Lexikos 14

Lexikos 14

Redakteur

Editor

J.C.M.D. du Plessis

Resensieredakteur

Review Editor

T. Harteveld

African Association for Lexicography

BURO VAN DIE WAT

STELLENBOSCH

Adviesraad / Advisory Board

Redaksiekomitee / Editorial Committee

Inhoud / Contents

Artikels / Articles

On Learners' Dictionaries

Projekte / Projects

Voorwoord

Vanjaar se uitgawe van *Lexikos* is die lywigste wat nog gepubliseer is sedert sy totstandkoming. Die rede hiervoor is die insluiting van 'n spesiale afdeling oor aanleerderswoordeboeke, bestaande uit die meeste voordragte aangebied op 'n tweedaagse seminaar wat op 11 en 12 November 2002 by die Universiteit van Pretoria gehou is. Ek wil graag die deelnemers bedank vir hul bereidwilligheid om hul lesings beskikbaar te stel vir insluiting in hierdie uitgawe van *Lexikos*.

Daar is besluit om af te sien van die klassifikasie van artikels in navorsingsartikels en beskouende artikels, soos die gebruik in *Lexikos* was sedert die vierde uitgawe. Hierdie besluit is geneem omdat baie artikels dikwels eienskappe van altwee soorte besit, wat dit soms moeilik maak om hulle onder 'n spesifieke afdeling in te deel. In vorige uitgawes sou sekere bydraes wat as beskouende artikels geklassifiseer is, ook onder navorsingsartikels ingedeel kon gewees het, en baie navorsingsartikels vertoon ook kenmerke van beskouende artikels.

Soos die geval verlede jaar was, het *Lexikos* vanjaar ook veel meer artikels ontvang as wat in hierdie uitgawe ingesluit kon word. Die bydraers moet bedank word vir hul begrip en bereidwilligheid dat hul artikels oorgedra word na volgende jaar se uitgawe.

Die keurders verdien 'n spesiale woord van dank vir hul belangrike bydrae om die artikels te beoordeel en op hierdie wyse die gehalte van *Lexikos* te verseker.

Op haar gewone entoesiastiese manier was me. Riette Ruthven verantwoordelik vir die professionele en vaardige setwerk van al die artikels. Sy het ook 'n groot deel van die korrespondensie betreffende die beoordeling en werwing van die artikels gehanteer. Me. Tanja Harteveld het tegniese hulp verleen. Hulle altwee het 'n aansienlike bydrae gelewer tot die suksesvolle afhandeling van hierdie uitgawe van *Lexikos*.

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

Foreword

This year's issue of *Lexikos* is the bulkiest that has been published since its establishment. The reason for this is the inclusion of a special section on learner's dictionaries, comprising most of the papers presented at a two-day seminar held on 11 and 12 November 2002 at the University of Pretoria. I wish to thank the participants for their willingness to make their papers available for inclusion in this issue of *Lexikos*.

It was decided to abandon the classification of articles into research articles and contemplative articles, as has been the practice in *Lexikos* since the fourth issue. This decision has been taken because many articles often possess qualities of both types which sometimes makes it difficult to classify them under a specific heading. In previous issues, some contributions classified as contemplative articles could also have been included under research articles, and many research articles also show features of contemplative articles.

As has been the case last year, *Lexikos* this year also received far more articles than could be incorporated in this issue. The contributors must be thanked for their understanding and willingness that their articles be carried over to next year's issue.

The referees deserve a special word of thanks for their important contribution to judge the articles and in this way to ensure the quality of *Lexikos*.

In her usual enthusiastic way, Ms Riette Ruthven was responsible for the professional and skilful typesetting of all the articles. She also handled a large part of the correspondence concerning the refereeing and requesting of articles. Ms Tanja Harteveld lent technical assistance. Both of them made a considerable contribution towards the successful completion of this issue of *Lexikos*.

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

'n Woord van AFRILEX

'n Taal voldoen aan die algemene definisie van standaardisering wanneer daar 'n grammatika, spel- en skryfreëls, woordeboeke en terminologie in daardie taal beskikbaar is. 'n Standaardtaal is gekodifiseer en eenvormig. Tale verander egter, ontwikkel en word gemoderniseer en daarom is standaardisering 'n voortdurende proses.

Aangesien taal dinamies is, is dit heeltemal natuurlik dat die woordeskat, spelling en skryfwyse van 'n taal sal verander. Hierdie veranderinge word weerspieël in die manier waarop veral die spelsisteem van 'n taal aanpas by die ontwikkeling van daardie taal. Gewoonlik bly die basiese spel- en skryfbeginsels onveranderd, maar taal ontwikkel en moderniseer en hierdie proses het gewoonlik 'n invloed op die spelling en skryfwyse.

Die spel- en skryfreëls van die amptelike Suid-Afrikaanse tale moet dringend hersien word. Die verskillende Nasionale Leksikografiese Eenhede (NLE's) van die Pan-Suid-Afrikaanse Taalraad (PanSAT), die Terminologiekoördineringsafdeling (TKA) van die Nasionale Taaldiens (NTD), Department van Kuns en Kultuur (DKK), en ander individue of organisasies wat besig is met die saamstel van algemene of vakwoordeboeke kan nie woordeboeke vir die verskillende amptelike tale maak sonder gestandaardiseerde spel- en skryfwyses nie. Vertalers kan nie vertaal en uitgewers kan nie publiseer sonder gestandaardiseerde spel- en skryfreëls nie. Daar is gevolglik 'n behoefte aan 'n betroubare dokument om te raadpleeg oor die skryf-, spel- en redigeerreëls vir elk van die amptelike tale.

Die Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns behartig die Afrikaanse spel- en skryfreëls. Hierdie Kommissie publiseer die *Afrikaanse Woordelys en Spelreëls* (AWS). Die AWS is alreeds verskeie kere hersien en elke hersiening weerspieël die veranderinge in die Afrikaanse taal oor die vorige dekades. Die verskillende Taalkommissies het probeer om by te bly by die dinamiese groei van Afrikaans oor die dekades deur die veranderinge in die spelling en skryfwyse in elke hersiene bewerking van die AWS te inkorporeer.

Dokumente wat die spel- en skryfreëls behandel, bestaan wel in al die Afrikatale, maar hierdie dokumente is verouderd en nie so omvattend as wat hulle veronderstel is om te wees ten einde in die hedendaagse behoeftes op die verskillende kommunikasiegebiede ten opsigte van kennisverspreiding te voorsien nie. Die spel- en skryfreëls moet dringend hersien en gemoderniseer word. Taal kan nie in geskrewe kommunikasie gebruik word of in woordeboeke gedokumenteer word as die spel- en skryfwyse nie gestandaardiseer is nie. Nuwe terme wat geskep word, benodig 'n stewige taalkundige basis met behoorlike woordvormingsbeginsels as norm.

PanSAT het alreeds met hierdie projek begin deur die Nasionale Taalliggame te versoek om die bestaande spel- en skryfreëls te hersien. Die PanSAT-strukture (bv. die Nasionale Taalliggame (NTL'e), die Provinsiale Taalkomitees (PTK's), en die Nasionale Leksikografiese Eenhede (NLE's)) wil graag met die nasionale en provinsiale Departemente van Kuns en Kultuur, van Onderwys en van Kommunikasie saamwerk. Die finale dokumente vir elk van die amptelike Afrikatale behoort produkte van konsultatiewe meganismes te wees waarvoor die verskillende sprekers van die amptelike tale regmatig eienaarskap kan aanvaar.

Die Nasionale Taalliggame is goed geposisioneer om 'n produk te lewer met normatiewe reëls oor spelling en skryfwyse vir elk van die amptelike Afrikatale, wat kommunikasie in Suid-Afrika sal verbeter. Die tegniese komitees (TK's) wat met die standaardisering toevertrou is, sal moet besluit oor die graad van hersiening wat vir die verskillende tale benodig word. Sommige van die Nasionale Taalliggame het alreeds met die hersiening van die spel- en skryfreëls vir hul onderskeie tale begin.

Die tegniese komitees (TK's) sal die bestaande dokumente oor spel-, skryf- en redigeerreëls nagaan en/of hersien. Hulle sal ook fokus op naamgewings- en woordvormingsbeginsels, en op afkortings- en punktuasiebeginsels. Hierdie proses behels standaardisering sowel as modernisering van die bestaande reëls.

Die uitset sal nege produkte wees wat maklike verwysing na die spel- en skryfreëls van die nege amptelike Afrikatale sal moontlik maak. Hulle sal gebruik word as vinnige verwysingsbronne wat ook omvattende woordelyste sal bevat wat die reëls illustreer.

Ek dink ek praat namens al die AFRILEX-lede wanneer ek sê dat ons almal daarna uitsien om hersiene spel- en skryfreëls vir al ons amptelike Suid-Afrikaanse tale te hê.

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

A Few Words from AFRILEX

A language complies with the common definition of standardisation if there are grammars, spelling and orthography rules, dictionaries and terminology available in that language. A standard language is codified and uniform. Languages, however, change, develop and are modernised and therefore standardisation is an ongoing process.

Since language is dynamic, it is quite natural for the vocabulary, spelling and orthography of a language to change. These changes are reflected in the way that especially the spelling system of a language adapts according to development in that language. Usually the basic spelling and orthography policies and principles remain unchanged but languages develop and modernise and this process usually has an effect on the spelling and orthography.

The spelling and orthography rules for the official South African languages need to be revised urgently. The various National Lexicography Units (NLUs) of the Pan South African Language Board (PanSALB), the Terminology Co-ordination Section (TCS) of the National Language Service (NLS), Department of Arts and Culture (DAC), and other individuals or organisations who are compiling general or technical dictionaries cannot make dictionaries for the various official languages without standardised spelling and orthography. Translators cannot translate and publishers cannot publish without standardised spelling and orthography rules. There is therefore a need for a reliable document to consult on orthography, spelling and editing rules for each of the official South African languages.

The Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns (the Language Commission of the South African Academy for Science and the Arts) deals with Afrikaans spelling and orthography rules. This Commission publishes the *Afrikaanse Woordelys en Spelreëls* (AWS). The AWS has been revised several times already and each revision reflects the changes in the Afrikaans language throughout the past decades. The various Language Commissions tried to keep abreast of the dynamic growth of Afrikaans throughout the decades by incorporating changes in the spelling and orthography in each revised version of the AWS.

Documents dealing with spelling and orthography rules do exist in all the African languages but these documents are outdated and are not as comprehensive as they should be in order to address today's needs in the various fields of communication with regard to knowledge dissemination. The spelling and orthography rules need to be revised and modernised urgently. Language cannot be used in written communication or documented in dictionaries if the spelling and orthography rules are not standardised. New terms that are created need a solid linguistic basis with proper word-forming principles as norm.

PanSALB has already embarked on this project by requesting the National Language Bodies to revise the current spelling and orthography rules. The PanSALB structures (e.g. the National Language Bodies (NLBs), the Provincial Language Committees (PLCs), and the National Lexicography Units (NLUs)) would like to collaborate with the national and provincial Departments of Arts and Culture, of Education, and of Communication. The final documents for each of the official African languages should be products of consultative mechanisms of which the different speakers of the official African languages could rightfully claim ownership.

The National Language Bodies are well positioned to deliver a product with normative rules on spelling and orthography for each of the official African languages, that would enhance communication in South Africa. The Technical Committees (TCs) entrusted with standardisation will have to decide on the degree of revision that is needed for the respective languages. Some of the National Language Bodies have already started with the revision of the spelling and orthography rules of their respective languages.

The Technical Committees (TCs) will revisit and/or revise existing documents on spelling, orthography and editing rules. They will also focus on naming and word forming principles, and on abbreviation and punctuation principles. This process entails standardisation as well as modernisation of the existing rules.

The output will be nine products that will make easy reference to the spelling and orthography rules of the nine official African languages possible. These will be used as quick reference tools that will also include appropriate word lists illustrating the rules.

