprehensible. The metatexts of the analysed dictionaries were satisfactory in helping users where to find the meaning descriptions, how the typographical and non-typographical markers work, and what different terms were used to describe the treatment of meaning descriptions (p. 246-247). One problem remains that the user sometimes has to look for the information in several metatexts, which is of course not user-friendly (p. 315, note 182).

Context is another aspect of a dictionary article which requires close attention, especially if the dictionaries claim that they are meant to help users in the *production of texts*. Bielińska (p. 263-264) finds that most of the dictionaries gave guidelines in their metatexts on their treatment of context, and their presentation of these items typographically and/or non-typographically. Sometimes, this information could only be obtained by means of the sample articles. The metatexts were not very helpful in giving guidelines on the ordering of items providing the context, or information on the descriptive language used to present this demonstrative part of the dictionary article, and in what ways the items giving the collocation differ from the items giving the examples (p. 263-264).

Fixed expressions and idioms are not sufficiently discussed in the metatexts of the dictionaries which Bielińska (p. 274-275) analysed. There are so many studies in literature on the phenomenon of fixed expressions and idioms in lexicography that it is hard to understand why the analysed dictionaries gave so little attention to it in the metatexts. No criteria for the selection of fixed expressions and idioms were given (which may only be important to researchers and not to the learners). The notation techniques for the presentation of these items were usually not introduced, which is important for the user-friendliness in the dictionary. In addition, there were no clear indications of the basic morpho-syntactic form of fixed expressions, which is of course important information for learners of a foreign language for producing texts.

The system used for marking is also very important in learners' dictionaries. In this context, Bielińska's first question was to determine whether the *system of marking* used by the dictionary was systematically and fully explained in the metatexts. The answer was "no" in all of the cases, which points to a serious shortcoming. The *format* and the *position* of the markers, on the other hand, were given by all the dictionaries.

The users' guidelines differ greatly from dictionary to dictionary. On page 123, Bielińska gives some examples of useless guidelines in a particular dictionary. It offers to help students in finding the right words, but instead, some of the headings are as follows: "Learn some new words early in the morning

when getting up and in the evening before going to sleep", or "Tape notes in your home or at your office to the furniture with the words on them that you want to learn", or "Listen to German music, and try to understand the individual words". These "tips" do not really contribute to the learning process of users or to guaranteeing successful dictionary use.

One aspect which Bielińska mentions time and again throughout her book is the fact that many lexicographers themselves do not believe in the usefulness of metatexts, or the outer matter, of the dictionaries they write. She suggests (p. 308) that they might perhaps think that they are wasting their time and effort, because users will in any case not read the outer texts. She quotes for example Landau (2001: 148), who writes: "it is widely believed among lexicographers that no one reads it [the front matter]."

The question remains as to what extent dictionary writers take into account the reviews on their dictionaries and the research into dictionary use, and why they ignore the suggestions made by researchers. Could it be that theoreticians have too high expectations of future dictionaries, or do they make suggestions which are impossible to put into practice? In addition, considerations outside of lexicography, such as market-related requirements, also have an influence on the making of a dictionary. The more commercial a dictionary is, the more influence "time, space and money" will have (p. 309).

Bielińska (p. 311) asserts that, if a dictionary clearly and explicitly indicates *where* the information can be found in the different metatexts, users will notice that, and will also start using the metatexts according to their purposes.

Bielińska's insistent requirement that there should be information for *researchers* of dictionaries in the metatexts (p. 311-312), makes one wonder whether that is really necessary. Although this may be an ideal situation, commercial interests of the publishing houses may always regard this as unnecessary. She therefore suggests that one might publish these theoretical expositions in a separate booklet or journal article, and just refer to this publication in the list of sources (p. 313), so that researchers can still have access to the information, without the dictionary becoming too thick and heavy.

It needs to be mentioned that Bielińska has an impressive list of references at the end of her book, which provides evidence of the thorough research that she has undertaken. In addition, she gives excellent theoretical overviews of matters regarding dictionary planning, especially on dictionary use and outer texts, and her use of lexicographical terminology is precise and consistent.

Even though Bielińska uses the acknowledged German metalexigraphical terminology, this book is written in straight-forward German. Her analyses are accompanied by comprehensible and user-friendly tables containing pluses, minuses and zeros indicating whether a particular feature is present in a dictionary or not. The book is a must-read for lexicographers compiling monolingual learners' dictionaries for foreign students. When used as a checklist of aspects in the metatexts, it could help improve the usefulness of such dictionaries for potential users.

References

- Bergenholtz, H. and S. Tarp.** 2005. Verteilungsstrukturen in Wörterbüchern. Barz, I., H. Bergenholtz and J. Korhonen (Eds.). 2005. *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*: 119-126. Frankfurt am Main: PeterLang.
- Gouws, R.H.** 2004. Outer Texts in Bilingual Dictionaries. *Lexikos* 14: 67-88.
- Hartmann, R.R.K.** 2000. European Dictionary Culture. The Exeter Case Study of Dictionary Use among University Students, against the Wider Context of the Reports and Recommendations of the Thematic Network Project in the Area of Languages (1996–1999). Heid, U. et al. (Eds.). 2000. *Proceedings of the Ninth Euralex International Congress, EURALEX 2000, Stuttgart, Germany, August 8th–12th, 2000. Vol. 1*: 385-391. Stuttgart: Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.
- Kipfer, B.A.** 1987. Dictionaries and the Intermediate Student: Communicative Needs and the Development of User Reference Skills. Cowie, A.P. (Ed.). 1987. *The Dictionary and the Language Learner. Papers from the EURALEX Seminar at the University of Leeds. 1–3 April 1985*: 44-54. Tübingen: Niemeyer.
- Köster, L. and F. Neubauer.** 1994. Langenscheidts Grosswörterbuch Deutsch als Fremdsprache und seine Benutzer. Henrici, G. and E. Zöfgen (Eds.). 1994. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des "de Gruyter Wörterbuchs Deutsch als Fremdsprache"*: 283-310. Tübingen: Niemeyer.
- Landau, Sidney I.** 2001²/1989. *Dictionaries. The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Langenscheidt** 1993 = Götz, D., G. Haensch and H. Wellmann (Eds.). 1993. Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Das neue einsprachige Wörterbuch für Deutschlernende. Berlin/Munich: Langenscheidt.
- Langenscheidt** 2008 = Götz, D., G. Haensch and H. Wellmann (Eds.). 2008. Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Das einsprachige Lernerwörterbuch für Einsteiger. Berlin: Langenscheidt.
- Ripfel, M.** 1989. Ergebnisse einer Befragung zur Benutzung ein- und zweisprachiger Wörterbücher. *Lexicographica* 1: 178-201.
- Ripfel, M. and H.E. Wiegand.** 1988. Wörterbuchbenutzungsforschung. Ein kritischer Bericht. Wiegand, H.E. (Ed.). 1988. *Studien zur neuhochdeutschen Lexikographie VI. Band 2*: 491-520. Hildesheim [et al.]: Georg Olms.
- Svartvik, J.** 1999. Corpora and Dictionaries. Herbst, T. and K. Popp (Eds.): *The Perfect Learners' Dictionary*: 491-520. Tübingen: Niemeyer.
- Wörterbuch Deutsch als Fremdsprache** 2000 = Kempcke, G. et al. (Eds.). 2000. *De Gruyter Wörterbuch Deutsch als Fremdsprache*. Berlin/New York: Walter de Gruyter.

Maria Smit
Independent Lexicographer and Language Practitioner
Stellenbosch
Republic of South Africa
(ria.eden@gmail.com)

Heberto H. Fernández Urdaneta. *Dictionaries in Spanish and English from 1554 to 1740: Their Structure and Development.* 2010, 401 pp. ISBN 84-96695-45-X (Paperback). Vertere. Monográficos de la Revista Hermēneus, Número 12. Saragossa: Pórtico Librerías, S.A. Price: €30.

This book is the twelfth in the collection of monographs Vertere, dealing with translation and translation studies. It is a revised, augmented edition of Fernández's dissertation submitted to the Department of Linguistics and Translation at the University of Montreal in April 2007.

The revival of learning during the Renaissance stimulated the compilation of grammars and dictionaries which, because of the invention of printing, could be circulated more easily and freely among travellers and merchants, in this way reaching larger groups of users. Spanish and English bilingual lexicography was an important tool in the teaching of Spanish in Tudor England.

The question Fernández sets out to answer is what can be deduced from the structure of early alphabetical and topical Spanish and English dictionaries, and their outside matter texts, about the compilation principles lexicographers followed and the purpose of their works. The investigation led first to a structural typology of these publications showing how the overall organization of topical dictionaries changed only slightly, while the position of grammars and alphabetical dictionaries was reversed. Both grammars and lexicographical products were used for pedagogical purposes. However, all the authors do not follow the same pedagogical approach, which influences the organization of the component parts of a specific author's work. Topics discussed by Spanish and English lexicographers comprise three categories: metalexicographical topics dealing with lexicography and the organization, function, intended users and features of the dictionary; metalinguistic topics treating phonetics, morphology, loan words and etymology; and extralinguistic topics informing about the genesis and printing of the dictionary, and the socio-political circumstances under which it was compiled.

The historical frame of the study extends from the two anonymous compilations of 1554 to the 1740 dictionary of Pedro Pineda. The corpus contains all the major and minor general bilingual Spanish and English dictionaries, alphabetically and topically arranged, published during this period. The body of the work therefore consists of a discussion of these dictionaries, grammars and vocabularies and their reprints and reissues: six from the sixteenth century, two from the seventeenth century and four from the eighteenth century. Of these twelve dictionaries twenty-two editions have been studied by Fernández.

Fernández follows an inductive method. By starting with an analysis of each work and then comparing it with preceding works, he arrives at a structural typology of the dictionaries and a classification of the subjects discussed in them. The organization of the sections dealing with each work are as follows: the discussion begins with introductory remarks of the compiler's life and work, followed by the genesis and compilation of the particular dictionary. After the

sources have been explored, the megastructure of the work is presented and the outside matter texts are listed. Then the macro- and microstructures are studied, comparing these with those of previous works. To make this comparison possible, samples from the beginning, middle and final parts of the dictionaries have been chosen. These samples also allow the verification of the prevailing ideas about a certain work. Short dictionaries and vocabularies could be studied in their entirety, but from larger ones samples had to be taken. When little or no research had been done on a certain work, aspects such as alphabetization, capitalization, use of articles, pronunciation and microstructural data are described. Studying the front and/or back matter texts last made a comparison of the subject matter in these texts in a particular dictionary with those in other dictionaries possible. In this way the interrelationships between the dictionaries as well as the evolution of the principles of the organization and compilation could be clarified. The analysis of the front matter texts reveal how the work was conceived and organized and compiled according to particular principles and points of view. Each discussion is rounded off by concluding remarks.

The book consists of seven chapters. In the first chapter the development of lexicography and the emergence of lexicography are reviewed, followed by an explanation of the fundamental concepts *megastructure*, *macrostructure* and *outside matter*, which make a uniform description of the constituent parts of the discussed works possible. A literature review is presented in chapter two, the first part dealing with the works of Roger J. Steiner and the second with other literature such as bibliographies, electronic resources and scholarly publications. Chapter three contains the problem statement, including the rationale for and objectives of the study and the corpus studied. Chapter four is historical, first explaining the role and influence of Antonio de Nebrija's works on modern lexicography in general and on Spanish and English lexicography in particular, and then giving an overview of the socio-political factors that led to the compilation of the first Spanish and English dictionaries during the second half of the sixteenth century. Chapters five, six and seven form the body of the book which deals with the corpus. Fernández classified the dictionaries first according to the century when they were published and then in order of the appearance of their first editions. Chapter five studies works from the sixteenth century: the two anonymous dictionaries from 1554, the *Book of English and Spanish* and the *Very Profitable Book to Learn English and Spanish*, together with those by John Thorius *The Spanish Grammar. With a Dictionarie adioyned vnto it* (1590), Richard Percyvall *Bibliotheca Hispanica. Containing a Grammar; with a Dictionarie in Spanish, English and Latine* (1591), William Stepney *The Spanish Schoole-master. Containing Seven Dialogues and a Vocabularie* (1591, 1619, 1620) and John Minsheu *A Dictionarie in Spanish and English* (1599, 1623). Chapter six discusses works from the seventeenth century: the short dictionary by Lewis Owen *The Key of the Spanish Tongue* (1605, 1606) and the polyglot dictionaries by Minsheu *Guide into the Tongues* (1617, 1625, 1626, 1627) and its companion *A*

Most Copious Spanish Dictionarie (1617). Chapter seven examines works from the eighteenth century: John Stevens' *A New Spanish and English Dictionary* (1706–05, 1726), the topical dictionaries by Félix Antonio de Alvarado *Spanish and English Dialogues. With the Construction of the Universe, and the Principal Terms of the Arts and Sciences* (1718, 1719) and Stevens *New Spanish Grammar. To Which Is Added, a Vocabulary of the Most Necessary Words* (1725, 1739), and Pedro Pineda's *New Dictionary, Spanish and English and English and Spanish* (1740).