I think I speak on behalf of all the AFRILEX members when I say that we are all looking forward to having revised spelling and orthography rules for all our official South African languages.

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

Redaksionele doelstellings

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

Die strewe van die AFRILEX-reeks is:

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

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

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

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

Editorial Objectives

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

The objectives of the AFRILEX Series are:

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

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

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

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

Redaktionelle Ziele

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

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

(1)   ein Medium schaffen für die nationale und internationale Diskussion, besonders aber der Lexikographie in Afrika mit seinen zahlreichen Sprachen dienen;
(2)   die Diskussion fördern, unter Lexikographen als auch zwischen 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 publiées sous forme de publications séparées dans la collection AFRILEX.

Die impak van kritiek op die samestelling van die *Woordeboek van die Afrikaanse Taal* (WAT)*

Willem Botha, *Woordeboek van die Afrikaanse Taal, Stellenbosch en Departement van Afrikaans en Nederlands, Universiteit van Stellenbosch, Stellenbosch, Republiek van Suid-Afrika (wfb@sun.ac.za)*

Opsomming: Die WAT se sensitiwiteit vir kritiek blyk duidelik uit die feit dat geeneen van die hoofpunte van kritiek ten opsigte van die eerste ag dele in die kritiek op dele IX en X herhaal word nie. Ensiklopediese behandeling van leksikale items, die ingewikkeldheid of onverstaanbaarheid van betekenisverklarings, die ondeurdagte aanwending van voorbeeldmateriaal, stadige alfabetiese vordering en die afwesigheid van volledige uitspraakleiding figureer derhalwe nie in die kritiek op dele IX en X nie. Die artikel gee 'n oorsig van die hoofpunte van kritek op dele IX en X en bespreek hoe die WAT in daaropvolgende dele op die kritiek gereageer het. In gevalle waar daar nie op geldige kritiek ag geslaan is nie, of waar die kritiek nie met praktiese voorstelle gepaardgegaan het nie, word aanpassings aan die hand gedoen. Sodoende word die wisselwerking tussen leksikografiese teorie en praktyk 'n dinamiese proses.

Sleutelwoorde: HOOFPUNTE VAN KRITIEK, OMGANGSTAAL, VOLGORDE VAN SITATE, UITSPRAAKLEIDING, ONBEREIKBARE VERWYSINGS, MISLEIDENDE SINONIEM-VERWYSINGS, WISSELWERKING TUSSEN LEKSIKOGRAFIESE TEORIE EN PRAKTYK

Abstract: The Impact of Criticism on the Compilation of the *Woordeboek van die Afrikaanse Taal* (WAT). The WAT's sensitivity to criticism is evident from the fact that none of the main points of criticism against the first eight volumes is repeated in the criticism of volumes IX and X. Encyclopedic treatment of lexical items, complexity or incomprehensibility of meaning descriptions, ill-judged use of citations, slow alphabetic progress and the absence of full pronunciation guidance therefore do not feature in the criticism on volumes IX and X. The article gives an overview of the main points of criticism against the ninth and tenth volumes and discusses the way in which the WAT reacted to the criticism in subsequent volumes. In instances where valid critique was disregarded, or where the critique was not accompanied by practical solutions, adaptations are suggested. In this way, the interaction between lexicographic theory and practice becomes a dynamic process.

Keywords: MAIN POINTS OF CRITICISM, COLLOQUIAL LANGUAGE, ORDER OF CITATIONS, PRONUNCIATION GUIDANCE, REFERENCES THAT CANNOT BE REACHED, MIS-

---

LEADING SYNONYM REFERENCES, INTERACTION BETWEEN LEXICOGRAPHIC THEORY AND PRACTICE.

1.   Inleiding

Die besware wat in die kritiek op die eerste ag dele van die *Woordeboek van die Afrikaanse Taal* uitgespreek is, het 'n beslissende impak op die samestelling van die nuwe-generasie-WAT-dele gehad. Dit blyk daaruit dat geeneen van die kernbesware ten opsigte van die eerste dele in die kritiek op dele IX en X herhaal word nie. Ensiklopedisiteit, die ingewikkeldheid of onverstaanbaarheid van betekenisverklarings, die ondeurdagte aanwending van voorbeeldmateriaal, stadige alfabetiese vordering en die afwesigheid van volledige uitspraakleiding figureer derhalwe nie in die kritiek op dele IX en X nie. Die heftigheid wat die kritiek op die eerste ag dele gekenmerk het, is afwesig in die kritiek wat op dele IX en X uitgespreek word. In die identifisering van die kernbesware is daar derhalwe gekonsentreer op gevalle waar dieselfde besware deur meer as een kritikus ten opsigte van 'n bepaalde deel uitgespreek is en punte van kritiek wat op albei dele betrekking het.

2.   Kernbesware uitgespreek in die kritiek op Dele IX en X en die impak daarvan op die samestelling van Dele X en XI

2.1   Gebrekkige opname van items uit die omgangstaal

Die beswaar is teen Deel IX ingebring dat die omgangstaal nie genoegsaam in die makrostruktuur weerspieël word nie. Die kritiek het veral betrekking op Engelse leenwoorde met 'n hoë gebruiksfrekwensie wat nie opgeneem is nie. Kritici het die tekortkoming daaraan gewyt dat die WAT se materiaalversameling nie verteenwoordigend genoeg is nie (vgl. Feinauer 1996: 236-238 en Gouws 1994). Volgens Feinauer sou die volgende leksikale items onder andere in Deel IX opgeneem kon word met die etiket (*geselstaal*): *lover, lift, looks, lagoon, lawn, lawnmower, lumberjacket* en *landlord*, asook meerwoordige leksikale items soos *love en charity* en *likes en dislikes*.

Die oorsaak van die onbevredigende opname van omgangswoorde, veral uit Engels, was inderdaad, soos die kritici beweer het, 'n onverteenwoordigende materiaalversameling ten opsigte van die omgangstaal. Die Buro het 'n daadwerklike poging aangewend om die probleem te bowe te kom deur verskeie inisiatiewe om die elektroniese gedeelte van die WAT-materiaalversameling te vergroot (Van Schalkwyk 2002) en ook deur redakteurs toe te laat om op eie inisiatief omgangstaalvorme op te neem. Die belangrikste ontwikkeling was egter toe redakteurs vanaf die samestelling van Deel XI via die internet elektroniese toegang tot die argiewe van Nasionale Pers verkry het. Dit het redakteurs in staat gestel om gebrekkige voorbeeldmateriaal aan te vul, en ook om

hul vermoedens oor die bestaan of die gebruiklikheid van 'n bepaalde omgangstaalvorm in Afrikaans te kontroleer met die oog op opname in die WAT. Die beskikbaarheid van die baie groot bykomende bron van elektroniese materiaal lê egter ook 'n verantwoordelikheid op die redakteurs en doen 'n beroep op hul intuïtiewe taalaanvoeling om te weet wanneer om die argiewe te raadpleeg.

Engelse leenvertalings en verafrikaanste spelvorme blyk wel goed in WAT XI verteenwoordig te wees, byvoorbeeld *naais* (naas *nice*), *newwermaaind, newwer-newwer* (op die *newwer-newwer*), *nip, nippel, oukei, ouky, opdollie, opdres, opdress, opfieks, opjack, opjazz, opjêk, oplyn, opsent* (verouderend), *optjek, optof, optop, olraait, ôraait, orraait, ouvers, overs*. In die vierde uitgawe van die HAT word leenwoorde en verafrikaanste vorme opgeneem wat in die spreektaal "van 'n groot groep sprekers van die standaardvariëteit van Afrikaans" voorkom (Odendal en Gouws 2000: v). Van die 24 lukraak aangehaalde voorbeelde uit die WAT XI kom vyf, nl. *nippel, oukei, opdollie, opsent* en *orraait* ook in die HAT voor. Die WAT en die HAT verteenwoordig tipologies uiteenlopende woordeboeke. As omvattende woordeboek moet die WAT probeer om die totale woordeskat van Afrikaans te weerspieël, terwyl dit van 'n handwoordeboek verwag word om veral die standaardtaalwoordeskat op te neem. Die HAT het egter op vernuwende wyse die taak aanvaar om ook die spreektaal van standaardtaalsprekers te weerspieël. Vanweë sy kleiner omvang kan hy dit nooit op die skaal van 'n omvattende woordeboek soos die WAT onderneem nie, wat die groot verskil in die getal opnames hierbo verklaar. Terselfdertyd wil dit tog uit die lukraak raakgeblaaide voorbeelde hierbo voorkom of die WAT waarskynlik besig is om sy groter opdrag, in vergelyking met 'n standaardwoordeboek, na te kom.

Vanaf Deel IX na Deel XI was daar 'n beduidende toename in die opname van Engelse leenvertalings en verafrikaanste spelvorms. 'n Elektroniese soektog is uitgevoer na lemmas waarby die frase "*n.d. Eng.*" in 'n etiket voorkom en uit die resultaat is daardie voorbeelde geïdentifiseer wat as leenvertalings of verafrikaanste spelwyses kwalifiseer. Met hierdie soektog word geen aanspraak op akkuraatheid gemaak nie, maar wil daar slegs op 'n beduidende tendens gewys word. In die onderskeie dele is ongeveer 25, 50 en 90 voorbeelde van leenvertalings en verafrikaanste spelvorms gevind. Dit verteenwoordig onderskeidelik 0,37%, 0,53% en 0,80% van die lemmas in Dele IX, X en XI. Die stygende tendens in die opname van Omgangsafrikaans in die WAT blyk ook uit 'n vergelyking met die lys van Omgangsafrikaans wat agter in die *Afrikaanse Woordelys en Spelreëls* van 2002 opgeneem is. In die AWS se lys van Omgangsafrikaans is 15 items onder L opgeneem waarvan net vier in die WAT is, terwyl daar 29 items deur die AWS onder O opgeneem is waarvan 22 in die WAT verskyn.

In die kritiek op die gebrekkige opname van Omgangsafrikaans is daar pertinent melding gemaak van die min frekwentgebruikte Engelse leenwoorde wat in Deel IX opgeneem is. Suiwer Engelse leenwoorde uit Omgangsafrikaans

kom egter steeds baie min in die WAT voor. In Deel XI is deur middel van 'n elektroniese soektog slegs vier opgespoor, nl. *nice, nonstop, off-beat* en *offer* (bet. I 4); in Deel IX slegs drie, nl. *lady, ladybird* en *lord,* en in Deel X vyf, nl. *madam, miss, mister, mod* en *movie* wat almal as (*geselstaal*) en (*Eng.*) geëtiketteer is.

Die opname van Engelse leenwoorde in 'n omvattende verklarende woordeboek soos die WAT wat oor verskeie dekades saamgestel word, is meer problematies as in die geval van 'n handwoordeboek wat ongeveer elke tien jaar hersien kan word. Dit is vir die leksikograaf moeilik om te voorspel of 'n Engelse leenwoord werklik in die spreektaal gevestig sal raak en of dit, soos in so baie gevalle gebeur, redelik gou in onbruik gaan raak. Die oordeelsfout van die opsteller van 'n omvattende woordeboek bly hom letterlik lewenslank by, terwyl die samesteller van 'n handwoordeboek sy fout waarskynlik binne tien jaar sal kan regstel. Die maker van 'n handwoordeboek het nie noodwendig 'n maklike taak in hierdie verband nie, maar slegs 'n minder moeilike een. Konsekwentheid ten opsigte van opname is steeds moeilik haalbaar soos blyk uit die feit dat *stunning* wel, maar *amazing, awesome* en *fabulous* nie in die vierde uitgawe van die *Verklarende Handwoordeboek van die Afrikaanse Taal* opgeneem is nie.

Die suinige opname van Engelse leenwoorde uit die omgangstaal moet aan die geldende opnamebeleid eerder as 'n ontoereikende materiaalversameling toegeskryf word. Indien 'n Engelse leenwoord 'n verafrikaanste teenhanger het wat meer gebruiklik of ewe gebruiklik as die leenwoord is (byvoorbeeld *newwermaaind*), neem die WAT tans slegs die verafrikaanste vorm op. Indien die Engelse leenwoord meer gebruiklik is en baie sterk in die omgangstaal gevestig is, word albei vorme opgeneem (byvoorbeeld *nice* en *naais*). In die bepaling van die frekwensie word hoofsaaklik op geskrewe bronne (in elektroniese of gedrukte vorm) en in 'n mindere mate op die redakteurs se intuïsie en parate taalkennis staatgemaak (Lombard 2002). Derhalwe is *newwermaaind* opgeneem, maar nie *never mind* nie, teenoor *nice* en *naais* wat albei opgeneem is. Engelse leenwoorde sonder 'n verafrikaanste teenhanger word slegs opgeneem in gevalle waar daar 'n baie hoë gebruiksfrekwensie is en daar geen twyfel oor die gevestigdheid in die omgangswoordeskat bestaan nie.

Daar is hierbo opgemerk dat die beskikbaarheid van die Naspers Argiewe 'n verantwoordelikheid op die redakteurs plaas en 'n beroep op hul intuïtiewe taalaanvoeling doen om te weet wanneer om die argiewe te raadpleeg. Die redaksie sal waarskynlik daarby baat as hulle voorafbeplanning doen vir die opname van Engelse leenwoorde, aangesien dit andersins onwaarskynlik is dat redakteurs bloot op intuïsie konsekwent die Naspers Argiewe sal raadpleeg om hul vermoedens omtrent die frekwensie van 'n leenwoord, wat nie in die WAT-korpus opgeneem is nie en wat hulle op eie inisiatief oorweeg vir opname, te kontroleer. Daar kan op 'n baie eenvoudige basis begin word deur 'n omgangstaallys op 'n sentrale plek te hou waarop redaksielede omgangstaalwoorde kan aanteken wat in gesprekke en elders gehoor word. Dit sou 'n eerste stap wees in die optekening van ten minste die standaardtaalspreker se om-

gangstaal. Die AWS se lys van Omgangsafrikaans kan ook betrek word. Voordat daar aan 'n bepaalde letter van die alfabet begin word, kan die aangetekende items bespreek word en aanpassings gemaak word. Daarná kan die redakteurs die items in die materiaalversameling en die Naspers Argiewe kontroleer soos hulle alfabeties vorder. Mettertyd kan daar dan uit hierdie prosedure 'n opnamebeleid kristalliseer waardeur Engelse leenwoorde in 'n ruimer mate op 'n verantwoordbare wyse opgeneem word.

2.2   Ontoereikendheid van die *Toeligting*

Hoewel groot lof uitgespreek is ten opsigte van die *Toeligting*, was daar ook heelwat voorstelle ter verbetering daarvan. In bykans alle gevalle is daar gehoor gegee aan die voorstelle van kritici.

Na aanleiding van kritiek dat die *Toeligting* van Deel IX geen inligting gee oor die stelsel waarvolgens verskillende betekenisonderskeidings georden word nie (vgl. Feinauer 1996: 267), is aanpassings in Deel XI gemaak. In die *Toeligting* van Deel XI (p. xx) word uiteengesit dat 'n empiriese metode van rangskikking gebruik word. Daarvolgens word die betekenisonderskeidings wat ten tyde van die samestelling van die woordeboek volgens die oordeel van die samestellers die gebruiklikste is, eerste aangebied, en die mins gebruiklike betekenisonderskeiding laaste.

Kritici het ook daarop gewys dat 'n lys van registeretikette verstrek word, maar dat geen aanduiding van elkeen se bestek gegee word sodat onderskei kan word tussen woorde wat gemerk word as skeltaal, vloek, plat en vulgêr nie (vgl. Gouws 1996). In Deel XI is gehoor gegee aan die versoek vir gebruiksleiding by registeretikette en word daar in afdeling 2.6.1 van die *Toeligting* uiteengesit in watter gevalle elke etiket toegeken word.

Aanbevelings van kritici (vgl. Gouws 1999) het ook tot aanvullings ten opsigte van die afdeling oor verwysings in die *Toeligting* van Deel XI gelei. Ten opsigte van homoniemnommers word onder 2.3 *Die lemma en sy vorm* genoem dat verwysings na homonieme in ouer dele volgens die nuwe stelsel aangebied word waarvolgens die homoniemnommer die lemmateken voorafgaan en nie daarop volg soos in die ouer dele nie. Daar word vanaf Deel XI onder die afdeling Verwysings ook uitvoering gegee aan 'n voorstel dat melding gemaak word van verwysings vanuit die sentrale woordelys na die *Toeligting* in die geval van rassistiese items wat die sentrale woordelys en die voortekste verbind.

Die *Toeligting* is ook aangevul ten einde helderheid te probeer verskaf oor bepaalde onduidelikhede by kritici. Kritici (vgl. Gouws 1994) het aangetoon dat die relatief min sitate teenoor die getal poëme by gewone woorde uit die spreek- en skryftaal vir hulle daarop dui dat die woorde nie so goed in die materiaalversameling verteenwoordig is nie. Vanaf Deel XI word in die *Toeligting* verklaar dat poëme redaksioneel saamgestelde voorbeeldmateriaal is wat

meestal op werklike sitate gebaseer is. Die sitate waarop die poëme gebaseer is, is gewoonlik om 'n verskeidenheid redes nie geskik om direk aangehaal te word nie, byvoorbeeld omdat 'n volledige bronverwysing ontbreek, of omdat die sitaat onnodig lank is en moeilik verkort kan word met behoud van die verstaanbaarheid daarvan. Dit kan ook 'n poging tot ruimtebesparing wees waar byvoorbeeld verskeie grammatikale vorme geïllustreer moet word. Hierdie opset is veral van toepassing op die hantering van gewone woorde uit die spreektaal. In die geval van nuwer omgangstaalvorme in WAT IX en WAT X kan die gebruik van poëme egter op 'n gebrekkige materiaalversameling dui aangesien daar wel leemtes in dié verband bestaan het wat grootliks opgelos is toe redakteurs met die samestelling van Deel XI elektroniese toegang tot die argiewe van Nasionale Pers verkry het.

2.3   Onduidelik geordende en onverteenwoordigende sitaataangawe

Resensente (vgl. Feinauer 1996: 256) wys daarop dat dit in Deel IX, soos in Deel VIII, steeds nie duidelik is volgens watter maatstaf sitate georden word nie. Sommige resensente staan 'n chronologiese ordening van sitate voor terwyl ander daarop wys dat die jaartalverspreiding van sitate soms nie verteenwoordigend van 'n item se werklike gebruik is nie (vgl. Feinauer 1996 en Gouws 1994).

Die WAT het hierop gereageer deur in die *Toeligting* van Deel XI (p. xx) te verduidelik dat voorbeeldmateriaal in dieselfde volgorde aangegee word as die betekeniselemente in definisies. Elke betekeniselement word egter nie noodwendig met voorbeeldmateriaal toegelig nie.

Indien die volgorde van sitate in die artikels van die negende en tiende dele nagegaan word, sal egter opgelet word dat die sitate dikwels die grammatiese inligting ten opsigte van die lemma illustreer in die volgorde waarin dit aangegee is. By selfstandige naamwoorde weerspieël die lemma, op enkele uitsonderings na, die enkelvoudsvorm en in die gleuf vir grammatiese gegewens word eers die meervoudsvorm en daarna die verkleinwoordsvorm verstrek. Dit kom duidelik na vore in die sitaatordening van die artikels van onder andere *lens* en *model*.

In 'n poging om die stelsel van sitaatordening in die *Toeligting* te verduidelik is dus nagelaat om te meld dat twéé faktore, nl. die volgorde van betekeniselemente in die definiens en die volgorde waarin die grammatiese gegewens aangegee word, medebepalend is vir die sitaatvolgorde. Die redaksionele stelsel laat heelwat vryheid in hierdie verband toe en redakteurs oordeel per geval watter volgorde die grootste illustratiewe waarde ten opsigte van die definiendum sal hê, en rangskik die sitate dan dienooreenkomstig. Die sitaatordening in die WAT dien derhalwe tot versterking van óf die wisselwerking tussen die sitate en die definiens, óf ter versterking van die wisselwerking tussen die sitate en die grammatiese gegewens. Vir 'n bespreking van die wisselwerking of

ondersteuning tussen die sitaatmateriaal en ander inligtingstipes kan Rade-
meyer (1993) en Lombard (1992) geraadpleeg word, alhoewel hulle nie uit-
sprake maak oor die ordening van sitate nie. Gouws (1991: 275) verwys na die
eis van kohesie binne 'n woordeboekartikel.

Ten einde duidelikheid oor die sitaatordening in die WAT te verskaf,
behoort die onvolledige uiteensetting in die *Toeligting* in die volgende deel reg-
gestel te word.

Soos reeds genoem, bepleit sommige resensente 'n chronologiese ordening
van sitate in die WAT. Hieroor kan opgemerk word dat 'n chronologiese orde-
ning van sitate 'n duidelik herkenbare ordening is en 'n oorsigtelike beeld van
die jaartalverspreiding gee, maar geen besondere bydrae lewer tot die sitate se
illustrasie van die betekeniselemente in die definiens of die grammatiese inlig-
ting nie. In 'n sinchroniese woordeboek het 'n chronologiese sitaatordening nie
dieselfde waarde as in 'n historiese woordeboek waar die eerste sitaat ook die
oudste opgetekende gebruiksgeval van 'n bepaalde leksikale item veronderstel
nie.

Die kritiek dat die jaartalverspreiding van sitate soms nie verteenwoordi-
gend is nie, dui waarskynlik op 'n chronologies onverteenwoordigende mate-
riaalversameling ten opsigte van bepaalde leksikale items. Dit is opvallend dat
daar by sommige lemmas slegs sitate uit die sestiger- en sewentigerjare aange-
gee word. Die toegang wat redakteurs vanaf die samestelling van Deel XI tot al
hoe meer elektroniese materiaal, veral die argiewe van Nasionale Pers het,
behoort hierdie reeds te lank voortdurende probleem grootliks op te los.

2.4   Foutiewe en onvoldoende uitspraakaanduiding

In Deel IX kom letteruitspraak dikwels voor (Feinauer 1996: 353 en Gouws
1994), veral by samestellings of samestellende afleidings waar die geskrewe
verdubbeling van konsonante in die transkripsie herhaal word. Die redaksie
het hierop gereageer deur geminasie vanaf Deel X soos volg aan te dui:
[la(x)'xas].

In gevalle van glyerinvoeging word slegs 'n [h] as glyer ingevoeg, terwyl
die [h] soms ook met 'n [j] of met 'n [w] vervang kan word (Feinauer 1996: 254).
Die redaksie het besluit om hom in beginsel by glyerinvoeging tot die [h] te
beperk, veral vanweë die uitdying van permutasies wat by die uitspraakaan-
duiding ontstaan. Dié besluit word in die lig van die kritiek in die *Toeligting*
van Deel XII verantwoord.

Resensente spreek verder die hoop uit dat volgende dele van die WAT
ook sillabeverdeling sal bevat (Feinauer 1996: 255 en Gouws 1994). Daar is oor
dekades in feitlik elke deel van die WAT meer uitspraakinligting as in die
vorige deel gegee. Die redaksie moet ernstig oorweging daaraan skenk om ook
dié een vorm van uitspraakinligting wat nog uitstaande is, nl. sillabeverdeling,
in die volgende deel van die WAT aan te bied.

3.   Geldige kritiek waarop ag geslaan behoort te word

3.1   Inleiding

Desnieteenstaande die aansienlike impak van kritiek op die samestelling van die tiende en elfde dele van die WAT is daar enkele voorstelle van verdienstelike praktiese waarde gemaak waarop nie gereageer is nie en daar sal voorts op die uitvoerbaarheid van hierdie voorstelle gefokus word. Daar sal ook aanpassings voorgestel word waar die kritiek nie met praktiese voorstelle gepaardgegaan het nie.