The general conclusion summarizes the historical circumstances which led to the beginning of Spanish and English lexicography. Spanish and English integrated books based on their core structure and peripheral texts are typologized. The evolution of each structure type is traced according to the basic distinction between alphabetical and topical dictionaries. This is followed by a classification of the subjects lexicographers discussed in the prefatory texts. The chart at the end of this section gives a comprehensive and updated overview of Spanish and English bilingual lexicography. The book is concluded by a bibliography containing primary sources comprising dictionaries, grammars and vocabularies discussed in the book, and secondary sources consisting of electronic resources, monographs and papers dealing with these primary sources.

This study brings an important corrective on and extension of Roger J. Steiner's 1970 published book *Two Centuries of Spanish and English Bilingual Lexicography (1590–1800)*, which concentrates on lexicographical works in the alphabetical tradition. Works studied by Fernández for the first time are the two anonymous dictionaries from 1554, which were not known to Steiner, the short dictionary by Lewis Owen (1605), the second and third editions of William Stepney's vocabulary (1619, 1620), the two editions of Felix de Alvarado's nomenclator (1718, 1719) and the two editions of John Stevens' vocabulary (1725, 1739). An exception to Fernández's rule of not including multilingual works is Minsheu's 1617 polyglot dictionary and its companion, which bring a better and fuller understanding of Minsheu's lexicographical practice and influenced subsequent bilingual dictionaries. Also new is a discussion of the potential relationships between the 1617 two-part volume by Minsheu, subsequent editions of the polyglot dictionary only (1625, 1626 and 1627), and the Stevens dictionary of 1706–05. Furthermore a fuller analysis than Steiner's of the 1590 Thorius dictionary, the 1591 edition of Stepney's vocabulary and Minsheu's *Most Copious Dictionary* of 1617 is given. This study is therefore an informative addition to the literature dealing with the historical description and evaluation of Spanish and English bilingual lexicography.

J.C.M.D. du Plessis
Bureau of the Woordeboek van die Afrikaanse Taal
Stellenbosch, South-Africa
(jcmddp@sun.ac.za)

Pedro A. Fuertes-Olivera (Editor). *Specialised Dictionaries for Learners*. 2010, IX + 237 pp. ISBN 978-3-11-023132-8 (Hardback). Lexicographica. Series Maior 136. Berlin/New York: Walter de Gruyter. Price: €99.95.

Introduction

In learner lexicography, there has been a tradition of focusing on general-language learner's dictionaries, not least the well-known monolingual English learner's dictionaries. However, with the information and knowledge society, increasing attention is now being paid to various types of specialised dictionaries, including specialised dictionaries for learners. Many (language) learners need assistance from dictionaries not just to cope with situations that involve use of everyday language; very often, they have a need for dictionary assistance for coping with problems in the process of learning a specific LSP or a specific subject field. It is therefore essential that pedagogical LSP lexicography be promoted to a status that equals the more traditional lexicographic disciplines. This volume of articles is a major contribution to that endeavour.

Purpose of the book

Specialised Dictionaries for Learners is meant as a Festschrift for Enrique Alcaraz Varó, a Spanish linguist, translator and lexicographer, who died in 2008. As stated by the editor in the preface, the purpose of the book is firstly "[to defend] a *function-based transformative approach* centred on the dictionary and the users, investigating which lexicographical theories and principles are best suited for learners such as those enrolled on LSP courses and/or Translation degrees" (p. 2). Secondly, it "honours Enrique Alcaraz's pioneering visions and daily activities as a teacher, translator of specialised texts, and lexicographer" (p. 2). Later the editor adds that "the organization [of the book] adds to the unity of purpose of the book: to open up new lines of research in terms of the construction of pedagogically-oriented specialised dictionaries" (p. 2).

Contents of the book

The main sections of the book are eleven chapters dealing with aspects of pedagogical specialised lexicography. Part 1 (with 5 chapters) is devoted to the contribution of function theory to the development of specialised dictionaries for learners, Part 2 (with 4 chapters) deals with the contribution of linguistics to the development of specialised dictionaries for learners, and Part 3 (with 2 chapters) deals with future challenges for specialised learner's lexicography. The other sections of the book include a preface by the editor, a tribute to Prof.

Alcaraz written by Eva Samaniego Fernández, a selected bibliography of works by Prof. Alcaraz, an introduction by the editor describing the research presented in the individual chapters, a bibliography, notes on contributors, a name index and a subject index.

Chapter 1

Henning Bergenholtz / Sven Tarp: LSP Lexicography or Terminography? The Lexicographer's Point of View

Bergenholtz and Tarp start out by refuting the often-heard claim that terminographers and lexicographers are engaged in different types of tasks. The claim is that terminographers produce tools which are based (among others) on a systematic macrostructure, describe LSP terms, and are designed for text production, whereas lexicographers produce tools which are based (among others) on an alphabetical macrostructure, describe general-language words, and are designed for text reception and translation. According to Bergenholtz and Tarp, nothing could be further from the truth, since both those who call themselves terminographers and those who call themselves lexicographers are engaged in producing utility tools to satisfy specific lexicographically relevant needs of specific users in specific types of social situations.

To support their claims, Bergenholtz and Tarp first take the user situation of translation as an example. The translation process is complex and requires both cognitive and communicative skills of those who may become dictionary users in the translation process. For example, translators often need some basic knowledge of the specialised field related to the text to be translated. To cater to such a need, specialised dictionaries can be provided with a systematic introduction to the field in question, as is the case in for example the two dictionaries *English Gene Technology Dictionary* and *Spanish Gene Technology Dictionary*.

Bergenholtz and Tarp then take the user situation of reception as their second example. In this context, the authors argue that different types of users need different types of encyclopaedic information about entries. In this connection, users are divided into laypeople, semi-experts and experts. This is again illustrated with examples from the two above-mentioned dictionaries, with entries designed for laypeople on the one and semi-experts on the other hand, whereas experts will be unlikely to consult a lexicographical dictionary in a receptive user situation.

Although none of this, including the refutation of the false distinction between terminography and lexicography, is really new, the chapter provides a good overview of the central claims of the modern theory of lexicographical functions.

Chapter 2

Sven Tarp: Functions of Specialised Learners' Dictionaries

The purpose of Tarp's contribution is to present and discuss the possible functions of specialised learners' dictionaries "with the aim of presenting ideas for developing specialised learners' dictionaries of a new type, designed to assist users in both the learning of practical skills and the acquisition of knowledge about a specific subject field" (p. 39).

First, however, Tarp finds it necessary to define a learners' dictionary which is "a dictionary compiled with the genuine purpose of assisting learners in an on-going learning process" (p. 40).

Tarp then makes a very important distinction between two types of learning, although there are important relationships between them: learning a skill (either a communicative/linguistic skill or a practical skill) and acquiring knowledge. For example, consulting the data of a dictionary in order to improve one's communicative skills in a foreign language may add to one's knowledge of that particular foreign language. Similarly, practical skills can be improved either indirectly by consulting lexicographical tools that will transmit knowledge about a given subject field, or directly by consulting lexicographical tools that provide information which can be used to solve specific problems.

With respect to cognitive user situations, Tarp argues for two things: text books should be planned in such a way that they can satisfy not only global information needs but also punctual information needs, and dictionaries (perhaps specialised dictionaries in particular) should be arranged in such a way that they give a systematic overview of or introduction to a specific subject field.

More than giving concrete and detailed recommendations for the structure and contents of "specialised learners' dictionaries of a new type, designed to assist users in both the learning of practical skills and the acquisition of knowledge about a specific subject field", Tarp's (rather abstract) contribution is more a scene-setter for discussions of the practical achievements of these goals.

Chapter 3

Rufus H. Gouws: The Monolingual Specialised Dictionary for Learners

Gouws's contribution starts by pointing to an essential aspect of the concept of 'learner' in the context of specialised dictionaries for learners, namely that 'learner' can be understood in two ways: as "learner of the language in which the dictionary is presented or learner of the subject field treated in the dictionary" (p. 56). This is a very important distinction to have in mind when planning specialised dictionaries for learners, and Gouws gives examples of dictionaries which have successfully taken account of this distinction.

Using this distinction, it is also possible to "classify" both of these two types of learners, language learners into beginners, intermediate and advanced, and subject field learners into laypeople, semi-experts and experts.

The rest of the chapter is devoted to a thorough discussion of a large number of structural parameters, which, if carefully considered in the dictionary planning process, will "ensure successful dictionary consultation procedures by enhancing the access to the desired data and ensuring an optimal retrieval of information by the intended target user" (p. 60).

The structures discussed include the data distribution structure, access structures, the microstructure, definitions and data on grammar, although specialised dictionaries, including those for learners, will usually exclude grammatical data since their functions are normally receptive and cognitive.

Gouws's contribution also includes a discussion of types of specialised monolingual dictionaries for learners (single-field, multi-field and subfield) and suggestions for the structural planning of (normally multi-field) dictionaries that could be integrated with other learning material into a study package.

As can be seen, Gouws's treatment of the requirements for the optimal specialised dictionary for learners is quite comprehensive, and, although practical examples of the recommendations offered are scarce (or actually non-existent), planners of such dictionaries will find here invaluable guidance to their decision-making processes.

Chapter 4

Sandro Nielsen: Specialised Translation Dictionaries for Learners

As the title of his contribution indicates, Nielsen deals with specialised translation dictionaries for learners. In LSP translation, focus has traditionally been on terms and terminology, but noting with Laurén (1993: 99-100) that "terms generally make up less than 20 percent of any LSP text", Nielsen calls for increased attention to the translation of what he describes as domain-specific syntactic structures and genre conventions. His main contention is that specialised translation dictionaries for learners "should be augmented reference tools" (p. 69) by incorporating data on language- and domain-specific syntactic structures and genre conventions.

He has to admit, however, that incorporating such data into multi-field specialised dictionaries would result in a much less thorough treatment of relevant syntactic and genre-related LSP phenomena, than what would be possible in a single-field specialised dictionary.

Nielsen gives a few examples of these LSP phenomena. For syntax, his illustration is the structure of a particular noun phrase found for example in legal genres in Danish and German, but structurally impossible in, for example, English and French. The ideal "augmented reference tool" would contain an

integrated outside matter section on all such language- and genre-specific syntactic structures.

There is no doubt that the users of Nielsen's augmented reference tool need such syntactic and genre-related data. However, there are no suggestions in Nielsen's contribution as to how these syntactic structures and genre conventions could be identified and presented. All that we are given is an index for such a section on p. 80.

Chapter 5

Ildikó Fata: The Bilingual Specialised Translation Dictionary for Learners

Fata's contribution first outlines the metalexical basis for the planning and production of the ideal specialised LSP dictionary for translation purposes. Noting that this requires an interdisciplinary approach based on a discussion of the translation process (as outlined in Tarp 2007), Fata concludes that if a translation-oriented specialised dictionary is to be of assistance in the translation process, it must "combine knowledge of specialised language and knowledge of the system, and convey structured knowledge of the given domain, in this way helping the translator through the various phases of translation" (p. 88).

Fata then proceeds to a discussion of the place of the bilingual specialised dictionary in a typology of translation auxiliary materials (based on Nord 2002), noting that everybody seems to agree that existing bilingual LSP dictionaries are not suitable for translation purposes, since they generally ignore such relevant features as text, situation and culture.

Fata further lists (in quite some detail) the types of information needed in the ideal specialised dictionary for translation purposes. Here some exemplification might have been in place.

The final section of Fata's contribution is a project report on a translation-oriented bilingual LSP dictionary of *pension insurance*, the *Magyar–német, német–magyar nyugdíjbiztosítási szakszótár*. Noting first that the dictionary is intended for a very wide range of user groups (from Hungarian and German translators and interpreters to students of translating and interpreting as well as pension experts, social politicians, interested laypeople and future pensioners) with widely differing levels of professional knowledge on the one hand and linguistic knowledge on the other, Fata then lists the types of user situations (however, specifying under knowledge-oriented and operational-oriented user situations not user situations, but data types), concluding with a discussion of the pedagogical dimension of the dictionary.