3.2   Verwysings wat nie opgevolg kan word nie

Optimale gebruikersvriendelikheid was een van die hoofdoelwitte met die ontwerp van die "nuwe" WAT. Die gebruikersvriendelikheid van die WAT word egter benadeel deur bepaalde gevalle van verwysing wat nie deur die gebruiker opgevolg kan word nie. In sommige gevalle is dit te wyte aan 'n inherente verskynsel kenmerkend van 'n omvattende woordeboek wat deelsgewys verskyn en in ander gevalle kan dit toegeskryf word aan die implementering van die nuwe redaksionele stelsel vanaf Deel IX.

3.2.1   Verwysings na dele wat nog moet verskyn

3.2.1.1   Verwysings na verbandhoudende inskrywings

Verwysingsinskrywings ingelei deur die verwysingsmerker *Vgl.* vestig die aandag op verbandhoudende inskrywings in 'n ander artikel waar addisionele taalkundige inligting, byvoorbeeld sterk semantiese ooreenstemming, gevind kan word. Wanneer die gebruiker na leksikale items verwys word wat in dele hanteer sal word wat nog moet verskyn, kan die verwysing egter nie opgevolg word nie. Gouws (1999: 22-23) verwys byvoorbeeld na die onopvolgbare verwysingsinskrywing wat by *mutimoord* aangetref word, nl.: Vgl. RITUELE MOORD (RITUEEL, uitdr.). Hy stel voor dat dit oorweeg kan word om in sulke gevalle met 'n strukturele merker aan te dui dat 'n bepaalde verwysingsadres nie bereik kan word nie.

Hierdie voorstel sou met die minimum ontwrigting toegepas kon word op onopvolgbare verwysings na verbandhoudende inskrywings. Die enigste nadeel verbonde aan die hanteringswyse is dat wanneer die betrokke deel wat die verbandhoudende inskrywing bevat, wel verskyn het, die verwysing met die ingeboude waarskuwing betekenisloos word.

3.2.1.2   Sinoniemverwysings

Vir die opstellers van 'n meerdelige woordeboek bied sinoniemdefinisies aan-

sienlik meer probleme as vir die opstellers van 'n eendelige woordeboek. Dit geld veral gevalle waar kruisverwysing na 'n lemma plaasvind wat nog in 'n volgende deel verklaar moet word. In Deel X van die WAT word *meervrouig* byvoorbeeld na *poligenies* verwys. Gebruikers moet dus tot die verskyning van Deel XII wag om die verwysing na *poligenies,* waar die definisie verskyn, te kan opvolg.

Gouws maak twee voorstelle vir die hantering van hierdie tipiese probleem van meerdelige omvattende verklarende woordeboeke. Reeds in sy resensie van Deel VII stel hy (Gouws 1985: 18) voor dat die WAT-redaksie dit miskien kan oorweeg om die definisie in sulke gevalle maar by die sinoniem wat alfabeties eerste kom te plaas om die gebruiker te help. Gouws (1989: 149) haal Burnett (1979: 23) in dié verband aan: "the scholarly community in general prefers published imperfection to unpublished perfection" en maak ook 'n verdere praktiese voorstel. Hy doen naamlik aan die hand dat waar die minder gebruiklike lemma die verklaring moet kry omdat die meer gebruiklike sinoniem in 'n volgende deel val, die meer gebruiklike vermeld moet word met 'n gepaardgaande aanduiding dat dit 'n hoër frekwensie het as die onderhawige lemma. Feinauer (1993: 296) noem ook dat sinoniemverwysing onvermydelik probleme oplewer in 'n meerdelige woordeboek soos die WAT, maar stel nie alternatiewe hanteringswyses voor nie.

In sinchroniese woordeboeke soos die WAT speel gebruiksfrekwensie 'n beslissende rol in alle besluite oor die redaksionele hanteringswyse van 'n leksikale item. In die WAT geld dit veral die lemma waaronder die definisie geplaas word en die ordening van betekenisonderskeidings. Voorts streef die WAT met sy nuwe redaksionele stelsel na eenvoud, 'n eenvormige werkswyse en 'n helder oordrag van taalkundige inligting. Ten einde aan hierdie basiese beginsels vir die samestelling van die WAT getrou te bly, lyk dit na die aanvaarbaarste opsie om nie een van Gouws se twee bogenoemde voorstelle te implementeer nie, maar wel sy voorstel ten opsigte van onopvolgbare verbandhoudende verwysings ook hier te oorweeg. Hiervolgens word met 'n strukturele merker aangedui dat 'n bepaalde verwysingsadres nie bereik kan word nie, in plaas daarvan dat met 'n bykomende merker aangedui word dat 'n vermelde sinoniem 'n hoër gebruiksfrekwensie as die onderhawige lemma het.

Die verwysing sou steeds nie opgevolg kan word nie, maar die gebruiker sou ten minste weet wat die ratio agter sodanige onopvolgbare verwysing is. Dit sou ook beteken dat daar op eenvormige wyse in die woordeboek aangedui word dat 'n verwysingsadres nie bereik kan word nie. Te veel strukturele merkers verwar gebruikers en navrae aan die Buro het getoon dat selfs gesofistikeerde gebruikers nie duidelikheid het oor gevestigde strukturele merkers in die WAT nie, soos byvoorbeeld die skeidingsteken "≈" wat by die voorbeeldmateriaal die kollokasies en poëme van die sitate skei.

Niks kan egter gedoen word om gebruikers te waarsku teen verwysings uit vroeër dele na niebestaande adresse in die nuwe-generasie-dele nie.

In die *Toeligting* kan onder Verwysings uiteengesit word dat enige verwy-

singsadres met 'n strukturele merker, byvoorbeeld "*", beteken dat die adres nie bereik kan word nie. Al die sodanig gemerkte adresserings na dele wat nog moet verskyn, is egter slegs geldig tot die verskyning van die betrokke deel waarna verwys word. Op grond van hierdie swakheid, lyk die uiteengesette hanteringswyse, ten spyte van noue aansluiting by die beginsels van eenvoud, eenvormigheid en voorkeur aan gebruiksfrekwensie, nie ideaal nie.

3.2.1.3   Wisselvormverwysings

Dit word gewoonlik misgekyk dat die probleem van onopvolgbare verwysings in meerdelige omvattende woordeboeke homself ook manifesteer by wissel-vormverwysings. In Deel V van die WAT word *kambuis* verwys na *kombuis* wat eers in Deel VII gelemmatiseer is. Soortgelyke voorbeelde is *kambers* en *kombers*, en *fonkel* (verouderd) en *vonkel*. Die verskynsel sal veral ter sprake kom by let-ters soos S en V wat, soos K, meer as een deel sal beslaan. Soos by ander ver-wysings na dele wat nog moet verskyn, sou by wisselvormverwysings ook per struktuurmerker aangedui kon word dat die adres nie bereikbaar is nie. Die-selfde nadele as wat hierbo uitgestippel is, is egter ook hier van toepassing.

3.2.2   Verwysings na reeds verskene dele

Vanaf Deel IX word verwysings uit vorige dele wat nie meer die taalwerklik-heid weerspieël nie, nie gehandhaaf nie. Verwysings vanaf nuwe-generasie-dele na die ouer dele weerspieël ook die taalwerklikheid van die nuwer dele, selfs as dit meebring dat die gebruiker na 'n niebestaande adres verwys word. Dié verbreking in kontinuïteit is veral sigbaar by sinoniem- en wisselvormver-wysings en by Vgl.-gevalle, maar ook by sinoniem-, antoniem- en wisselvorm-vermelding.

Soms word die gebruiker na 'n niebestaande adres in 'n vorige deel ver-wys. Dit kom voor wanneer die lemma of betekenisonderskeiding waarna ver-wys word, nog nie bestaan het toe die vorige deel verskyn het nie, of bloot nie opgeneem is in die betrokke deel nie. So word daar in Deel IX by *loslopie* die volgende verwysing aangetref: Vgl. BYLOPIE. Deel I bevat egter nie die lemma *bylopie* nie en die gebruiker kan nie die adres bereik nie. In Deel X word die gebruiker onder andere by een van *muur* se betekenisonderskeidings en by *magnetofoon* deur middel van sinoniemverwysings na onderskeidelik *Berlynse muur* en *bandopnemer* verwys, wat nie een in Deel I opgeneem is nie, aangesien dié deel voor die oprigting van die muur én die uitvinding van die bandopne-mer verskyn het. By *literatuurlys* in Deel IX word *bronnelys* as sinoniem ver-meld, maar *bronnelys* is nie in Deel I opgeneem as lemma nie en by *luitenant* (bet. 3) word *adjudant* as sinoniem vermeld, maar in Deel I is daar nie sodanige betekenisonderskeiding by *adjudant* opgeneem nie.

In al bogenoemde gevalle sou per struktuurmerker aangedui kon word dat die verwysingsadres in 'n vorige deel nie bereik kan word nie, byvoor-

beeld: Vgl. BYLOPIE*; Bandopnemer*; Berlynse muur (BERLYNS, uitdr.)* en sin. *bronnelys*.

Gemerkte verwysings na reeds verskene dele het nie die nadeel van soortgelyke verwysings na toekomstige dele nie, wat slegs geldig is tot met die verskyning van die betrokke deel van die WAT.

In sommige gevalle kan die adres wel bereik word, maar vind 'n onbevredigende oordrag van inligting plaas. In Deel III word die gebruiker by *geboortemerk* na die sinoniem *moedervlek* verwys. Die gebruiker verwag dus om 'n definisie by *moedervlek* te kry met die vermelding van *geboortemerk* as sinoniem. By die samestelling van Deel X was die gebruiksfrekwensie van *geboortemerk* egter aansienlik hoër as dié van *moedervlek*. Op grond hiervan is *moedervlek* na *geboortemerk* verwys, sodat daar by geeneen van die lemmas 'n definisie verskyn nie, en die gebruiker in 'n sirkelgang gelei word. Dit is te betwyfel of daar in hierdie gevalle per struktuurmerker aangedui moet word dat die verwysing nie opgevolg kan word nie, aangesien die gebruiker na 'n lemma met 'n volledige bewerking, maar sonder 'n betekenisverklaring, verwys word. Die betrokke kategorie van onsuksesvolle opvolging van verwysings maak 'n eenvormige hantering problematies en noodsaak 'n meer pragmatiese oplossing.

3.2.3 Voorgestelde hanteringswyse van verwysingsadresse wat nie bereik kan word nie

3.2.3.1 Verwysings na dele wat nog moet verskyn

Enige verwysing na 'n WAT-deel wat nog moet verskyn waarby aangedui word dat dit nie opgevolg kan word nie, het 'n beperkte tydperk van geldigheid. Dit geld Vgl.-gevalle, sinoniem- en wisselvormverwysings en sinoniem-, antoniem- en wisselvormvermeldings. Ná die verskyning van die nuwe deel skep die verwysing verkeerdelik die indruk by die gebruiker dat die verwysingsadres nie bereik kan word nie en verhoed derhalwe die suksesvolle oordrag van inligting.

Vir die gebruiker kom die grootste verlies aan inligting deur die tydelike onbereikbaarheid van verwysings voor by sinoniem- en wisselvormverwysing. Die gebruiker moet naamlik tydelik sonder 'n betekenisverklaring klaarkom. Indien 'n betekenisverklaring by 'n ongewone of verouderde sinoniem of wisselvorm, of by 'n streektaalvorm in plaas van 'n standaardtaalvorm verskaf word om vir die tydelike verlies te kompenseer, word die ordeningsfunksie van gebruiksfrekwensie by die samestelling van 'n sinchroniese woordeboek verydel. In die WAT bepaal gebruiksfrekwensie onder andere of 'n leksikale item opgeneem word, die lemma waar betekenisverklaring verskaf word, die ordening van betekenisonderskeidings, die ordening van 'n sinoniemparadigma en die ordening van homonieme in die makrostruktuur. Die gebruikers is gewoond daaraan om hulle in hulle soektogte op gebruiksfrekwensie as ordeningsbeginsel te verlaat.

Dit lyk derhalwe na gesonde leksikografiese praktyk om ten opsigte van verwysings na toekomstige WAT-dele die status quo te handhaaf en nie te probeer kompenseer deur byvoorbeeld per struktuurmerker aan te dui dat 'n verwysingsadres onbereikbaar is of deur 'n betekenisverklaring by 'n minder gebruiklike sinoniem of wisselvorm te verskaf nie.

By minder gebruiklike of verouderde sinonieme en wisselvorme wat na 'n meer gebruiklike lemma in 'n toekomstige deel verwys word, moet die beperkings op hul gebruik konsekwent per etiket aangedui word, byvoorbeeld (*minder gebruiklik*) of (*verouderd*). Sodoende weet die gebruiker dat gebruiksfrekwensie as ordeningsbeginsel toegepas word. Indien sinonieme of wisselvorme min of meer ewe gebruiklik is, behoort voorkeur gegee te word aan daardie item wat alfabeties eerste kom.

Dit moet ook in gedagte gehou word dat die taalwerklikheid dinamies is en dat adresserings tussen nuwe-generasie-dele in die toekoms ook onbereikbare status mag verkry. Hieroor het leksikograwe egter min beheer aangesien hulle slegs beperkte insigte in die toekomstige veranderinge in die taal het.

3.2.3.2   Verwysings na reeds verskene dele

Ter wille van konsekwentheid is WAT-redakteurs oor dekades gedwing om 'n skewe voorstelling van die taalwerklikheid te perpetueer. Vanaf Deel IX is hulle nie meer gebonde aan verwysings uit vorige dele wat nie 'n getroue weergawe van die taalwerklikheid is nie. Verwysings na vorige dele weerspieël die taalwerklikheid tydens die samestelling van die nuwer deel, ongeag daarvan of dit in botsing verkeer met inligting in die vorige dele.

In Deel VII is *kontroversieel* van 'n definisie voorsien met *omstrede* as sinoniemvermelding. Die redaksie van Deel XI het egter die taalwerklikheid anders weerspieël deur by *omstrede* te herdefinieer en *kontroversieel* as sinoniem te vermeld.

In die strewe om die taalwerklikheid tydens samestelling weer te gee, word gebruikers soms na lemmas verwys wat nooit in 'n vorige deel opgeneem is nie, of waarby nie 'n betekenisverklaring verskaf is nie. Dit geld Vgl.-gevalle, sinoniem- en wisselvormverwysing en sinoniem-, antoniem- en wisselvormvermelding. So word *littérature engagée* na *betrokke letterkunde* (BETROKKE, uitdr.) verwys alhoewel *betrokke letterkunde* nie in Deel I opgeneem is nie. In Deel I is *geboortemerk* na *moedervlek* verwys, maar by die samestelling van Deel X was die gebruiksfrekwensie van *geboortemerk* aansienlik hoër as dié van *moedervlek*. Op grond hiervan is *moedervlek* na *geboortemerk* verwys, sodat daar by geeneen van die lemmas 'n definisie verskyn nie, en die gebruiker in 'n sirkelgang gelei word.

Dit lyk nie of dit op die kort of lang termyn in die belang van gebruikers is om doelbewus onbereikbare verwysings te skep ten einde die taalwerklikheid te reflekteer nie. 'n Minder meganiese, maar meer pragmatiese beleid ten opsigte van verwysings na reeds verskene dele kan gevolg word.

Dit kan in beginsel aanvaar word dat gebruikers nooit van die een sinoniem of wisselvorm na die ander verwys behoort te word, terwyl daar nie by een van die lemmas 'n betekenisverklaring verskaf word nie. Dit beteken noodwendig dat daar 'n betekenisverklaring by 'n minder gebruiklike of verouderde sinoniem of wisselvorm verskaf sal word. By die ongewone item *plakprent* moet die gebruiker byvoorbeeld na regte na die meer gebruiklike *collage* verwys word. *Collage* is egter nie in Deel I opgeneem nie en *plakprent* kan derhalwe soos volg hanteer word:

> **plakprent** … (*ongewoon*) Collage — prent wat saamgestel word deur stukkies van verskillende materiale, veral papier, op 'n oppervlak te plak.

Hierdie voorstel is reeds deur die redaksie aanvaar en sal in Deel XII toegepas word. By die uiteindelike herbewerking van die WAT sal die definisie dan in al hierdie gevalle telkens by die sinoniem met die hoogste gebruiksfrekwensie verstrek word.

Onbereikbare verwysingsinskrywings met *Vgl.* na reeds verskene dele kan gestaak word totdat daar met die uiteindelike herbewerking begin word na voltooiing van die WAT tot by Z.

3.3   Misleidende sinoniemverwysings

Sinoniemverwysinginsinskrywings bevat dikwels geen verwysingsmerkers en geen tipografiese of ander strukturele merkers nie, soos blyk uit die artikel van die lemma *melkriem*:

> **melkriem** … Koeiriem.

Gouws (1999: 10) wys daarop dat die sinoniemverwysing *koeiriem* deur die gebruiker wat nie bekend is met die inhoud van die *Toeligting* nie, as 'n beknopte definiens beskou kan word in plaas van 'n verwysingsadres.

Wanneer 'n polisemiese lemma betrokke is, word daar voorsiening gemaak vir betekenispresisering:

> **onsegbaar** … Onuitspreeklik (ONUITSPREEKLIK 1).

Gouws (1999: 25) stel voor dat bg. gevalle van sinoniemverwysingsinskrywings vereenvoudig word tot:

> **onsegbaar** … Onuitspreeklik 1.

Gouws se voorstel het wel 'n vereenvoudigde sinoniemverwysingsinskrywing tot gevolg, maar los nie die probleem van gevalle sonder 'n betekenispresise-

ring (soos *Koeiriem*) op, wat vir baie gebruikers nie as 'n verwysingsadres her-
kenbaar mag wees nie.

Dit kan oorweeg word om alle sinoniemverwysings in klein hoofletters
weer te gee, byvoorbeeld:

> **melkriem** … KOEIRIEM.

> **onsegbaar** … ONUITSPREEKLIK 1.

Indien sinoniemverwysings in klein hoofletters aangegee word, skakel dit die
moontlikheid uit dat gebruikers dit nie as 'n verwysing sal herken nie. Dit sal
ook meebring dat verpligte verwysings, nl. sinoniem- en wisselvormverwys-
ings, naastenby eenders hanteer word in die WAT.

3.4   Verwysingsverhouding tussen los- en vasgeskrewe idiomatiese uit-drukkings

Gouws (1999: 18) merk op dat waar leksikale items sowel los- as vasgeskryf
kan word, die losgeskrewe variasie slegs vermeld en nie gelemmatiseer word
nie, byvoorbeeld:

> **mooipraatjies …** Ook soms *mooi praatjies*.

Aangesien die betrokke inskrywing geen verdere woordeboekraadpleging van
die gebruiker kan inisieer nie, kan dit nie as deel van die mediostruktuur van
die woordeboek beskou word nie. Hy beveel derhalwe aan dat dit nie in die
Toeligting onder die verwysingstipes bespreek word nie.

By idiomatiese uitdrukkings soos *mooi praatjies* met 'n vaste teenhanger
mooipraatjies word 'n volledige bewerking by elke variasie aangebied net soos
in die geval van variasies van die kernkomponent of soekwoord van 'n idio-
matiese uitdrukking, soos in die *Toeligting* onder **5 Verwysings** uiteengesit
word.

Alhoewel die bogenoemde verwysing na *mooi praatjies* dus wel opgevolg
kan word, is die verwarring rondom verwysing by vas- en losgeskrewe leksi-
kale items wat hier na vore tree, waarskynlik die gevolg van 'n gebruikerson-
vriendelike verwysingstelsel. 'n Moontlike oplossing sou wees om gevalle soos
mooipraatjies soos volg te hanteer:

> **mooipraatjies** s.nw. Ook soms *mooi praatjies* (MOOI, uitdr.) Niksseg-
> gende, waardelose praatjies of verskonings, wat …

Die wysiging moet in die *Toeligting* onder 5.3 *Verwysingsmetode by los- en vasge-
skrewe vorme* aangebring word en die hele uiteensetting oor los- en vasgeskrewe

vorme verder behou word sodat gebruikers 'n oorsigtelike beeld van los- en vasgeskrewe items kan verkry.

3.5   Spellingverskil tussen lemma-aangawe en voorbeeldmateriaal

Gouws (1996) wys daarop dat dit dikwels gebeur dat die optrede van die lemma in die voorbeeldmateriaal anders gespel word as die lemma-aangawe self. Die lemma-aangawe *minibustaxi* word sonder koppelteken geskryf, maar albei optredes van die lemma in die sitaat kry 'n koppelteken tussen die twee onmiddellike stukke.

Dit is opvallend dat in baie van die samestellings met *mini-*, byvoorbeeld *minidrag*, *miniduikboot* en *minigholf* die voorbeeldmateriaal ook, of slegs, 'n spelwyse met 'n koppelteken weergee, nl. *mini-drag*, *mini-duikboot* en *mini-gholf*. Die betrokke voorbeeldmateriaal is alles uit die sestiger- en sewentigerjare afkomstig en dui op 'n tendens in daardie tydperk om samestellings met *mini-* met 'n koppelteken te skryf. Dit sou verwarring uitskakel indien die taalwerklikheid in sodanige gevalle weerspieël word deur die betrokke voorbeeldmateriaal na 'n aanloop weer te gee, byvoorbeeld: ■ Ook, verouderd, *mini-drag*. Ten opsigte van *minibustaxi* blyk dit egter dat daar 'n gebrek aan goeie voorbeeldmateriaal was en daar toe van 'n sitaat met 'n afwykende spelwyse gebruik gemaak is. In sulke gevalle sou die WAT waarskynlik beter vaar deur die spelling stilswyend te wysig.

Indien die spelwyse met 'n koppelteken by 'n bepaalde vorm 'n bereëlde alternatief is, kan dit in die lemmaveld vermeld word soos in die geval van items wat met 'n hoof- of 'n kleinletter gespel word: (*ook met 'n koppelteken* ekonomieseklas-passasier). Frekwensie behoort te bepaal watter vorm gelemmatiseer word. Indien *ekonomieseklas-passasier* die frekwentste voorkom en gelemmatiseer word, word die vaste spelwyse in die lemmaveld vermeld: (*ook vasgeskryf* ekonomieseklaspassasier).

3.6   Los- en vasskryf van woorde

Gouws (1996) wys daarop dat volgens die 1991-uitgawe van die AWS adjektiwiese woordgroepe van die tipe *kop teen kop botsing* sonder koppeltekens geskryf moet word. Deel X gee egter voorbeelde soos *man-tot-man-opleiding.* Die AWS bepaal ook dat verbindings wat bestaan uit voornaamwoorde en die versterkende element *self* los of vas geskryf kan word. Die WAT gee egter by *myself* slegs een vorm, nl. dié wat vas geskryf word.

Dit wil voorkom asof die WAT in gevalle soos die bogenoemde ook streng by sy deskriptiewe aard hou deur slegs te weerspieël wat in die voorbeeldmateriaal voorkom. Hierdie beginsel kan verwarring by die gebruiker skep, veral wanneer dit die WAT in botsing met 'n spel*reël* bring, en nie slegs met die spelling van 'n bepaalde item in die woordelys nie. Daar word egter in Deel XI

by *onsself* aangetoon dat dit ook los geskryf kan word, nl. *ons self*. Die WAT moet voortgaan om in soortgelyke gevalle 'n groter mate van buigsaamheid aan die dag te lê.

4.   Slot

Sommige van die voorstelle hierbo ten opsigte van die kritici se besware, is reeds deur die redaksie aanvaar is en sal in WAT XII geïmplementeer word. Ander voorstelle in geldige kritiek van resensente sal hopelik ook tot veranderinge in die WAT lei sodat die wisselwerking tussen teorie en praktyk dinamies kan bly en die woordeboek steeds die stand van die metaleksikografie sal weerspieël.