Notwithstanding a few misunderstandings of aspects of the functional theory of lexicography, Fata seems to have conceived a dictionary solidly based on the theoretical framework of that theory.

Chapter 6

Aquilino Sánchez: The Treatment of Cultural and/or Encyclopaedic Items in Specialised Dictionaries for Learners

Sánchez notes: "Dictionaries [...] have rarely paid direct attention to the cultural dimension of words. The emphasis has been on the linguistic component disregarding the potential of cultural features for a more complete understanding of the language" (p. 111). Therefore he first emphasises the importance of including cultural elements in various types of dictionaries, including specialised bilingual dictionaries. He then provides a large number of examples where dictionaries will present the dictionary user with a full understanding of the entry word by including, in some way or another, a note explaining the cultural dimension(s) of the entry word in the individual dictionary articles.

This would apply not only to monolingual learner's dictionaries, but more importantly to bilingual dictionaries, since terms in two different languages are often part of differently structured semantic fields. A bilingual Inuit-English dictionary, for example, would not serve the needs of its users, if, for the many terms of *snow* in Inuit, it gave only the equivalent *snow* for all of them. This would not capture the fact that each individual term carries a unique extended meaning, specifying what kind of snow is covered by the term.

It is possible to find single instances of dictionaries trying to accomplish this, for example the *Longman Dictionary of English Language and Culture*, where we are told for example about the adjective *fat* that "in Britain and the US it is not considered attractive to be fat, and many people, esp. women, worry that they are too fat and try to lose weight". However, the very format for these cultural notes, a note attached as "an extra bonus" to the strictly linguistic definition, may have the consequence of reinforcing the perception that language and culture are two separate things.

When it comes to specialised lexicography, Sánchez stresses the need for inclusion of the cultural dimension of entry words and asserts that "specialised dictionaries have been more sensitive than general purpose dictionaries in the inclusion of cultural features as a necessary ingredient for grasping the meaning of the words" (p. 127).

However, Sánchez has to conclude that there have been few analyses and proposals in this area, and he himself offers very few, if any, concrete recommendations for the way in which cultural dimensions can be incorporated into any type of dictionary, for example regarding the format for such features.

Chapter 7

Geart van der Meer: The Treatment of Figurative Meaning in Specialised Dictionaries for Learners

Van der Meer investigates the feasibility of making users of specialised

learner's dictionaries aware of the link between the metaphorical use(s) and the literal use of a given word, claiming that this will give the dictionary user an understanding, comparable to that of a native speaker, of the full meaning of the word.

Examples taken mainly from the *Oxford Dictionary of Business English for Learners of English* demonstrate that no special effort has been made to link metaphorical use(s) of words to the literal use. In *Harrap's Dictionary of Business and Finance*, there are however occasional attempts to indicate such links implicitly as in the following definition of the metaphorical use of *bubble* where there are implicit links to the literal sense in the form of the verbs 'bursts' and 'exploded':

Industry or trend with no substance to it. A bubble usually bursts with more-or-less disastrous consequences for those financially involved. Probably the most famous bubble was the South Sea Bubble which exploded in 1720.

If a dictionary entry for the metaphorical sense of *bubble* should include an explicit link to the literal sense, van der Meer would formulate it as follows:

Industry or trend like buying shares that is like a real bubble (made of soap and water) in that it has no substance and is of short duration; like real bubbles, such bubbles are also said to *explode* or *burst*.

According to van der Meer, however, explicit links to literal senses in definitions of metaphorical senses would require rather much space and would also change the dictionary's character of being a specialised dictionary. He therefore recommends that specialised learner's dictionaries employ "sense definitions using a vocabulary (e.g. collocations) that at least strongly hints at the field of discourse from which the metaphor was originally taken" (p. 139). This also means that in cases where the literal sense of the word is not used or where the origin of the metaphorical sense is rather vague, the metaphorical sense will have to remain unexplained, a solution which the author finds unsatisfactory.

Chapter 8

Marie-Claude L'Homme: Designing Terminological Dictionaries for Learners Based on Lexical Semantics: The Representation of Actants

L'Homme reports on a project designed to transform a terminological database containing French terms relating to computing and the Internet into a specialised learner's dictionary.

L'Homme first notes that specialised dictionaries (including specialised dictionaries for learners) have focused on conceptual and encyclopaedic information, neglecting users' needs for data on the function of terms as linguistic units. Users need such data in order to solve for example communicative

problems in connection with (foreign) language production, (foreign) language reception, translation, etc.

In her contribution, L'Homme deals specifically with how to handle and represent actants (or arguments) mainly of verbs, adjectives and (deverbal) nouns "in a more user-friendly version" (p. 142) than the database.

L'Homme's terminological database, called DiConInfo, represents actants in a variety of forms, including actantial roles, which are not judged to be transparent to users of a learner's dictionary. L'Homme's suggestion for the specialised learner's dictionary, based on the DiConInfo database, is that actantial (or semantic) roles be supplemented with typical terms, instantiating the actants. For the verb *cliquer* (Eng. to click), it would mean that the semantic roles of Agent, Patient and Instrument are supplemented with typical terms for these actants, thus (p. 151):

CLIQUER 1, v.tr.

Structure actancielle: cliquer : **Agent** {utilisateur 1} ~ sur **Patient** {icône 1, fichier 1} avec **Instrument** {souris 1}

L'Homme's suggestions are very useful not only of course for the transformation of the DiConInfo database into a functionally adequate specialised learner's dictionary. It may also serve as a model for projects involving the transformation of functionally and user-type restricted databases of similar kinds to make them accessible to a wider user group with specific needs.

Chapter 9

Lynne Bowker: The Contribution of Corpus Linguistics to the Development of Specialised Dictionaries for Learners

Bowker's contribution starts out with an overview of a number of central lexicographical concepts such as specialised dictionary, learner's dictionary and specialised dictionary for learners in order to pave the way for a discussion of the contribution that corpus-based approaches have made to pedagogical lexicography in general and specialised lexicography in particular. This overview is very useful, because "these concepts do not necessarily mean the same thing to all people" (p. 155). In the discussion of specialised dictionaries for learners, Bowker mentions the hybrid learner's dictionaries, which combine data on general-language words and specialised terms.

Bowker then embarks on a thorough discussion of the contribution of corpus-based lexicography to learner's dictionaries. One fundamental aspect that receives special attention is the information about frequency that can be derived from a corpus-based approach, and which can be turned to good use in the structure of the dictionary, both at macro- and micro-level. Bowker mentions for example that the frequency data can be of assistance in the selection of lexical items to be included as headwords. More specifically, she claims that

"learners are typically interested in finding out about the words and phrases that are in common use, rather than those which are rare or obscure" (p. 163). Bowker also claims that frequency data can assist the lexicographer in deciding the order of meanings in cases where a lexical item has multiple meanings. The claim is here that "the majority of learners will typically be seeking information about the most common meaning of a word, so these meanings can be listed in an order based on frequency of occurrence in the corpus, with the most frequent meaning coming first" (p. 164).

In this reviewer's opinion, these claims can be questioned. It may be true that users of learner's dictionaries may more often seek data about the most frequent words and the most frequent meanings of words in situations involving language production, but when it comes to language reception, it may be argued that users (at least advanced users) will more often seek data about the less frequent words and less frequent meanings of words.

On the whole, however, the section provides us with very convincing arguments for relying on corpus-based approaches to lexicography in general and to learner's lexicography in particular.

Bowker ends with a discussion of the potential of a corpus-based approach for enhancing specialised dictionaries for learners, emphasising that compilers of such dictionaries can benefit from the insights gained in the use of corpora in compiling general-language learner's dictionaries.

Chapter 10

Zhang Yihua/Guo Qiping: An Ideal Specialised Lexicography for Learners in China Based on English-Chinese Specialised Dictionaries

Zhang and Guo start by noting a great need in China for bilingual specialised dictionaries for learners, since existing dictionaries of this kind cater only for the needs of professionals.

The authors have undertaken a study to investigate whether existing English–Chinese specialised dictionaries satisfy the needs of students who either attend compulsory courses in English for their specialty or who attend specialised courses taught in English. The results show that there is indeed a need for English–Chinese specialised learner's dictionaries that include data on matters such as pronunciation, examples, grammatical patterns, etc., all data types that are prominent in general learner's dictionaries.

Zhang and Guo outline a number of requirements for aspects of the ideal English–Chinese LSP dictionary using first a discussion of distinguishing features between on the one hand specialised learner's dictionaries and specialised general dictionaries and between on the other hand specialised learner's dictionaries and general learner's dictionaries, and then a discussion of communicative and cognitive functions and structural features (megastructure, microstructure, data distribution structure and access structure).

To begin with, they outline three so-called definition principles: the multi-dimensional definition principle, based on the findings of cognitive semantics and requiring definitions to accurately represent their cognitive domain; the holistic definition principle, based on the findings of Frame Semantics and recognising that words are not isolated linguistic units, but rather systematically related, which should be reflected in their definitions; and finally the pertinence principle, stating that definitions should be written in conformity with the corresponding academic field. Much of this is based on linguistic-semantic theories and is not related to any reflections on actual user needs in this respect.

In the concluding section on translation strategies of culture-bound words, Zhang and Guo come up with what seem to be very valuable recommendations concerning cases where Chinese lacks an equivalent for a specific English term. Their recommended solution is to use various types of calque translations (semantic, morphological, and phonological plus combinations of these) although this would involve the lexicographer in the invention of new terms.

Chapter 11

Pedro A. Fuertes-Olivera: Lexicography for The Third Millennium: Free Institutional Internet Terminological Dictionaries for Learners

After an introduction giving a brief discussion of the present state of lexicography (and its relationship with terminology), Fuertes-Olivera reviews approaches to typologies for electronic dictionaries. De Schryver's three-step typology (De Schryver 2003) is supplemented with three further criteria: Who has compiled the dictionary? Is it free or accessible through subscription? For whom has the dictionary been compiled?, which leads Fuertes-Olivera to propose a distinction between two main types of Internet dictionaries (with subtypes): (a) institutional Internet reference works and (b) collective free multiple-language Internet reference works. Disregarding the latter, he further subclassifies the former into restricted institutional Internet dictionaries and free institutional Internet reference works, which are the ones focused on in the remainder of Fuertes-Olivera's contribution.

Using an analysis and criticism of the quality of a number of such reference works, Fuertes-Olivera makes a number of proposals for the construction of pedagogically-oriented, free institutional Internet terminological dictionaries. The central question the author deals with is "how we can reconcile theoretical demands related to different users and their specific needs in different social situations with the reality of free institutional Internet terminological dictionaries which have usually been compiled for all users, regardless of their needs and use situations" (p. 202).

The proposals fall into two classes: lexicographical requirements and Internet requirements. The lexicographical requirements deal with issues such

as lemma selection, indications of meaning, semantic relationships, grammatical information and access structure, whereas the Internet requirements deal with issues such as visibility, placement of search field, readability of articles, instant results, reduction in the number of results, etc. (based on Almind 2005).

All these requirements seem quite logical. However, they are characterised by Fuertes-Olivera as obligatory requirements. It is an open question whether they should apply in their entirety to each and every Internet reference work, or whether there is the possibility that they can be applied selectively according to the function, target user group, etc. of the individual reference work.

Conclusion

As a lexicographical discipline in its own right, specialised learner's lexicography is still in its infancy, and very few specialised lexicographical tools for learners have been produced so far. However, *Specialised Dictionaries for Learners* is a major contribution to the establishment of a theoretical foundation for this relatively new lexicographical discipline. As observed already, very few of the contributions offer much in the way of practical recommendations for the compilation of specialised lexicographical tools for learners, but this collection of articles certainly not only paves the way for continued theoretical discussions within the discipline, but also lays a strong theoretical basis for more practically-oriented endeavours to produce tools for those who may seek lexicographical assistance in cognitive, communicative or operational situations involving specialised language.