Geraadpleegde bronne

Burnett, L.S. 1979. Lexicographical Problems in the Treatment of some Linguistic Terms in a *Supplement* to the *OED*. *ITL Review of Applied Linguistics* 45/46: 19-24.

Feinauer, I. 1993. Die agste deel van die *Woordeboek van die Afrikaanse Taal*. *Tydskrif vir Geesteswetenskappe* 33(4): 286-298.

Feinauer, I. 1996. Die negende deel van die *Woordeboek van die Afrikaanse Taal*. *Lexikos* 6: 233-271.

Gouws, R.H. 1985. Die sewende deel van die *Woordeboek van die Afrikaanse Taal*. *Standpunte* 38(4): 13-25, Augustus.

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

Gouws, R.H. 1991. Resensie van Hausmann, F.J., O. Reichmann, H.E. Wiegand en L. Zgusta (Reds.): *Wörterbücher: Ein internationales Handbuch zur Lexikographie/Dictionaries: An International Encyclopedia of Lexicography/Dictionnaires: Encyclopédie internationale de lexicographie*. Berlyn: W. de Gruyter, 1989–1991. *Lexicographica* 7: 268-290.

Gouws, R.H. 1994. *Woordeboek van die Afrikaanse Taal, Deel IX*. Bespreking gelewer op die SAUK-radioprogram *Skrywers en Boeke*.

Gouws, R.H. 1996. *Woordeboek van die Afrikaanse Taal, Deel X*. Bespreking gelewer op die SAUK-radioprogram *Skrywers en Boeke*.

Gouws, R.H. 1999. Mediostructural Representation, Textual Condensation and User-orientation in the *WAT X*. *Lexicographica* 15: 4-37.

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

Lombard, F.J. 2002. Onderhoud, Stellenbosch, 12 Junie.

Odendal, F.F. en R.H. Gouws (Reds.). 2000. *Verklarende Handwoordeboek van die Afrikaanse Taal.* Midrand: Perskor.

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

Taalkommissie van die S.A. Akademie vir Wetenskap en Kuns. 2002⁹. *Afrikaanse Woordelys en Spelreëls*. Kaapstad: Pharos.

Van Schalkwyk, D.J. 2002. Onderhoud, Stellenbosch, 12 Junie.

Requirements for an "Ideal" Bilingual $L_1 \rightarrow L_2$ Translation-Oriented Dictionary

Igor Burkhanov, *Section of Applied Linguistics, Institute of English Studies, University of Rzeszów, Rzeszów, Poland (burkhan@univ.rzeszow.pl)*

Abstract: The major aim of this article is to outline the requirements for an "ideal" bilingual $L_1 \rightarrow L_2$ dictionary of the general vocabulary specifically designed for the purposes of professional translation. The article challenges three commonly accepted beliefs: (a) a bilingual dictionary equals a translation dictionary; (b) a bilingual dictionary is a source of immediately insertable lexical equivalents of lemmata; and (c) a bilingual dictionary solely furnishes semantic-pragmatic equivalents, whereas a monolingual dictionary always defines meanings of lexical items. It has been claimed that bilingual lexicography should be based on a clear-cut conception of the future reference work specified in terms of lexicographic parameters: (a) the intended user group; (b) the purpose of the dictionary; etc. An "ideal" bilingual $L_1 \rightarrow L_2$ translation-oriented dictionary should be a reference work that is intended to serve the purpose of text production in L_2. In the situation of professional translation, L_2 text production is subject to two types of constraints: (a) constraints superimposed by the receptor language and culture; and (b) constraints superimposed by the source text. If constraints of the second type cannot be, in principle, envisaged, those of the first type can, and should be accounted for in a bilingual dictionary designed for the professional translator. The article specifies some of the requirements for such a reference work.

Keywords: BILINGUAL DICTIONARY, TRANSLATION DICTIONARY, SEMANTIC-PRAGMATIC EQUIVALENCE, INTERLINGUAL EQUIVALENCE, INTRALINGUAL EQUIVALENCE, DEFINING TECHNIQUE, EQUIVALENCE DEFINITION, PERIPHRASTIC DEFINITION, EXPLANATORY DEFINITION, PROFESSIONAL TRANSLATION, TRANSLATION THEORY, UNIT OF TRANSLATION.

Opsomming: Vereistes vir 'n "ideale" tweetalige $L_1 \rightarrow L_2$- vertaal-georiënteerde woordeboek. Die hoofdoel van hierdie artikel is om die vereistes te skets vir 'n $L_1 \rightarrow L_2$-woordeboek van die algemene woordeskat, spesifiek beplan vir die doeleindes van professionele vertaling. Die artikel weerlê drie algemeen aanvaarde opvattings: (a) 'n tweetalige woordeboek is dieselfde as 'n vertaalwoordeboek; (b) 'n tweetalige woordeboek is die bron van onmiddellik invoegbare leksikale ekwivalente van lemmata; en (c) 'n tweetalige woordeboek verskaf uitsluitlik semanties-pragmatiese inligting, terwyl 'n eentalige woordeboek altyd betekenisse van leksikale items omskryf. Daar word beweer dat tweetalige leksikografie gebaseer behoort te wees op 'n duidelik omskrewe begrip van die toekomstige naslaanwerk, gespesifiseer in terme van leksikografiese parameters: (a) die beoogde gebruikersgroep; (b) die doel van die woordeboek; ens. 'n "Ideale" tweetalige $L_1 \rightarrow L_2$- vertaal-georiënteerde woordeboek behoort 'n naslaanwerk te wees wat beplan is om die doel van teksproduksie in L_2 te dien. By die omstandigheid van professionele vertaling is

L$_2$-teksproduksie onderhewig aan twee soorte beperkings: (a) beperkings opgelê deur die teikentaal en -kultuur; en (b) beperkings opgelê deur die bronteks geskryf in L$_1$. As beperkings van die tweede soort in beginsel nie voorsien kan word nie, kan en behoort dié van die eerste soort verantwoord te word in 'n tweetalige woordeboek wat vir die professionele vertaler beplan is. Die artikel spesifiseer 'n aantal vereistes van so 'n naslaanwerk.

Sleutelwoorde: TWEETALIGE WOORDEBOEK, VERTAALWOORDEBOEK, SEMANTIES-PRAGMATIESE EKWIVALENSIE, INTERLINGUALE EKWIVALENSIE, INTRALINGUALE EKWIVALENSIE, DEFINIEERTEGNIEK, EKWIVALENSIEDEFINISIE, PERIFRASTIESE DEFINISIE, VERKLARENDE DEFINISIE, PROFESSIONELE VERTALING, VERTAALTEORIE, VERTAALEENHEID

This article forms part of a large-scale project that aims to outline the basic properties of "ideal" reference works intended to serve the purposes of translation and translator training, i.e. the intended user of those reference materials is either a person professionally involved in the process of mediated bilingual communication or a trainee translator.

This major objective can be achieved by attaining smaller goals, such as: (a) to specify the notion of professional translation; (b) to highlight the notion of equivalence in translation theory and lexicography; (c) to clarify the correlation between metalexicographic categories "bilingual dictionary" and "translation dictionary"; (d) to evaluate the state of the art in bilingual lexicography with respect to the purposes of professional translation; and (e) to spell out innovations that can be introduced in bilingual dictionaries in order to adjust them for the needs of the translator-user. These matters will be discussed in the appropriate sections of this article.

Pertinent questions will be asked, though not fully answered, in this article. It primarily aims to bring problems to the fore and illustrate the way in which translation-oriented dictionary-making should be developing. We shall mostly concentrate on translation-oriented bilingual dictionaries of the general vocabulary; the issues of terminological lexicography are beyond the scope of this article.

Before we proceed with these matters, the function of the adjective "ideal" in the context of the present article needs to be clarified. As I will try to show, reference works specifically produced for the purposes of professional translation are yet to be designed. The very term "ideal" translation-oriented dictionary" is intended to emphasize this idea.

Firstly, the adjective "ideal" is used here to show that the existing bilingual reference works do not totally satisfy the needs of a practicing translator. Secondly, the considerations below are not meant to be regarded as a description of a particular lexicographic project or projects, but rather as an attempt to outline at least certain basic principles of translation lexicography. Thirdly, a given dictionary, whether existing or intended to be compiled, is, from this viewpoint, an approximation to the non-existent ideal as specified by the require-

ments. Fourthly, we are not going to evaluate the feasibility of compiling translation-oriented bilingual dictionaries of a given pair of languages. Such an endeavour will require detailed analysis of the state of the art in bilingual (and monolingual) lexicography in both linguistic communities.

1.   The Notion of Professional Translation and Interlingual Equivalence in Translation Theory and Lexicography

A good deal of inconsistency in metalexicographic considerations concerning bilingual lexicography can be attributed to the variety of interpretations of the notion "translation". In both theoretical publications and introductions to actual dictionaries, it is not always clear what kind of translation the bilingual reference work is supposed to serve, since the notion of translation is erroneously believed to be self-explanatory.

It should be mentioned that professional translation differs from other activities that are also loosely referred to as "translation", primarily: (a) "school translation", i.e. finding native-language equivalents of foreign-language lexical items in order to facilitate acquisition of the latter; and (b) "translation for reading" which implies translating parts of a text so that those parts can be read more easily when needed (for a detail analysis see Gile 1991: 188). Thus, the delimiting adjective "translation-oriented" in terms such as "translation-oriented lexicography", "translation-oriented dictionary", etc., is intended to emphasize that the reference works under consideration should be *suited for the purposes of professional translation*. In our further considerations, unless specified, the term "translation" will be used in reference to professional translation only.

It should also be emphasized that the word "translation" may be used in at least two senses: (a) as an umbrella term designating mediated bilingual communication both in the spoken and written mode; and (b) in the narrower sense referring to translation in the written mode only, cf. translation vs. interpreting (e.g. see Gutknecht 2001). Naturally, bilingual dictionaries are extensively implemented in the process of translating in the second sense, i.e. mediated bilingual communication in the written mode. Conversely, it is difficult to make use of them in the course of consecutive or simultaneous interpreting, for interpreters usually do not have time to consult reference works.

From this it follows that bilingual lexicography primarily serves the purposes of professional translation in the written mode of communication. As for interpreting, reference works of this type, or any other reference materials for that matter, can be used solely in interpreter training.

A matter of primary significance for the issues under consideration is to highlight major aspects of professional translation. This category presupposes a specific form of communication in which there are several primary participants: (a) Communicant$_1$, i.e. a person (or a group of people) who uses the source language (L_1) and functions within the cultural setting of the appropri-

ate linguistic community; (b) Communicant$_2$, i.e. a person (or a group of people) using the target language (L$_2$), i.e. a different ethnic language that functions within another cultural background; (c) Communicant$_3$, the translator, a bilingual and bicultural person whose job is to negotiate linguistic and cultural discrepancies in the course of bilingual communication.[1]

The point to be noted here is that the basic assumption of lexicographers engaged in the production of bilingual dictionaries is that there is at least one obligatory L$_2$ lexical equivalent of the L$_1$ lexical item. It should be realized that this assumption is perfectly in line with the linguistically and/or semiotically oriented models of mediated bilingual communication that view the process of translating as encoding and decoding, i.e. a purely linguistic operation of substituting L$_1$ linguistic signs, particularly lexical items, with corresponding L$_2$ units (e.g. Catford 1965: 20; for a detailed analysis of these models see Burkhanov 2003a: 36-41). One of the consequences of this approach is the assumption of translatability, i.e. inter-replaceability of those signs with no loss of meaning (see, e.g. Petrilli 2001: 278). In practical terms, it means that the target text in L$_2$ can be easily "back-translated" into L$_1$ and the resulting text will be identical to the original. Obviously this assumption contradicts the hard reality of professional translation. Those models, though attractive due to their simplicity, do not account for all the aspects of the multifarious process of translating.