Bibliography

Dictionaries

- English Gene Technology Dictionary* = Kaufmann, Uwe and Henning Bergenholtz (Eds.) in cooperation with B. Stummann, S. Tarp, L. de la Rosa Marabet, G. la Serna Miranda, and N. la Serna Torres. 1998. *Encyclopedic Dictionary of Gene Technology. Volume 1. English (with Spanish Equivalents)*. Toronto: Lugus Libros LatinAmerica Inc.
- Harrap's Dictionary of Business and Finance* = *Harrap's Dictionary of Business and Finance*. 1988. London: Harrap's Reference.
- Longman Dictionary of English Language and Culture* = *Longman Dictionary of English Language and Culture*. 1992. London: Longman.
- Magyar-német, német-magyar nyugdíjbiztosítási szakszótár* = Fata, Ildikó. 2005. *Magyar-német, német-magyar nyugdíjbiztosítási szakszótár*. Szeged: Grimm.
- Oxford Dictionary of Business English for Learners of English* = Tuck, Allene (Ed.). 1993. *Oxford Dictionary of Business English for Learners of English*. Oxford: Oxford University Press.
- Spanish Gene Technology Dictionary* = Kaufmann, Uwe and Henning Bergenholtz (Eds.). 1998. *Diccionario Enciclopédico de Ingeniería Genética. Español (con equivalentes in inglés)*. Toronto: Lugus Libros LatinAmerica Inc.

Other literature

- Almind, Richard.** 2005. Designing Internet Dictionaries. *Hermes, Journal of Linguistics*: 34: 37-54.
- De Schryver, Gilles-Maurice.** 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.
- Laurén, Christer.** 1993. *Fackspråk. Form, innehåll, function*. Lund: Studenterlitteratur.
- Nord, Britta.** 2002. *Hilfsmittel beim Übersetzen: Eine empirische Studie zum Rechercheverhalten professioneller Übersetzer* (FASK/Johannes-Gutenberg-Universität: Reihe A, Abhandlungen und Sammelbände 32). Frankfurt am Main: Peter Lang.
- Tarp, Sven.** 2007. Qué requisito debe cumplir un diccionario de traducción del siglo XXI? Fuertes Olivera, Pedro A. (Ed.). 2007. *Problemas lingüísticos en la Traducción Especializada*: 227-256. Valladolid: Universidad de Valladolid.

Birger Andersen
Centre for Lexicography
Department of Business Communication
School of Business and Social Sciences
Aarhus University, Aarhus
Denmark
(ba@asb.dk)

Jan Posthumus. *Woordenboeken en hun lotgevallen: Studies over lexicografiese praktijk (ca. 1850 tot heden).* 2009, x + 331 pp. ISBN/EAN 978-90-8880-007-8 (Stichting Neerlandistiek VU) / ISBN/EAN 978-3-89323-534-6 (Nodus Publikationen) (Sagteband). Cahiers voor Taalkunde, nr. 28. Amsterdam: Stichting Neerlandistiek VU / Münster: Nodus Publikationen. Prys: €43.

Jan Posthumus. *Lexicografie in Nederland: Peilingen in de negentiende en twintigste eeuw.* 2009, xiv + 291 pp. ISBN/EAN 978-90-8880-022-1 (Stichting Neerlandistiek VU) / ISBN/EAN 948-3-89323-762-3 (Nodus Publikationen) (Sagteband). Uitgaven Stichting Neerlandistiek VU 62. Amsterdam: Stichting Neerlandistiek VU / Münster: Nodus Publikationen. Prys: €43.

Jan Posthumus het bekend geword vir sy navorsing oor bekende Nederlandse woordeboeke vanaf ongeveer die middel van die 19de eeu, en sy artikels daarvoor wat veral in die tydskrifte *Trefwoord* en *De Woordenaar* verskyn het. 'n Veertiental van hierdie artikels is in hersiene vorm versamel in *Woordenboeken en hun lotgevallen*. Twee langer artikels wat oor die leksikografiese aktiwiteite van die familie Calisch en die inhoud en inrigting van Koenen se *Verklarend Handwoordenboek* handel, kon nie in hierdie bundel ingesluit word nie. Hulle is, voorsien van 'n inleiding, saam met vier langer artikels uit *Woordenboeken en hun lotgevallen*, naamlik die een oor Kramer se Franse, Duitse en Engelse vertaalwoordeboeke, en die drie oor Prick van Wely se leksikografiese werk, byeengebring in *Lexicografie in Nederland* wat by die Vrije Universiteit, Amsterdam as 'n akademiese proefskrif aangebied is. Daar is dus 'n oorvleueling wat die inhoud van hierdie twee werke betref.

In die "Inleiding" tot *Lexicografie in Nederland* plaas Posthumus die studie in 'n wyer konteks. Nadat hy 'n beskouing van die bestaande oorsigte oor die Nederlandse leksikografie gegee het, bespreek hy die inhoud van F. Claes en P. Bakema se *Bibliography of Dutch Dictionaries* (1995). 'n Opsomming van die studies ingesluit in hierdie werk van Posthumus toon hul besondere bydrae tot die algemene Nederlandse leksikografiese gesprek.

Die belangrike rol van die uitgewer word aangedui in die oorsig van die sakeondernemings wat op die een of ander tyd 'n aansienlike verskeidenheid woordeboeke gepubliseer het. Ten slotte meld Posthumus dat sekere van die bekende woordeboeke wat bespreek word, nou skynbaar hul nuttigheids waarde begin verloor, wat moontlik ook 'n aanduiding van die beëindiging van 'n besondere tydperk in die Nederlandse leksikografie is.

Posthumus se boek bied 'n aantal gevallestudies aan wat sekere aspekte van die Nederlandse leksikografiese publikasieterrein gedurende die neëntiende en twintigste eeu belig. Die hoofstukke oor die leksikografiese aktiwiteite van die Calische, Kramers en Prick van Wely handel almal oor die oorsprong en ontwikkeling van Nederlandse tweetalige woordeboeke, 'n gebied wat tot dusver min aandag geniet het. Die slothoofstuk toon hoe Koenen se

Nederlandse verklarende handwoordeboek homself vir meer as 'n eeu bly handhaaf het.

Die eerste hoofstuk handel oor die leksikografiese aktiwiteite van die familie Calisch. Nadat die vader M.I. Calisch (1718–1842) die voortou geneem het met sy *Nouveau Dictionnaire Français–Hollandais et Hollandais–Français* (1830) het sy seun I.M. Calisch (1808–1884) en sy skoonseun N.S. Calisch (1819–1891) op nog uitgebreider skaal by die publikasie van woordeboeke betrokke geraak. Hulle het nie alleen opvolguitgawes van die Franse woordeboek versorg nie, maar onder die naam Calisch het ook woordeboeke vir Duits en Engels verskyn, en een vir Nederlands.

Die hoofstuk "Van Jaeger naar Kramers" beskryf die voorgeskiedenis van die woordeboeke wat Prick van Wely by Van Goor tot 'n hoë peil ontwikkel het. Dit skets die geleidelike omvorming van die sakwoordeboekie tot 'n volwaardige woordeboek, volledig met 'n beskrywing van die ingewikkelde reeks uitgawes wat daarvan verskyn het. Twee artikels wat hierby aansluit, is opgeneem in *Woordenboeken en hun lotgevallen*. In die artikel "Van Tauchnitz naar Jaeger" word getoon dat die oorsprong van hierdie woordeboeke lê in 'n drietal ouere sakwoordeboekies uit die fonds van die Duitse uitgewer Tauchnitz wat Van Goor tot sy eie voordeel gebruik het. 'n Portret van hierdie ondernemende uitgewer, die grondleër van die uitgewery wat sy naam sou kry, word geskets in die artikel "Gerrit Benjamin van Goor (1816–1871)" wat terselfdertyd 'n beeld gee van die destydse uitgewersbedryf.

In die opskrif van die driedelige studie oor die lewe en werk van F.P.H. Prick van Wely (1867–1926) word hy beskryf as "een gedreven woordenaar". Die eerste deel handel oor sy lewensloop en wetenskaplike ontwikkeling, terwyl die volgende twee dele gewy is aan sy leksikografiese werk bestaande uit die totstandbrenging van tweetalige woordeboeke. Die materiaal hiervoor het Posthumus verkry uit die korrespondensie van die uitgewer Van Goor. Behalwe dat dit die stryd met Van Goor se groot mededinger J.B. Wolters weerspieël, gee dit ook 'n insiggewende beeld van die Nederlandse woordeboekbedryf in die eerste kwart van die twintigste eeu.

In die laaste hoofstuk beskryf Posthumus die lotgevalle van 'n enkele werk, die eentalige *Verklarend Handwoordenboek der Nederlandsche Taal*, saamgestel deur M.J. Koenen (1847–1920). Dit toon hoe dit vir 'n eeu lank suksesvol in opeenvolgende uitgawes waarvoor verskillende redakteurs van wie H.J.E. Endepols (1877–1962) die bekendste geword het, in stand gehou is. Na die eerste verskyning in 1897 het Koenen gou die Nederlandse woordeboek geword wat sonder veel mededinging 'n gevestigde plek in die middelbare onderwys ingeneem het.

Lexicografie in Nederland word afgesluit met 'n lys publikasies waarna in die teks verwys word, 'n register van persoonsname, en 'n Engelse vertaling van die opsomming van die inhoud van die verskillende hoofstukke wat in die "Inleiding" gegee word.

Van die oorblywende artikels in *Woordenboeken en hun lotgevallen* wat nie hierbo vermeld is nie, handel die eerste twee oor bepaalde aspekte van die

woordeboekbedryf. In "Woordenboekcriminaliteit", 'n term afkomstig van die leksikograaf F.J. Hausmann, word getoon met watter metodes uitgewers en ook soms die leksikograwe wie se werk hulle publiseer, die koperspubliek op soms subtiele, soms openlik bedrieglike manier mislei. In die kritiese resensie-artikel "Een bibliografie met gebreken" van die *Bibliography of Dutch Dictionaries* (1995) word aangedui watter probleme voorkom by die totstandbrenging van 'n insiggewende beskrywende oorsig van Nederlandse woordeboeke en word verskillende aspekte genoem wat verbetering verdien.

'n Ondersoek in die argief van die uitgewer J.B. Wolters het Posthumus in staat gestel om gegewens oor twee kleiner woordeboekies van Koenen op te spoor. Die een beskryf die ontwikkelingsgang saam met 'n uiteensetting van die ondeursigtige nommering van die uitgawes van Koenen se *Verklarend Zakwoordenboekje* (1882–1902) waarvan daar uiteindelik drie miljoen eksemplare verkoop is. Die ander is 'n aanvulling hierby wat 'n beskrywing gee van 'n verdere tweetal kleiner woordeboekies van Koenen wat vir die onderwys bedoel was, naamlik die *Geïllustreerd Woordenboekje der Nederlandsche Taal* en die *Beknopt Handwoordenboek der Nederlandsche Taal*.

Die artikel oor C.R.C. Heckenrath (1867–1935), veral bekend as die samesteller van J.B. Wolters se Franse woordeboek, toon hom ook as iemand met 'n maatskaplike belangstelling. 'n Artikel wat inligting oor H. Picard (1810–1858) se lewe en werk verskaf, bring sy *Pocket Dictionary* nie net in verband met enkele soortgelykes uit die tweede helfte van die neëntiende eeu nie, maar toon ook hoe dit bygedra het tot die ontwikkeling van die tweetalige Japanse leksikografie. Die artikel "J.B. Wolters' woordenboeken klassieke talen" gee, benevens 'n inventaris van die nege uitgawes van hierdie woordeboeke, 'n beskrywing van hul doelstelling en gebruikaspekte, ook deur die aanwending van gegewens in die samestellers se voorwoorde.

Die slotartikel "Lexicografie van het Australiese Engels" handel oor die aard en status van Australiese Engels, en die weerspieëling daarvan in die woordeboeke bedoel vir die Australiese publiek.

Woordenboeken en hun lotgevallen word afgesluit met 'n lys verwysings, gevolg deur drie indekse: 'n indeks van woordeboeke, 'n indeks van persone, en 'n indeks van onderwerpe.

Saam vorm hierdie twee publikasies van Posthumus 'n belangrike bydrae tot die geskiedenis van veral die vertalende Nederlandse leksikografie. Ten minste een van die woordeboeke wat hy behandel, is in die meeste Nederlandse (en ook talle Afrikaanse!) huise te vinde. Hopelik sal hy sy navorsing verder voortsit om nog die onbeskrewe leksikografiese werk van D. Bomhoff (1792–1860) en S. de Bruin (1821–1901) te belig.

J.C.M.D. du Plessis
Buro van die Woordeboek van die Afrikaanse Taal
Stellenbosch, Suid-Afrika
(jcmddp@sun.ac.za)

Herbert Ernst Wiegand und M^a Teresa Fuentes Morán. *Estructuras lexicográficas. Aspectos centrales de una teoría de la forma del diccionario.* (Colección Lexicografía 2). 2010, 485 Seiten. ISBN 978-84-613-4428-4. (Taschenbuch). Granada: Ediciones Tragamato. Preis: €45.00.

Herbert Ernst Wiegand hat in den vergangenen Jahren — ja: Jahrzehnten — lexicographische Strukturkonzepte ausgearbeitet, weiterentwickelt, und seine Erkenntnisse in einer Vielzahl von Publikationen in unterschiedlichen Publikationsorganen, Ländern und Sprachen veröffentlicht. Lexicographische Strukturen beschreiben den Aufbau von Wörterbüchern und von Komponenten und Zusammenhängen innerhalb von Wörterbüchern. Kenntnis um die lexicographischen Strukturen ist für die Lexikographie essentiell, da der Wörterbuchaufbau eine wesentliche Rolle bei der Informationserschließung spielt und ein besseres Verständnis dieses Aufbaus zu einer verbesserten, benutzerfreundlicheren Darstellung der Daten im Wörterbuch führen kann. Mit dem rezensierten Werk legt Wiegand gemeinsam mit M^a Teresa Fuentes Morán eine in spanischer Sprache verfasste Monographie vor, deren Titel auf Deutsch lautet: *Lexicographische Strukturen. Zentrale Aspekte einer Theorie der Wörterbuchform*, und von der man erwarten kann, dass sie die thematisch relevanten Arbeiten und die wichtigsten Erkenntnisse zu lexicographischen Strukturen an einem Ort zusammenbringt und zusammenfasst.

Das Buch besteht nach einer einleitenden *Vorstellung (Presentación)* aus elf Kapiteln. In den ersten beiden Kapiteln werden relativ kurz übergreifende Termini eingeführt (*Wörterbuchgegenstand, Wörterbuchgegenstandsbereich*) und die Theorie der Wörterbuchform, in der die Strukturen ihren Platz haben, in einer allgemeinen Theorie der Lexikographie verortet. Die folgenden neun Kapitel behandeln jeweils unterschiedliche lexicographische Strukturen und dazugehörige Phänomene. Alle Ausführungen beziehen sich auf Printwörterbücher (vgl. S. 17), elektronische Wörterbücher werden nicht thematisiert.

In Kapitel 3 (24 S.) werden *das Wörterbuch als Textverbund* und die *Textverbundstrukturen* betrachtet. Noch ohne konkretes reales Vorbild werden alle zentralen Aspekte anhand der Komponenten eines fiktiven Wörterbuchs charakterisiert; aufbauend darauf wird ein Ausschnitt aus einer Typologie von Textverbundstrukturen gezeigt. Bevor die Textverbundstrukturen genauer beschrieben werden, erfolgt anhand der fiktiven Komponenten eine Erläuterung grundlegender Aspekte von Strukturen: Es wird verständlich eingeführt, was konkrete und abstrakte Strukturen sind und welche Mengen und welche (partitiven oder Präzedenz-)Relationen bekannt und definiert sein müssen, um hierarchische Strukturen zu konstruieren. Auch wird der Aufbau und die Verwendung von Baumgraphen, mit denen die hierarchischen Strukturen visualisiert werden können, knapp dargelegt und wesentliche Charakteristika von Typologien genannt. Damit vermittelt dieses Kapitel gleichzeitig allgemeine Kenntnisse, die auch in den Folgekapiteln wichtig sind.

Kapitel 4 (26 S.) behandelt *Artikelkonstituentenstrukturen* in kondensierten Wörterbuchartikeln, m.a.W. den Aufbau von Wörterbuchartikeln. Anhand eines konkreten (deutschsprachigen) Artikels („Bütte“ aus dem *Handwörterbuch der deutschen Gegenwartssprache*) wird die funktional-positionale Segmentation der Kommentare, der elementaren und nichtelementaren Angaben sowie der nichttypographischen Strukturanzeiger ausführlich veranschaulicht. Diese bildet die Grundlage für eine (konkrete oder abstrakte) hierarchische Darstellung des Aufbaus bzw. der Zusammengehörigkeiten im Artikel, wozu wieder mit partitiven Relationen und Präzedenzrelationen gearbeitet wird. Außerdem wird genauer als in Kap. 3 erläutert, welche weiteren Mengen benötigt werden, um Strukturen zu definieren — damit werden wiederum allgemeine Kenntnisse zu Strukturen vermittelt und vertieft. Einige kritische Anmerkungen, in denen Defekte in der Artikelkonstituentenstruktur konkreter Wörterbücher aufgezeigt werden, schließen das Kapitel ab.

An die Artikelkonstituentenstrukturen schließen sich in Kapitel 5 (71 S.) die *Mikrostrukturen* an, diese sind Teilstrukturen der Artikelkonstituentenstrukturen. Auch hier wird, analog zu den Artikelkonstituentenstrukturen und wiederum überwiegend mithilfe des Artikels „Bütte“, erklärt, wie Mikrostrukturen herausgearbeitet werden und welche Mengen und Relationen dafür notwendig sind. Zunächst sind die Erläuterungen auf Wörterbuchartikel zu monosemen Ausdrücken beschränkt, dann werden sie auf Artikel zu polysemen Ausdrücken ausgeweitet; integrierte und teilintegrierte Mikrostrukturen werden behandelt (keine nicht-integrierten Mikrostrukturen) und Hinweise zur grafischen Darstellung von Mikrostrukturbildern gegeben. An die Ausführungen zu einfachen Mikrostrukturen schließen Ausführungen zu erweiterten Mikrostrukturen mit einer Übersicht über deren unterschiedliche Subtypen an, bevor schließlich nach den reinen auch die hybriden Mikrostrukturen mit funktionalen Angabezusätzen behandelt werden.

Im Kontext der Angabezusätze werden die *Textarchitekturen* in Kapitel 5 bereits kurz angesprochen, in Kapitel 6 (26 S.) werden sie dann genauer beleuchtet und — diesmal überwiegend anhand spanischer Wörterbuchartikel — veranschaulicht; sie sind ein Aspekt der räumlichen Anordnung von Wörterbuchdaten. Eine Typologie der Textarchitekturen von Wörterbuchartikeln wird präsentiert, die notwendigen Mengen und Relationen (neu: textarchitektonische Oberhalb-/Unterhalb-Relationen) zur Beschreibung von Artikelkonstituenten- bzw. Mikroarchitekturen benannt und Hinweise für die grafische Darstellung von Architekturbildern gegeben. Eine kurze Darstellung der *Textverbundarchitekturen* beschließt das Kapitel.

Kapitel 7 (76 S.) behandelt *Adressierung und Adressierungsstrukturen* — den Aspekt, wie Daten im Wörterbuch aufeinander bezogen sind und bezogen werden müssen, um Informationen korrekt zu erschließen. Unterschiedliche Aspekte und Typen der lemmatischen, sublemmatischen und nicht-lemmatischen Adressierung werden gezeigt und die Adressierungsrelation zu anderen textuellen Relationen in Bezug gesetzt. Artikelinterne Adressierung wird

für Angaben und Angabezusätze erläutert; außerdem die Fernadressierung. Adressierungsstrukturen sind nicht-hierarchische Strukturen, was sie von den zuvor besprochenen Strukturen unterscheidet. Die Beispiele sind deutsch, spanisch und zweisprachig mit unterschiedlichen Sprachen.

In Kapitel 8 (27 S.) werden anhand von bereits bekannten Artikeln die *Angabenstrukturen* beleuchtet, in denen Artikelkonstituentenstrukturen bzw. Mikrostrukturen mit den Adressierungsstrukturen zusammengefasst werden. Der Aufbau reiner und hybrider sowie exhaustiver und nicht-exhaustiver Angabenstrukturen mit den zugehörigen Relationen und Relationstermen (partitiv, präzedentiv, oberhalb/unterhalb, Adressierung) wird für Artikel und Kommentare gezeigt.

Kapitel 9 (16 S.) hat die *Makrostrukturen* zum Thema. Nach einer Unterscheidung alphabetischer Makrostrukturen von alphabetischen äußeren Zugriffsstrukturen (über ihren genuinen Zweck, eine Aussage über die makrostrukturelle Abdeckung zu geben) werden zunächst formgeprägte (monoalphabetische sowie polyalphabetische) Makrostrukturen, dann nicht-formgeprägte (systematische und randomisierte) Makrostrukturen dargestellt. Ein Ausschnitt aus einer Typologie der Makrostrukturen beschließt das Kapitel.

Kapitel 10 (86 S.) behandelt *Zugriffsstrukturen* und *Zugriffspfade*. Nach einer einführenden Unterscheidung direkter und indirekter sowie mediostruktureller und nicht-mediostruktureller externer Akzessivität und interner Akzessivität werden zunächst die reinen äußeren Zugriffsstrukturen beschrieben — die alphabetischen am Beispiel des Alphabets des deutschen Schriftsystems unter Einbeziehung der Zeichen <Ä/ä>, <Ö/ö>, <Ü/ü> und <ß>, jedoch ohne Berücksichtigung spanischer Eigenheiten wie den Digraphen <ch> und <ll>, deren Status im Zugriffsalphabet in der jüngeren spanischen Lexikographie eine Änderung erfahren hat und die sich zur kontrastiven Darstellung angeboten hätten — und gemäß der Art der Anordnung ihrer Einträge typologisiert (glatt-, nischenalphabetisch); daran anschließend die hybriden äußeren Zugriffsstrukturen (nestalphabetische und desultorisch-alphabetische). Äußere Umtext-Zugriffsstrukturen werden ebenso abgehandelt wie Registerzugriffsstrukturen und der Zugriff über Seitenzahlen. Nach den äußeren folgen die inneren Zugriffsstrukturen und die Schnellzugriffsstrukturen, bevor abschließend die Zugriffspfade angesprochen werden, die keine statischen Strukturen sind, sondern die den individuellen Weg jedes individuellen Benutzers und Benutzungsvorgangs nachzeichnen.

Kapitel 11 (89 S.) schließlich handelt von den *Mediostrukturen*. Nach der Erläuterung der Verweisvoraussetzungen werden terminologische Fragen zur Unterscheidung zwischen den Textsegmenten (z.B. Verweisangabe, Verweissprungsangabe, Verweiskennzeichnung, Verweisbeziehungsangabe, Verweisadressenangabe) und den Relationen (z.B. Verweisrelation, Adressierungsrelation) im Kontext von Verweisungen sowie dem Ausdruck *Verweis/remisión* an sich (der hier, wie in anderen Arbeiten Wiegands, als kognitive, nicht als textuelle Einheit aufgefasst wird) geklärt. Jeder dieser Aspekte wird im Folgen-

den detaillierter behandelt, die Aspekte werden miteinander in Beziehung gebracht und teilweise durch Typologisierung strukturiert präsentiert. Schließlich werden einfache und erweiterte Mediostrukturen anhand deutscher und spanischer Wörterbuchartikel dargestellt.

Auf das Kapitel über Mediostrukturen folgt direkt das Literaturverzeichnis, es gibt kein abschließendes Kapitel (Schlusswort, Fazit, Ausblick o.Ä.).