Though numerous and heterogeneous, sometimes even contradictory in nature, more recent translation theories view the process of translating as manipulation of the source text to adjust it to the requirements of the target culture, a specific form of human action, a communicative process which takes place within a given social context, an act of intercultural communication, etc. (see Gentzler 1993, Venutti 2000, Burkhanov 2003a: 41-55). In those theories the major emphasis is put not so much on the semantic-pragmatic content of lexical units, but on the inferencing of utterance meaning on the basis of contextual information, on the specific constraints as imposed by the discourse type, register, etc., as well as on other multifarious culture-specific aspects of communicative interaction.

In view of those developments in the theoretical study of translation, quite a number of experts have come to the conclusion that the very concept of translation equivalence is expendable, if not misleading (e.g. see Snell-Hornby 1988). Simultaneously, those theoretical solutions are, in turn, vulnerable to criticism, since in most cases they fail to draw a demarcation line between translation proper and adaptation. The starting point of our further considerations is that the category of equivalence is indispensable in both translation studies and bilingual lexicography, but it should not be understood as a homogeneous interlingual correspondence between lexical items belonging to L$_1$ and L$_2$.

It is important to note that the notion of equivalence as a major category of translation studies has evolved for the last fifty years. Firstly, various kinds of translation equivalence have been distinguished. It is noteworthy that the term

"equivalence" is usually provided with the delimiting adjective: "formal", "dynamic", "semantic", "pragmatic", "communicative", "functional", etc. From this it follows that an L_1 linguistic expression may be a semantic equivalent of an L_2 linguistic expression, but not pragmatically equivalent to the latter.

Secondly, it has been observed that translation equivalence varies in degree. For instance, according to Kade (1968), four types of equivalence can be distinguished: (a) *total equivalence* characteristic for standardized specialized terminology; (b) *facultative equivalence* that presupposes a one-to-many correspondence between L_1 and L_2 lexical units; (c) *approximating equivalence* that accounts for one-to-part-of-one correspondence; and (d) *zero equivalence* observable in the case of culture-bound lexical items.

Thirdly, and no less importantly, we may note the impact of correlated categories, such as "unit of translation", "translation adequacy" and "translational norms", which not only contributed to the development of translation studies, but also delimited the scope of application of the notion "equivalence" (for details see Burkhanov 2003a: 91-114).

A unit of translation is a functional-semantic category that is not restricted to the linear boundaries of linguistic structures and involves macro-textual parameters. So, the unit of translation cannot possibly be equated to an interlingual correlation describable in terms of one-to-one correspondences between L_1 and L_2 lexical items; it is a dynamic conceptual configuration spread over a large stretch of the text (Newmark 1988; Rabadán 1991). Simultaneously, the very notion of translation adequacy, as an optimal level of mediated bilingual communication in given circumstances of verbal interaction, implies selecting the most appropriate equivalents out of a variety of linguistic choices in the course of translating (Reiss and Vermeer 1984; Shveitser 1993). Moreover, translation strategies are dependent on historically-motivated cultural, literary and translational norms of the receptor culture (Toury 1980; Heylen 1993).

As for the notion of equivalence in lexicographic theory, Piotrowski (1994: 104) rightly pointed out that, though there are a large number of publications on bilingual lexicography, the notion of bilingual equivalence has not been fully specified yet. The greater majority of metalexicographers and practicing lexicographers assume that this basic category of bilingual lexicography is self-explanatory, or superficially formulate it in terms of translation equivalence understood as cross-linguistic correlations between lexical items, or ignore the problem altogether. One of a few attempts to fill the gap was undertaken by Wiegand (2003 and earlier publications). Nevertheless, though very interesting, his discussion tends to be limited to only one aspect of the issue under consideration, namely the lexicographic presentation of L_1 and L_2 lexical items, whereas their semantic-pragmatic equivalence is supposed to be determined at the earlier stages of lexicographic description.

The foregoing brief exposition has demonstrated that there is a considerable gap between the interpretation of equivalence, not to mention similar categories, in translation studies and what can be found in bilingual lexicogra-

phy. Bringing the concept of equivalent in bilingual lexicography into conformity with that of translation equivalent as developed in contemporary translation theory will only be profitable for this branch of lexicographic activity. Ideally, the optimal translation-oriented bilingual dictionary should be applicable in the translator's attempts to produce an adequate target text that conforms to the requirements of translational norms, which implies facilitating his/her search of appropriate units of translation.

2.   Bilingual Dictionary and Translation Dictionary: General Considerations

Our major goal in this section is to clarify the notion of bilingual dictionary and juxtapose it with that of translation-oriented dictionary. It is no secret that for the majority of experts in metalexicography those two terms are interchangeable (e.g. see Hartmann and James 1998: 146, Hannay 2003, Swanepoel 2003: 67-69). As has been mentioned, this line of reasoning presupposes that a monolingual dictionary provides definitions, whereas a bilingual dictionary furnishes interlingual equivalents, i.e. for each L_1 lexical unit that may be found in the word list as a lemma there is an L_2 lexical equivalent (or equivalents) presented in the right-hand side of the dictionary entry. We may say that the long-standing tradition of bilingual lexicography has been based on the assumption that for each and every L_1 lexical item there is at least one most appropriate L_2 lexical equivalent, and the characteristic features of bilingual reference works are specified in contrast with those of monolingual ones (cf Zgusta 1984: 147):

> The dictionary [bilingual dictionary — I.B.] should offer not explanatory paraphrases or definitions, but real lexical units of the target language which, when inserted into the context, produce a smooth translation. This is a perfectly natural requirement.

It should be noted that this statement not only contradicts the aforementioned major postulates of contemporary translation theory. Even more importantly, the ever-increasing vast number of reference works intended to meet the needs of different user groups do not fit into the neat classificatory framework presented above. This state of affairs has not escaped the attention of experts in lexicographic theory. Nevertheless, in metalexicography those reference works that display the features of both supposedly basic types are usually regarded as either hybrid works of reference or some sort of deviation from the general principle. For instance, Hartmann and James (1998) distinguish, in addition to a bilingual dictionary and a monolingual dictionary, another kind of lexicographic product: a bilingualised dictionary. A bilingualised dictionary is a reference work which is "based on a monolingual dictionary whose entries have been translated in full or in part into another language" (Hartmann and James 1998: 14).

On the basis of the definitions of pertinent notions provided in a dictionary of lexicographic terminology (Bergenholtz et al. 1997) and his own considerations, Tarp (2002) specifies the following three types of dictionaries: (a) a monolingual dictionary — a dictionary with only one object language, i.e. one language described; (b) a monolingual dictionary with a bilingual dimension — a dictionary that has one object language and another description language; and (c) a bilingual dictionary — a dictionary that has two object languages and provides equivalents in the target language for each word and expression in the source language. Obviously "a monolingual dictionary with a bilingual dimension" in the trichotomous division presented by Tarp corresponds to what Hartmann and James refer to as "a bilingualised dictionary".

Those and similar terminological distinctions have been introduced in order to save the aforementioned commonly accepted dichotomy: a monolingual dictionary furnishes definitions, i.e. defines, specifies, explicates meanings of lexical items, whereas a bilingual dictionary provides L$_2$ lexical items characterized by semantic-pragmatic equivalence to the L$_1$ lexical items represented by the lemmata. References to linguistic and cultural anisomorphism (discrepancies between a pair of languages arising due to their semantic, pragmatic, grammatical and cultural differences), which leads to the fact that "translation equivalents are typically partial, approximative, non-literal and asymmetrical (rather than full, direct, word-for-word and bidirectional)" (Hartmann and James 1998: 51), are usually made to pay the tribute to practicing lexicographers' efforts and to illustrate the difficulties they face rather than to draw conclusions that have far-reaching metalexicographic consequences and may lead to constructive solutions in bilingual lexicography. In the following, I will tentatively suggest a different approach that appears to be more realistic and takes into consideration the actual state of the art in dictionary-making.

First and foremost, the monolingual dictionary definition does not always *specify* the meaning of the entry word in the sense of providing its paraphrase or a sentential explication. One method of lexicographic explanation is based on intralingual lexical-semantic equivalence and consists in enumerating lexical items with similar meanings in the definiens, primarily synonymous equivalents and opposites in negation, particularly those that came to be known as complementaries or non-gradable antonyms. Cf.:

> **invent** *v.t.* <u>Devise</u>, <u>originate</u> (new method, instrument, etc.) … (COD4);
>
> **open** adj. **1** <u>not closed</u>; allowing (things, persons) to go in, out, through … (ALDCE3). [2]

It is evident that a definition on the basis of intralingual equivalence does not explicate the meaning of a lexical item, as much as periphrastic and/or explanatory definitions do, and in this respect is not very much different from a typical bilingual dictionary definition based on interlingual lexical semantic-pragmatic equivalence. Cf.:

> **legible** = <u>разборчивый</u>; <u>четкий</u>; <u>удобочитаемый</u> (о почерке, шрифте и т.п.) (NERD).

The use of synonyms and/or antonyms in negation in monolingual works of reference obviously leads to circularity of definitions and has been severely criticized by experts in metalexicography. Nevertheless, practicing lexicographers know that sometimes it is unavoidable. It is particularly the case when the lexicographer does not make use of explanatory, sentential definitions, and periphrastic definitions on the basis of a hyperonym cannot be provided due to the specificity of the lexical-semantic content of lemmata.

Moreover, if from the viewpoint of logic the defining techniques based on intralingual equivalence are commendable, there are purely applied-linguistic advantages in this way of defining. For instance, they afford the possibility to account for the lexical-semantic relations in the lexicon, particularly synonymy and oppositeness in meaning, which may be considered advantageous with regard to particular lexicographic functions[3]; hence most appropriate for certain types of lexicographic reference works.

Simultaneously, contrary to popular belief, there are bilingual dictionaries that furnish detailed definitions of the source-language lemmata in the target language. Very often those definitions are not only periphrastic, but also of an explanatory character. By way of example we may cite the German–Russian dictionary of synonyms (DRSW) where the discrimination of synonyms and usage specifications are provided in L_2, i.e. in Russian. Cf.:

> **Peinlich** неприятный тем, что связан с тягостным ощущением неловкости, смущения *и т. п.* [unpleasant, because it is associated with a painful feeling of being in an awkward, embarrassing situation, etc.].[4]

Meaning specifications of L_1 synonyms in L_2, in addition to the listing of their L_2 equivalents, may be said to be a characteristic feature of the English–Russian dictionary of synonyms (ERDS). Each dictionary article in that reference work provides an explanatory definition in Russian intended to specify the elements of meaning shared by all the synonyms, and a detailed explication of each member of the synonym series with meaning discrimination as well as notes that account for the senses of the synonyms at issue which are close to the meaning under consideration and comments concerning their figurative senses (for details see Burkhanov 2003b). In fact, those lexicographic works were produced in accordance with Ščerba's original idea that a learner-oriented bilingual dictionary should provide definitions of the foreign language lexical items in the learner's native language.

In addition, it should be emphasized that in case of a lexical gap, i.e. the situation when there is no even approximate L_2 equivalent of an L_1 lexical item whose meaning is often of a culture-specific nature, lexicographers unavoidably have to use a short interlingual paraphrase in a bilingual dictionary. Cf.:

> **casebook** = журнал для записи больных, клиентов, посетителей и *т.п.* [a book in which patients, clients, visitors, etc. are noted] (NERD);

> **buńczuk** = (oznaka wladzy [symbol of power]) horse-tail ensign (GPED).

Though defining in L$_2$ is often accompanied by furnishing even approximate L$_2$ semantic-pragmatic equivalents (if there are any), the very existence of the technique of lexicographic specification exemplified above is tell-tale evidence that all bilingual dictionaries not only provide interlingual equivalents, but may, and do define.