Bei der Lektüre fällt ein Ungleichgewicht bei der Wahl der Beispiele auf. Die Verfasser des rezensierten Werks sind ein deutscher Wörterbuchforscher und eine spanische Wörterbuchforscherin. Es ist festzustellen, dass von den insgesamt rund 150 Abbildungen von Wörterbuchartikeln, Artikelstrecken oder anderen Ausschnitten aus Wörterbüchern, die als Beispiele oder zur Veranschaulichung herangezogen werden, etwa zwei Drittel aus einsprachigen deutschen Wörterbüchern stammen und nur etwas über ein Fünftel aus einsprachigen spanischen Wörterbüchern (die restlichen Abbildungen sind aus zweisprachigen oder anderssprachigen Wörterbüchern). Von den rund 40 Strukturdarstellungen, in denen (auch) konkrete Angaben genannt werden, beziehen sich ca. 60 % auf deutsche Artikel und nur ca. 36 % auf spanische Artikel (ca. 4 % andere/zweisprachige). Manche Beispielartikel werden in späteren Kapiteln wieder aufgegriffen, so dass ein Wiedererkennungseffekt eintritt. Dass relativ viele Phänomene an deutschsprachigen Beispielen erklärt werden, ist wissenschaftlich selbstverständlich legitim und lässt sich teilweise auch in der Sache begründen, da bestimmte Phänomene wie — um nur zwei Beispiele zu nennen — die funktionalen Angabezusätze zur Kennzeichnung des Wortakzents und der Vokalquantität oder die nichtelementare Morphologieangabe anhand deutscher Beispiele sehr gut veranschaulicht werden können (das Spanische hingegen hat relativ regelmäßige Betonungsregeln, deren Abweichungen in der Schriftform durch diakritische Akzentzeichen kenntlich gemacht werden, weshalb eine gesonderte Kennzeichnung im Wörterbuch durch „künstliche“ funktionale Angabezusätze der genannten Art nicht notwendig ist; auch die Bildung der Pluralformen von Substantiven — einen synthetisch gebildeten Genitiv gibt es im Spanischen nicht — ist relativ regelmäßig, so dass diese Angabe im spanischen Wörterbuch ebenfalls oft wegfällt). In der Darstellung geht es Wiegand und Fuentes Morán erkennbar um die auch auf Wörterbücher anderer Sprachen übertragbare *Methodik* der Segmentierung und der Ermittlung von Angaben, Angabezusätzen etc. und deren Klassen und nicht um einzelsprachliche Besonderheiten. Dieser Aspekt der Beispielwahl wird daher nicht negativ beurteilt — obwohl es für spanische Leser(innen) sicherlich angenehm wäre, die Entwicklung der Konzepte anhand von mehr Beispielen in der eigenen Sprache verfolgen zu können (vgl. auch die Anmerkungen zur den Ausführungen zum Zugriffsalphabet in Kap. 10).

Bedauerlich ist aber, dass zahlreiche der Abbildungen, insbesondere der grafischen Strukturdarstellungen, grafisch unschön präsentiert werden und/oder sogar Fehler aufweisen. Zu den formalen Mängeln gehört, dass manche Strukturdarstellungen aus Platzgründen (das Buchformat ist A5, der Satzspie-

gel entsprechend noch kleiner) nicht in einer angemessenen Größe abgedruckt werden, was dazu führt, dass mancher Text innerhalb der Darstellungen fast unleserlich klein wird oder dass textuelle Bezeichnungen von grafischen Linien (Rahmen, Kanten) überdeckt werden (z.B. Fig. 5-8, S. 93; 5-13, S. 104; 5-32, S. 123; 6-12, S. 162; 8-3, S. 257; 8-8, S. 267). Der schwerwiegendste Kritikpunkt ist jedoch, dass in einer nicht unwesentlichen Zahl von Strukturdarstellungen inkorrekte Bezeichnungen zu finden sind, wozu im Folgenden nur wenige Beispiele gegeben werden:¹

- In *Figura* 3-5 (S. 45) liegt vor: eine fehlerhafte Nummerierung beim letzten Typologiekriterium („1a-2“ müsste „1c-2“ heißen); ein Fehler in der Bezeichnung des zweiten Typologiekriteriums („SconTx“ müsste durch „SNM“ ersetzt werden); eine unvollständige Bezeichnung bei „t₆“ (hier liegt zusätzlich Rechtserweiterung vor).
- In *Figura* 4-6 (S. 64) sind die konkreten Angaben teilweise inkorrekt wiedergegeben: statt „|“ müsste es „-“ lauten, statt „n“ „-n“.
- In *Figura* 5-23 (S. 116) müsste das zweite Auftreten der Klassenbezeichnung „Ind-ForSing.c“ (kondensierte Angabe zur Bildung der Singularform), die zur Angabe „o. Pl.“ gehört, statt dessen „Ind-STan.c“ (kondensierte Angabe eines Singularetantum, vgl. S. 268) lauten; dahingehend müsste auch der linke tiefgestellte Index des folgenden Angabeblassymbols geändert werden.
- In *Figura* 8-7 (S. 265) sind die Polysemieangaben durchgehend inkorrekt mit „Ind-PS“ (*Indicación de paráfrasis de significado* ‚Bedeutungsparaphrasenangabe‘) anstelle „Ind-Pol“ beschriftet.

Solche Fehler müssen benannt werden, obwohl sie das Verständnis des fachlichen Inhalts nicht wesentlich erschweren oder gar verunmöglichen. Dies ist dem inhaltlichen Aufbau und der textuellen Darstellung zu verdanken, die sehr positiv herauszustellen sind: Das gesamte Buch ist ebenso wie die einzelnen Kapitel in sich schlüssig aufgebaut, die Ausführungen sind trotz der relativ komplexen metalexikographischen Fachterminologie verständlich formuliert und sehr gut und flüssig zu lesen — dies ist sicherlich auch das Verdienst von Fuentes Morán, die, wie der *Vorstellung* zu entnehmen ist, den von Wiegand auf Deutsch verfassten Text vervollständigt und ins Spanische übersetzt hat (vgl. S. 15). Obwohl über einige Strecken hinweg viel mit Variablen, Abkürzungen und formelhaften Ausdrücken gearbeitet wird, bleibt die Darstellung übersichtlich, da die neuen Ausdrücke deutlich eingeführt werden — lediglich wenn Abkürzungen für Angabeklassen, die in vorhergehenden Kapiteln eingeführt wurden, später unkommentiert wieder aufgegriffen werden, wünscht man sich ein entsprechendes Abkürzungsverzeichnis im Anhang. Grundlegende Kenntnisse der Mengenlehre und ihrer Formelsprache werden jedoch vorausgesetzt; auch ist für den deutschsprachigen Leser eine gewisse Vertrautheit

mit der von Wiegand (auf Deutsch) geprägten Terminologie der Wörterbuchforschung sicherlich vorteilhaft, wenn auch nicht erforderlich, um die einzelnen Teilaspekte besser zuordnen zu können.

Inhaltlich lässt sich zusammenfassend feststellen: Es handelt sich bei dem rezensierten Werk um eine aktuelle Darstellung der lexikographischen Strukturkonzepte nach Wiegand. Abweichende Konzepte oder frühere Ideen (auch von Wiegand selbst), die im Laufe der vergangenen Jahre zum heutigen Stand hin geändert oder verworfen wurden, werden kaum berücksichtigt. Selbst die bei Wiegand sonst fast obligatorischen Hinweise auf Josette Rey-Debove (Makro- und Mikrostruktur) oder „Nicolas Bourbaki“ (Strukturkonzept im Rahmen der Mengenlehre) fehlen hier; bei Letzterem ist lediglich die Rede vom „más moderno concepto matemático de estructura basado en la teoría de conjuntos“ (S. 27). Dieser Verzicht auf einen historischen Blickwinkel wird von Lesern, die eine Gesamtdarstellung erwarten, bedauert werden, er ermöglicht aber eine umso geradlinigere und stringentere Darstellung der zentralen Aspekte nach aktuellem Stand. Diese gelingt hier vortrefflich: Die Lektüre erschließt dem Leser/der Leserin ein breites Wissen über die unterschiedlichen lexikographischen Strukturen und darüber, wie diese aufeinander zu beziehen sind. Gemeinsam mit den oben genannten Punkten des schlüssigen Aufbaus und der klaren Formulierung führt dies zu einer unbedingt positiven Gesamtbeurteilung.

Als Adressaten, die von der Lektüre des vorliegenden Werks profitieren werden, sind zunächst die Experten in der lexikographischen Praxis und der theoretischen Wörterbuchforschung zu nennen — erstere, weil ihnen in den *Estructuras lexicográficas* konkrete Hinweise zu unterschiedlichen Aufbaumöglichkeiten von Wörterbuchartikeln, zu den Bezügen der Daten in Wörterbüchern und etwa auch Hinweise auf Fehler bei der Strukturierung gegeben werden, letztere, weil ihnen nun eine zusammenfassende Darstellung eines wichtigen Teilbereichs der Wörterbuchforschung in systematischer Form vorliegt, auf deren Grundlage weitere Forschungen aufbauen können. Als weitere Adressatengruppe können engagierte Semiexperten — nämlich: Studierende — genannt werden; engagiert, da sie bereit sein müssen, sich in eine stellenweise komplexe und im Spanischen in vielen Bereichen neue Terminologie einzuarbeiten, wobei sie aber mit der wie erwähnt gut verständlichen Präsentation starke Unterstützung finden.

Das umfangreiche terminologische System, das von Wiegand und Kolleginnen und Kollegen im Bereich der lexikographischen Strukturen entwickelt wurde, wird derzeit auch im *Wörterbuch zur Lexikographie und Wörterbuchforschung (WLWF)*² beschrieben, so dass bald in deutscher (und englischer) Sprache ebenfalls eine vollständige und zentrale Quelle für das systematische Nachschlagen und Recherchieren von Termini und lexikographischem Fachwissen (nicht nur) zu Strukturen zur Verfügung stehen wird. Dennoch wäre eine verständliche, ausführliche, aber kompakt in einem Band zusammengefasste Darstellung des Teilbereichs der lexikographischen Strukturen, wie sie

mit dem hier rezensierten Werk auf Spanisch vorliegt, aus der Sicht des deutschen Muttersprachlers auch für die deutsche Wörterbuchforschung — in deutscher Sprache — sehr zu wünschen; sie wäre für die Forschung, für die Praxis und auch für die Lehre ein großer Gewinn — ein Gewinn, zu dem der spanischen Lexikographie schon jetzt zu gratulieren ist.

Anmerkungen

1. Weitere vergleichbare Fehler sind etwa zu finden in den *Figuras* 4-3 (S. 51), 5-8 (S. 93), 5-28 (S. 119), 5-35 (S. 126), 5-41 (S. 138), 5-44 (S. 143), 6-7 (S. 157), 6-12 (S. 162), 8-10 (S. 270), 11-22 (S. 432), 11-32 (S. 450).
2. Wörterbuch zur Lexikographie und Wörterbuchforschung. Dictionary of Lexicography and Dictionary Research. Mit englischen Übersetzungen der Umtexte und Definitionen sowie Äquivalenten in neun Sprachen. Hrsg. u. bearb. v. Herbert Ernst Wiegand, Michael Beißwenger, Rufus H. Gouws, Matthias Kammerer, Angelika Storrer, Werner Wolski. 4 Bde. Bd. 1: Systematische Einführung, A–C. Berlin/New York 2010.

Michael Mann
Friedrich-Alexander-Universität Erlangen-Nürnberg
Department Germanistik und Komparatistik
Bismarckstr. 1
D-91054 Erlangen
(Michael.Mann@ger.phil.uni-erlangen.de)

Publikasieaankondigings / Publication Announcements

Anne-Marie Beukes en Marné Pienaar. *Veeltalige vertaalterminologie./Multilingual Translation Terminology*. 2010, xvi + 256 pp. ISBN 978 0 627 02815 1 (Sagteband). Pretoria: Van Schaik Uitgewers/Publishers. Prys: R239,95.

Monika Bielińska. *Lexikographische Metatexte. Eine Untersuchung nichtintegrierter Außentexte in einsprachigen Wörterbüchern des Deutschen als Fremdsprache*. 2010, 375 pp. Danziger Beiträge zur Germanistik 32. ISBN 978-3-631-60500-4 (Hardback). Frankfurt am Main: Peter Lang. Price: €61.50. (Review in this issue.)

A.J. Brink en J. de V. Lochner. *Woordeboek vir die Gesondheidswetenskappe/Dictionary for the Health Sciences*. Tweede hersiene en bygewerkte uitgawe. 2011, xxii + 1191 pp. ISBN 978-1-86890-113-5 (Hardeband). Kaapstad: Pharos Woordeboeke. Prys: R550,95.

Aletta Elizabeth Cloete. *Basic Afrikaans. The Top 1000 Words and Phrases. Afrikaans–English*. 2011, vi + 121 pp. ISBN 978-1-86890-116-6 (Paperback). Cape Town: Pharos Dictionaries. Price: R99,95.

Julie Coleman. *A History of Cant and Slang Dictionaries Volume IV 1937–1984*. 2010, 536 pp. ISBN 978-0-19-956725-6 (Hardback). Oxford: Oxford University Press. Price: £75.

Johanna de Wet (Samesteller). *Wat praat jy! Afrikaanse idioome, gesegdes en vaste uitdrukkings met verklarings*. Vierde uitgawe. 2010, x + 496 pp. ISBN 978-1-86919-388-1 (Sagteband). Pretoria: Protea Boekhuis. Prys: R195.

Lucien Ditougou. *On Est Ensemble: 852 mots pour comprendre le français du Gabon*. Préface par Guy Rossatanga-Rignault. 2009, 154 pp. ISBN 978-2-912776-85-3. Hommes et sociétés. Libreville: Éditions Raponda Walker. (Article de revue dans ce numéro.)

Heberto H. Fernández Urdaneta. *Dictionaries in Spanish and English from 1554 to 1740: Their Structure and Development*. 2010, 401 pp. ISBN 84-96695-45-X (Paperback). Vertere. Monográficos de la Revista Hermēneus, Número 12. Saragossa: Pórtico Librerías, S.A. Price: €30. (Review in this issue.)

Pedro A. Fuertes-Olivera (Editor). *Specialised Dictionaries for Learners*. 2010, IX + 237 pp. ISBN 978-3-11-023132-8 (Hardback). Lexicographica. Series Maior 136. Berlin/New York: Walter de Gruyter. Price: €99.95. (Review in this issue.)

NgoSemzara Kabuta. *Nkòngamyakù Cilubà–Mfwàlànsa*. 2008, XI + 365 pp. ISBN 2-930393-07-06. Gand: Recall. (Article de revue dans ce numéro.)

F.J. Labuschagne en L.C. Eksteen. *Verklarende Afrikaanse woordeboek*. Negende, hersiene uitgawe. 2010, xiv + 1580 pp. ISBN 978-1-86890-100-5 (Hardeband). Kaapstad: Pharos Woordeboeke. Prys: R350.

Paul Achille Mavoungou. *Lexicographie et Confection des Dictionnaires au Gabon*. 2010, 356 pp. ISBN 978-1-920109-89-9 (Paperback). Stellenbosch: SUN PRESS/SUN MeDIA. Price: R200.

Paul Achille Mavoungou et Bernard Plumel. *Dictionnaire Yilumbu–Français*. 2010, 801 pp. ISBN 978-2-912776-98-3 (Paperback). Libreville: Éditions Raponda-Walker.

Multilingual Soccer Terminology List English/Afrikaans/IsiZulu/IsiXhosa/Siswati/IsiNdebele/Setswana/Sepedi/Sesotho/Tshivenda/Xitsonga. 2010, iv + 240 pp. No ISBN (Paperback). Department of Arts and Culture, RSA. Free for personal or educational use at http://www.dac.gov.za/doc/2010/Soccer_Terminology_2010.pdf

Jan Posthumus. *Lexicografie in Nederland: Peilingen in de negentiende en twintigste eeuw*. 2009, xiv + 291 pp. ISBN/EAN 978-90-8880-022-1 (Stichting Neerlandistiek VU) / ISBN/EAN 948-3-89323-762-3 (Nodus Publikationen) (Sagteband). Uitgaven Stichting Neerlandistiek VU 62. Amsterdam: Stichting Neerlandistiek VU / Münster: Nodus Publikationen. Prys: €43. (Resensie in hierdie uitgawe.)

Jan Posthumus. *Woordenboeken en hun lotgevallen: Studies over lexicografische praktijk (ca. 1850 tot heden)*. 2009, x + 331 pp. ISBN/EAN 978-90-8880-007-8 (Stichting Neerlandistiek VU) / ISBN/EAN 978-3-89323-534-6 (Nodus Publikationen) (Sagteband). Cahiers voor Taalkunde, nr. 28. Amsterdam: Stichting Neerlandistiek VU / Münster: Nodus Publikationen. Prys: €43. (Resensie in hierdie uitgawe.)

Anton F. Prinsloo. *Die aap in jou koffie. Afrikaanse eponieme van A tot Z*. 2011, 351 pp. ISBN 978-1-86919-463-5 (Sagteband). Pretoria: Protea Boekhuis. Prys: R200.

Pumlani Sibula. *Economic and Management Sciences Terminology/Ekonomiese en Bestuurswetenskappe Terminologie/isiGama sobuNzululwazi kwezoQoqosho NoLawulo*. 2009, vi + 383 pp. ISBN 978-0-620-44475-0 (Paperback). Stellenbosch: IZiko leLwimi le Univesithi yaseStellenbosch/Stellenbosch University Language Centre/Universiteit Stellenbosch Taalsentrum.

Pumlani Sibula. *IsiGama seSoka/Soccer Terminology/Sokker-Terminologie*. 2010, iv + 46 pp. ISBN 978-0-620-46837-4 (Paperback). Stellenbosch: IZiko leeLwimi le Univesithi yaseStellenbosch/Stellenbosch University Language Centre/ Universiteit Stellenbosch Taalsentrum.

Wiegand, Herbert Ernst und M^a Teresa Fuentes Morán. *Estructuras lexicográficas. Aspectos centrales de una teoría de la forma del diccionario*. (Colección Lexicografía 2). 2010, 485 Seiten. ISBN 978-84-613-4428-4 (Taschenbuch). Granada: Ediciones Tragacanto. Preis: €45. (Rezension in dieser Ausgabe.)

VOORSKRIFTE AAN SKRYWERS

(Tree asseblief met ons in verbinding (lexikos@sun.ac.za) vir 'n uitvoeriger weergawe van hierdie instruksies of besoek ons webblad: <http://www.wat.co.za>)

A. REDAKSIONELE BELEID

1. Aard en inhoud van artikels

Artikels kan handel oor die suiwer leksikografie of oor implikasies wat aanverwante terreine, bv. linguistiek, algemene taalwetenskap, rekenaarwetenskap en bestuurskunde vir die leksikografie het.

Bydraes kan onder enigeen van die volgende rubrieke geklassifiseer word:

(1) **Artikels:** Grondige oorspronklike wetenskaplike navorsing wat gedoen en die resultate wat verkry is, of bestaande navorsingsresultate en ander feite wat op 'n oorspronklike wyse oorsigtelik, interpreterend, vergelykend of krities evalueerend aangebied word.

(2) **Resensieartikels:** Navorsingsartikels wat in die vorm van 'n kritiese resensie van een of meer gepubliseerde wetenskaplike bronne aangebied word.

Bydraes in kategorieë (1) en (2) word aan streng anonieme keuring deur onafhanklike akademiese vakgenote onderwerp ten einde die internasionale navorsingsgehalte daarvan te verseker.

(3) **Resensies:** 'n Ontleding en kritiese evaluering van gepubliseerde wetenskaplike bronne en produkte, soos boeke en rekenaarprogramme.

(4) **Projekte:** Besprekings van leksikografiese projekte.

(5) **Leksikonotas:** Enige artikel wat praktykgerigte inligting, voorstelle, probleme, vrae, kommentaar en oplossings betreffende die leksikografie bevat.

(6) **Leksikovaria:** Enigee van 'n groot verskeidenheid artikels, aankondigings en nuusvrystellings van leksikografiese verenigings wat veral vir die praktiserende leksikograaf van waarde sal wees.

(7) **Verslae:** Verslae van konferensies en werksessies.

Bydraes in kategorieë (3)-(7) moet almal aan die eise van akademiese geskifte voldoen en word met die oog hierop deur die redaksie gekeur.

2. Wetenskaplike standaard en keuringsprosedure

Lexikos is deur die Departement van Onderwys van die Suid-Afrikaanse Regering as 'n gesubsidieerde d.w.s. inkomstegenererende navorsingstydskrif goedgekeur.

Artikels sal op grond van die volgende beoordeel word: taal en styl; saaklikheid en verstaanbaarheid; probleemstelling, beredenering en gevolgtrekking; verwysing na die belangrikste en jongste literatuur; wesenlike bydrae tot die spesifieke vakgebied.

3. Taal van bydraes

Afrikaans, Duits, Engels, Frans of Nederlands.

4. Kopiereg

Nóg die Buro van die WAT nóg die African Association for Lexicography (AFRILEX) aanvaar enige aanspreeklikheid vir eise wat uit meewerkende skrywers se gebruik van materiaal uit ander bronne mag spruit.

Outeursreg op alle materiaal wat in *Lexikos* gepubliseer is, berus by die Direksie van die Woordboek van die Afrikaanse Taal. Dit staan skrywers egter vry om hulle materiaal elders te gebruik mits *Lexikos* (AFRILEX-reeks) erken word as die oorspronklike publikasiebron.

5. Oorspronklikheid

Slegs oorspronklike werk sal vir opname oorweeg word. Skrywers dra die volle verantwoordelikheid vir die oorspronklikheid en feitelike inhoud van hulle publikasies.

6. Gratis oordrukke en eksemplare

Skrywers ontvang vyf gratis oordrukke van elke artikel of resensieartikel van hulle wat gepubliseer is asook een gratis eksemplaar van die uitgawe waarin sodanige artikel(s) verskyn het. Skrywers van

suiwer evaluerende resensies en van bydraes tot die rubrieke Leksikonotas, Leksikovaria, Projekte en Verslae ontvang vyf gratis oordrukke van hulle bydraes. In laasgenoemde vier kategorieë kan die redaksie egter, afhangend van die aard en omvang van die bydraes, besluit om ook 'n eksemplaar van die betrokke uitgawe aan 'n skrywer toe te ken.

7. Uitnodiging en redaksionele adres

Alle belangstellende skrywers is welkom om bydraes vir opname in *Lexikos* te lewer en aan die volgende adres te stuur:

Die Redakteur: LEXIKOS
Buro van die WAT
Posbus 245
7599 STELLENBOSCH
Republiek van Suid-Afrika

B. VOORBEREIDING VAN MANUSKRIP

Die manuskrip van artikels moet aan die volgende redaksionele vereistes voldoen:

1. Lengte en formaat van artikels

Bydraes moet verkieslik nie 20 getikte A4-bladsye met teks in dubbelspasiëring en ruim kantlyne (ongeveer 2,5 cm) oorskry nie. Manuskrip moet verkieslik in elektroniese formaat as ASCII-teks, as volledig geformateerde Microsoft Word (DOS of Windows) lêers of as WordPerfect (DOS of Windows) lêers op rekenaarskyf (360 KB tot 1.44 MB) voorgelê word. 'n Rekenaardrukstuk van die artikel moet die skyf vergesel. Elke artikel moet voorsien wees van 'n Engelse opsomming van tussen 150 en 250 woorde, sowel as tussen 10 en 30 Engelse sleutelwoorde.

2. Grafika

Een stel duidelike oorspronklike illustrasies, tabelle, grafieke, diagramme, of kwaliteitsafdrukke daarvan, moet voorgelê word. Die plasing van grafika binne die teks moet duidelik aangedui word.

3. Bibliografiese gegewens en verwysings binne die teks

Kyk na onlangse nommers van *Lexikos* vir meer inligting.

INSTRUCTIONS TO AUTHORS

(For a more detailed version of these instructions, please contact us (lexikos@sun.ac.za)
or refer to our website: <http://www.wat.co.za>)

A. EDITORIAL POLICY

1. Type and content of articles

Articles may deal with pure lexicography or with the implications that related fields such as linguistics, general linguistics, computer science and management have for lexicography.

Contributions may be classified in any one of the following categories:

(1) **Articles:** Fundamentally original scientific research that has been done and the results that have been obtained, or reflecting existing research results and other facts in an original, synoptic, interpretative, comparative or critically evaluative manner.

(2) **Review articles:** Research articles presented in the form of a critical review of one or more published scientific sources.

Contributions in categories (1) and (2) are subjected to strict anonymous evaluation by independent academic peers in order to ensure the international research quality thereof.

(3) **Reviews:** An analysis and critical evaluation of published scientific sources and products, such as books and computer software.

(4) **Projects:** Discussions of lexicographical projects.

(5) **Lexiconotes:** Any article containing practice-oriented information, suggestions, problems, questions, commentary and solutions regarding lexicography.

(6) **Lexicovaria:** Any of a large variety of articles, announcements and press releases by lexicographic societies which are of particular value to the practising lexicographer.

(7) **Reports:** Reports on conferences and workshops.

Contributions in categories (3)-(7) must all meet the requirements of academic writing and are evaluated by the editors with this in mind.

2. Academic standard and evaluation procedure

The Department of Education of the South African Government has approved *Lexikos* as a subsidized, i.e. income-generating research journal.

Articles will be evaluated on the following aspects: language and style; conciseness and comprehensibility; problem formulation, reasoning and conclusion; references to the most important and most recent literature; substantial contribution to the specific discipline.

3. Language of contributions

Afrikaans, Dutch, English, French or German.

4. Copyright

Neither the Bureau of the WAT nor the African Association for Lexicography (AFRILEX) accepts any responsibility for claims which may arise from contributing authors' use of material from other sources.

Copyright of all material published in *Lexikos* will be vested in the Board of Directors of the Woordeboek van die Afrikaanse Taal. Authors are free however to use their material elsewhere provided that *Lexikos* (AFRILEX Series) is acknowledged as the original publication source.

5. Originality

Only original contributions will be considered for publication. Authors bear full responsibility for the originality and factual content of their contributions.

6. Free offprints and copies

Authors will receive five free offprints of each of their articles or review articles published, as well as one complimentary copy of the issue containing such article(s). Authors of purely evaluative

reviews and of contributions to the categories Lexiconotes, Lexicovaria, Projects, and Reports receive five free offprints of their contributions. In the case of the latter four categories, the editors may, however, depending on the nature and scope of the contributions, decide to grant the author a copy of the issue concerned.

7. Invitation and editorial address

All interested authors are invited to submit contributions for publication in *Lexikos* to:

The Editor: LEXIKOS
Bureau of the WAT
P.O. Box 245
7599 STELLENBOSCH
Republic of South Africa

B. PREPARATION OF MANUSCRIPTS

Manuscripts of articles must meet the following editorial requirements:

1. Length and format

Contributions should not exceed more than 20 typewritten A4 pages with double spacing and ample margins (about 2,5 cms). Manuscript should preferably be in electronic form on a (360 KB to 1.44 MB) floppy disk as either ASCII text, fully-formatted Microsoft Word (DOS or Windows) or WordPerfect (DOS or Windows) files. A computer printout of the article should accompany the disk. Each article must be accompanied by an English abstract of 150 to 250 words, and between 10 and 30 English keywords.

2. Graphics

One set of clear original drawings, tables, graphs, diagrams or quality prints thereof must be submitted. The locations of graphics must be clearly indicated in the text.

3. Bibliographical details and references in the text

Examine recent issues of *Lexikos* for details.

HINWEISE UND RICHTLINIEN FÜR AUTOREN

(Nehmen Sie bitte uns Kontakt auf (lexikos@sun.ac.za) für eine ausführlichere Wiedergabe dieser Hinweise oder besuchen Sie unsere Webseite: <http://www.wat.co.za>)

A. REDAKTIONELLE ZIELSETZUNGEN

1. Art und Inhalt der Artikel

Es können Artikel aufgenommen werden, die sich mit Themen der Lexikographie befassen oder mit Zusammenhängen, die zwischen der Lexikographie und benachbarten Fachgebieten wie z.B. Linguistik, allgemeiner Sprachwissenschaft, Lexikologie, Computerwissenschaft und Management bestehen.

Die Beiträge sollten einer der folgenden Kategorien entsprechen:

(1) **Artikel**, die grundlegend über neue Forschungsansätze und deren Ergebnisse berichten, oder die bestehende Forschungsergebnisse und andere Informationen selbständig, interpretativ, vergleichend oder kritisch bewertend wiedergeben.

(2) **Rezensionsartikel**, die in der Form eines Forschungsartikels eine oder mehrere veröffentlichten wissenschaftlichen Quellen kritisch rezensieren.

Beiträge in Kategorien (1) und (2) werden streng anonym von unabhängigen wissenschaftlichen Experten begutachtet, um ein internationales fachliches Niveau in *Lexikos* zu gewährleisten.

(3) **Rezensionen**, die veröffentlichte wissenschaftliche Quellen und Produkte, wie z.B. Bücher und Software, analysieren und kritisch bewerten.

(4) Lexikographische **Projekte**, die vorgestellt werden.

(5) **Notizen zum Lexikon**, die praxisbezogene Informationen, Vorschläge, Probleme, Fragen, Kommentare und Lösungen hinsichtlich der Lexikographie enthalten.

(6) **Lexikovaria**, die unterschiedliche Beiträge, Ankündigungen und Pressemitteilungen lexikographischer Vereinigungen, die dem praktischen Lexikographen wichtig sein können, einschließen.

(7) **Berichte** über Konferenzen und Workshops.

Beiträge in Kategorien (3)-(7) müssen im akademischen Stil abgefaßt werden. Sie werden von der Redaktion unter diesem Gesichtspunkt beurteilt.

2. Wissenschaftliche Standards und das Beurteilungsverfahren

Das Erziehungsministerium der südafrikanischen Regierung hat *Lexikos* als eine subventionierte, d.h. einkommenerzeugende Forschungszeitschrift anerkannt.

Artikel werden auf Grund der folgenden Gesichtspunkte bewertet: Sprache und Stil; Sachlichkeit und Verständlichkeit; Problembeschreibung, Argumentation und Schlußfolgerung; Hinweise auf die neueste und wichtigste Literatur; wesentlicher Beitrag zum besonderen Fachgebiet.

3. Sprache der Beiträge

Afrikaans, Deutsch, Englisch, Französisch oder Niederländisch.

4. Das Urheberrecht

Weder das Büro des WAT noch die African Association for Lexicography (AFRILEX) übernehmen Verantwortung für Ansprüche, die daraus entstehen könnten, daß Autoren Material aus anderen Quellen benutzt haben.

Das Urheberrecht aller in *Lexikos* publizierten Artikel wird dem Direktorium unseres Büros übertragen. Es steht Autoren jedoch frei, ihren Beitrag anderweitig zu verwenden, vorausgesetzt, *Lexikos* (AFRILEX-Serie) wird als Originalquelle genannt.

5. Originalität

Nur Originalbeiträge werden begutachtet. Autoren tragen die volle Verantwortung für die Originalität und den sachlichen Inhalt ihrer Beiträge.

6. Sonderdrucke und Freixemplare

Autoren erhalten fünf Sonderdrucke ihrer veröffentlichten Artikel oder Rezensionsartikel gratis

sowie ein Freiemplar der betreffenden Ausgabe. Rezensenten und Autoren von Beiträgen zu den Kategorien Notizen zum Lexikon, Lexikovaria, Projekte und Berichte erhalten fünf Sonderdrucke ihrer Beiträge gratis. Die Redaktion kann sich jedoch, abhängig von der Art und dem Umfang der Beiträge der letztgenannten vier Kategorien, vorbehalten, dem Autor ein Freiemplar der Ausgabe zu überlassen.

7. Einladung und redaktionelle Adresse

Alle Autoren, die interessiert sind, Beiträge für *Lexikos* zu liefern, sind herzlich willkommen. Sie werden gebeten, ihre Artikel an die folgende Adresse zu schicken:

Der Redakteur: LEXIKOS

Buro van die WAT

Postfach 245

7599 STELLENBOSCH

Republik Südafrika

B. VORBEREITUNG DES MANUSKRIPTS

Ein Artikelmanuskript muß den folgenden redaktionellen Anforderungen entsprechen:

1. Umfang und Format

Beiträge sollen nicht länger als 20 getippte A4-Seiten in zweizeiligem Abstand und mit Randabständen von ca. 2,5 cm sein. Das Manuskript sollte möglichst als elektronischer Text auf einer (360 KB bis 1.44 MB) Diskette vorgelegt werden, entweder im ASCII-Format, oder in formatiertem Microsoft Word (DOS oder Windows) bzw. WordPerfect (DOS oder Windows). Ein Ausdruck des vollständig formatierten Artikels soll mit der Diskette eingereicht werden. Jedem Artikel ist eine Zusammenfassung im Umfang von 150-250 Wörtern beizufügen. Ferner sollen etwa 10-30 inhaltskennzeichnende Stichwörter zu jedem Artikel angegeben werden.

2. Abbildungen

Ein reproduktionsfähiger Satz der originalen Abbildungen, Illustrationen, Tabellen, Graphiken und Diagramme oder Qualitätsabdrucke muß vorgelegt werden. Der Text selber sollte klare Hinweise auf die Position der Abbildungen enthalten.

3. Bibliographische Einzelheiten und Hinweise im Text

Zu Einzelheiten des bibliographischen Systems sind neuere Ausgaben von *Lexikos* einzusehen.

INSTRUCTIONS AUX AUTEURS

(Pour une version plus détaillée de ces instructions, contacter le Bureau du WAT (lexikos@sun.ac.za)
ou consulter notre website: <http://www.wat.co.za>)

A. POLITIQUE ÉDITORIALE

1. Caractéristiques et contenu des articles

Les articles seront consacrés à la lexicographie pure, ou aux rapports entre la lexicographie et les disciplines voisines telles que la linguistique, la linguistique générale, l'informatique et le management.

Les contributions pourront appartenir à l'une des catégories suivantes:

(1) **Articles:** Recherches scientifiques originales, avec leurs résultats; ou présentations originales, synoptiques, interprétatives, comparatives, évaluatives et critiques des résultats de recherches en cours;

(2) **Articles bilans:** Articles de recherche présentés sous forme de bilan critique de travaux scientifiques déjà publiés.

Les contributions appartenant aux catégories (1) et (2) seront soumises de manière anonyme à des experts spécialistes indépendants afin d'en assurer la qualité scientifique au niveau international.

(3) **Recensions:** Analyses et évaluations critiques de travaux de recherche et de productions scientifiques, telles que livres ou logiciels;

(4) **Projets:** Présentations de projets lexicographiques;

(5) **Lexiconotes:** Textes contenant des informations pratiques, ou des suggestions, des problèmes, des questions, des commentaires et des solutions concernant des activités lexicographiques;

(6) **Lexicovaria:** Articles, annonces, communiqués de presse émanant de centres de lexicographie et qui revêtent un intérêt particulier pour les lexicographes;

(7) **Rapports:** Rapports sur des colloques et ateliers.

Les contributions dans les catégories (3) à (7) devront répondre aux exigences de qualité des publications scientifiques et seront évaluées dans cette optique.

2. Critères et procédures d'évaluation

La revue *Lexikos* est reconnue et subventionnée par le Ministère de l'Éducation du gouvernement Sud-Africain comme revue devant générer des revenus.

Les articles seront évalués selon les critères suivants: langue et style, concision et clarté, formulation de la problématique, raisonnement et conclusion, référence aux travaux les plus importants et les plus récents, contribution substantielle à la discipline.

3. Langue des contributions

Afrikaans, allemand, anglais, français ou néerlandais.

4. Copyright

Le Bureau du WAT ou l'*African Association for Lexicography* (AFRILEX) décline toute responsabilité en cas de réclamations motivées par l'utilisation d'autres sources par les auteurs.

Les droits d'auteurs des documents publiés dans *Lexikos* appartiennent au Conseil d'administration du *Woordeboek van die Afrikaanse Taal* (WAT). Cependant, les auteurs sont libres d'utiliser leurs textes dans d'autres publications, à condition d'indiquer *Lexikos* (collection AFRILEX) comme source de la publication originale.

5. Originalité

Seules les contributions originales seront acceptées pour la publication. Les auteurs conservent l'entière responsabilité de l'originalité et du contenu factuel de leur texte.

6. Tirés-à-part et exemplaires gratuits

Les auteurs recevront gratuitement cinq (5) tirés-à-part de chaque article ou article de bilan, ainsi

qu'un exemplaire gratuit de la publication contenant cet article.

Les auteurs des recensions et des publications dans les catégories 'Lexiconotes', 'Lexicovaria', Projets et Rapports recevront cinq (5) tirés-à-part de leur texte. Dans ces quatre dernières catégories, le responsable de la revue pourra néanmoins décider, en fonction de la nature et des dimensions des publications, d'accorder un exemplaire gratuit de la revue à leurs auteurs.

7. Adresse de la revue

Les auteurs intéressés sont invités à soumettre leurs propositions à:

L'éditeur: LEXIKOS
Bureau du WAT
Boîte postale 245
7599 STELLENBOSCH
République d'Afrique du Sud

B. PRÉSENTATION DES MANUSCRITS

Les manuscrits se conformeront aux exigences suivantes:

1. Longueur et format

Les contributions ne devraient pas excéder 20 pages dactylographiées, de format A4, avec double espacement et marges suffisantes (environ 2,5 cm), si possible en format électronique (disquette 360KB à 1,44MB), sous forme de fichier ASCII, complètement formaté sous Microsoft Word ou sous WordPerfect (DOS ou Windows). La disquette sera accompagnée d'un tirage papier. Chaque article sera pourvu d'un résumé en anglais de 150 à 200 mots, et de 10 à 30 mots-clés.

2. Tableaux et graphiques

Les dessins, tableaux, graphiques et diagrammes seront envoyés, soit sous leur forme originale soit sous forme d'une copie de bonne qualité. Leur place dans le texte devra être clairement indiquée.

3. Bibliographie et références dans le texte

Voir les exemplaires récents de *Lexikos*.