The aforementioned lexicographic facts warrant a different approach in typology of reference works in metalexicography. Instead of a binary opposition "monolingual dictionary" vs. "bilingual (or multilingual) dictionary" with supposedly opposing purposes, the differences between the actual reference works can be described in terms of two distinctive features or lexicographic parameters: "number of languages" and "defining technique".[5]

The lexicographic parameter "number of languages" distinguishes reference works in which only one language is implemented (monolingual) and those with more than one language (multilingual), the bilingual dictionary being a kind of multilingual work of reference. Thus, the terms "monolingual", "multilingual", "bilingual", etc. within this framework characterize only the number of languages that can be found in the reference work under consideration and are devoid of any connotations concerning its purpose.

The parameter "defining technique" refers to the actual way of specification of lexical meaning irrespective of the language (or languages for that matter) used in the right-hand side of the dictionary. Various types of lexicographic definitions (with regard to monolingual dictionaries, of course) have been discussed in many sources (e.g. see Hanks 1987; Nakamoto 1998; Geeraerts 2003). To keep this article within manageable proportions, we are not going to consider these matters here[6]. Suffice it to say that, in principle, a definition of any type (or a mixed one) can be implemented in both monolingual and bilingual reference works. By way of example let us consider the basic types:

(a) equivalence definition

(i)   in a monolingual dictionary (on the basis of intralingual equivalence):
      invent *vt* … 1 <u>create</u> or <u>design</u> (sth not existing before) … 2 <u>make up</u>, <u>think of</u> … (ALDCE3);

(ii) in a bilingual dictionary (on the basis of interlingual equivalence):
      legible = <u>czytelny</u> (GEPD);

(b) periphrastic definition

(i)   in a monolingual dictionary:
      coat = an outer garment with sleeves (NHLD);

(ii)   in a bilingual dictionary:
        exterminator = środek tępiący (robactwo itd.) [a substance used to exter-
            minate (insects, etc.)] (GEPD);

(c) ostensive definition

 (i)   in a monolingual dictionary:
        cutlery = knives, forks and spoons, used for eating and serving food
            (ALDCE6);

(ii)   in a bilingual dictionary:
        nephew = bratanek; siestrzeniec ['one's brother's son'; one's sister's son]
            (GEPD);

(d) descriptive definition

 (i)   in a monolingual dictionary:
        muggy. If the weather is muggy, it is unpleasantly warm and damp, and
            makes you feel uncomfortable or tired (COBUILD);

(ii)   in a bilingual dictionary:
        Eid присяга *по сравнению с* **Schwur** *подчеркивает публичный, официаль-
            ный характер клятвы* … [oath, *compared to* **Schwur**, *emphasizes the
            public, formal character of the oath*] (DRSW).

The choice of an appropriate combination of the number of languages and
defining techniques for a particular reference work should be primarily deter-
mined by the intended user and lexicographic function or functions, for in-
stance: (a) to provide potential candidates for translation equivalents (produc-
tion dictionary for translation purposes); (b) to ensure understanding of the
target text (reception dictionary for reading purposes); (c) to facilitate the
acquisition of a non-native language; etc.

     For instance, in the case of two foreign languages studied by students, a
possible solution can be a trilingual reference work where there are two object
languages and the students' native language is used to explicate similarities
and differences between them, thus performing the function of a metalanguage
of lexicographic description (Burkhanov 1993).

     Though by no means comprehensive, the foregoing discussion has pro-
vided sufficient background for the specification of at least certain require-
ments for translation-oriented bilingual dictionaries of the general vocabulary,
which will be our concern in the next section.

3.   Prospective Innovations in Translation Lexicography

A number of experts are very much aware of the fact that the existing bilingual
$L_1 \rightarrow L_2$ reference works do not completely satisfy the needs of professional

translation and/or are not sufficient for adequate translational performance. For instance, Tarp (2002) approaches the problem from a different perspective and argues that monolingual dictionaries and bilingual dictionaries "the other way around", i.e. $L_2 \rightarrow L_1$ dictionaries, can better solve certain problems arising in the course of translating from L_1 into L_2.

The cornerstone of the following considerations is that an "ideal" $L_1 \rightarrow L_2$ translation-oriented dictionary is always designed for text production. A brief and by no means exhaustive account of the requirements for a reference work of this kind presented below follows from this assumption.

First and foremost, even the slightest discrepancies in the semantic and/or pragmatic aspects of the source-language lexical item and its target-language equivalent or equivalents should be explicitly specified in the dictionary entry. Since all the cases of semantic-pragmatic lexical equivalence of a facultative, approximating or zero character may cause translation errors, they deserve the translator's attention and have to be described in the dictionary entry. Conversely, total equivalence does not have to be specifically noted, since, due to long-standing lexicographic convention, the lack of any commentary, usage note and/or lexicographic indicator of whatever kind will be correctly interpreted by the intended user. The translation-oriented bilingual dictionary will thus furnish exhaustive information concerning semantic-pragmatic discrepancies between interlingual equivalents.

It is evident that this matter is closely interrelated with the issue of lexicographic representation of pragmatic, culture-specific and encyclopaedic information in a bilingual dictionary of the general vocabulary, which has already been touched on by practicing lexicographers (e.g. Sundström 1996).

This goal can be achieved by the unrestricted use of periphrastic and/or explanatory definitions in L_2 of the conceptual content and usage of L_1 lexical units. It should be emphasized that this technique of lexicographic presentation is undoubtedly underused in contemporary bilingual dictionaries, which may be attributed to the aforementioned generally-accepted belief that lexicographic works of this kind are supposed to provide lexical equivalents only. Thus, paraphrases and explanatory definitions in L_2 should not be "smuggled" into a translation-oriented bilingual dictionary, but treated as an efficient technique of explicating the semantic-pragmatic content of L_1 lexical items.

Nowadays the translator is supposed to be intuitively aware of the potential chances of meaning that may arise in the course of translating (particularly unavoidable "overtranslation" and "undertranslation"), if the recommended interlingual equivalents are implemented. It is obvious that lexicographic specification of potential complications that may appear in the process of transposition of the source text into another language and culture will be only appreciated by practicing translators.

To cite but one example: The aforementioned Polish lexical items *bratanek* 'one's brother's son' and *siostrzeniec* 'one's sister's son' have the same English translation equivalent. The unavoidable use of *nephew* as an only equivalent of

any of those two words may lead to undertranslation in the course of rendering a Polish text into English. The translator can and should compensate for the deficiency in the other parts of the target text. From this viewpoint, a useful innovation in a Polish–English translation-oriented dictionary can be specifications of such instances of partial equivalence.

Simultaneously, information of this kind is indispensable in the course of translating from English into Polish. The fact that the L_1 lexical item *nephew* should be accounted for by one of its two "interlingual co-hyponyms", i.e. either *bratanek* or *siostrzeniec*, warrants the necessity to check the whole text and/or other available materials in order not to commit a translation error. A good solution in cases of approximating equivalence of this kind can be the use of lexicographic indicators like WARNING or ATTENTION intended to highlight those interlingual lexical-semantic discrepancies. It is important to note that an extensive lexicographic treatment of such cases is not required in a bilingual $L_1 \rightarrow L_2$ reception-oriented dictionary designed for reading L_1 texts.

If the available translational solutions may cause undue domestication of the target text, the translator can resort to borrowing. The use of this technique of translating has far-reaching lexicographic consequences. Among others, it warrants the necessity to suggest transliteration in a translation dictionary when the L_1 linguistic community uses a different alphabet (or the same alphabet, but enriched by letters that cannot be found in L_2 writing) or a different writing system. This is particularly the case when culture-specific terms are accounted for. The already mentioned Polish word *buńczuk* may serve as an example.

Firstly, the short interlingual paraphrase "horse-tail ensign" provided in a dictionary does not seem to be sufficient. In fact, the lexical item at issue was used in two senses; and, to serve the purposes of professional translation, the dictionary should provide explications of both. One sense designated a horse tail used as a decoration on the helmet or girdle. The second denoted a wooden shaft with a ball on top and a cross-bar with horse tails attached to it. It was used as a symbol of authority vested with *hetmans*, Polish commanders-in-chief, Cossack leaders (cf. Russian *бунчук*), Turks (cf. Turkish *bunğuk*), Kazakhs, and Tatars.

At first sight it seems irrelevant if, in the course of translating into English a Polish historical novel or a museum booklet, the word will be borrowed in its original spelling or will be transliterated in accordance with the rules of English orthography as *bunchuk*. Nevertheless, adapted spelling seems to be preferable in this case, since, as has been shown above, the same culture-specific term may be used in the source texts written in other languages.

Generally speaking, it seems appealing to suggest that the more potential lexical-semantic equivalents are provided either in the entry or by means of cross-referencing (ideally including interlingual hyperonyms and/or paraphrases, and suggested loanwords), the better the reference work will serve its purpose. Naturally, in view of the foregoing considerations, none of those solu-

tions may prove to be applicable in a particular instance of translating. Nevertheless, with a larger range of L$_2$ equivalents there is a better chance of furnishing a more adequate lexical item for a particular context, not to mention the possibility of leading the translator to the right choice by association.

For instance, Birkenhauer and Birkenhauer (1989: 91) maintain that literary translators primarily need monolingual dictionaries with onomasiological arrangement of their word list. Cf.:

> A dictionary that would present its material in our ideal way — as a network of interrelated meanings, ranging from true synonyms to mere neighbouring concepts that just share one or two semes — would be invaluable for translators (and other professional writers). For only in such a network does the translator have the chance of hitting upon potentially new equivalents that can do more justice to the phonological, syntactic, and semantic constraints of his original text.

It is difficult to contradict the statement that the translator's activity is similar to that of an L$_2$ writer. Nevertheless, there is a very significant difference, namely: the original writer has to comply with only those constraints, which are imposed by the linguistic peculiarities of L$_2$ and its cultural setting. As for the translator, his/her linguistic choices are additionally subject to the impact of the constraints superimposed by the source text.

In view of the considerations above, Snell-Hornby's proposals seem very promising (Snell-Hornby 1990: 222-224). She believes that the lexicographer should aim at the specification of lexical items both paradigmatically (intralingually) and contrastively (interlingually), i.e. both against other members of the appropriate lexical fields in the source language and in contrast with lexical equivalents of the target language. In her opinion, the major aim of translation lexicography is to design a contrastive dictionary of synonyms.

Moreover, I would go on to claim that interlingual semantic-pragmatic equivalence of lexical items is not sufficient to produce an adequate translation. From the standpoint of semantic typology of languages, it is clear that an L$_1$ lexical item can be translationally equivalent to a linguistic sign of a different kind: a morpheme, a functional word, a syntactic structure, or a linguistic expression of a complex nature. For instance, the Russian functional word *ведь* and its Polish equivalent *przecież* are pragmatically equivalent to a tag question with falling intonation in English (for details see Burkhanov 1997 and 2002).

Another instance of interlingual semantic-pragmatic correspondences can be exemplified by the designation of the notions 'DYING' and 'KILLING'. These two concepts are lexicalized in English and practically all European languages, cf. *die* vs. *kill*. In the Turkic languages, the notions at issue are designated by the opposition of two word forms of the same lexeme, cf. *ölmäk* vs. *öldürmäk* (Azerbaijani); *ölmek* vs. *öldürmek* (Turkish). Bilingual dictionaries have not been very efficient in the lexicographic representation of semantic-prag-

matic equivalence of units of various symbolic levels, i.e. lexical items, morphemes and syntactic structures.

Translation-oriented bilingual dictionaries of the general vocabulary should be compiled within the framework of an overall applied-linguistic description of both languages, with semantic typology of languages being the theoretical basis of such a description. In practical terms, it will mean that in cases when the semantic-pragmatic content of a source-language lexical item is not lexicalized in the target language, the equivalent inflectional word form(s) and/or syntactic structure(s) (with appropriate intonation patterns, if needed) should be furnished in the bilingual translation-oriented dictionary.

Moreover, it should be clear from the exemplification above that a dictionary specifying intralingual and interlingual lexical-semantic correspondences is not sufficient. Probably an optimal solution is a reference work featuring semantic and pragmatic correlations of not only lexical units but also those of all the symbolic units within semantic domains of various kinds, i.e. frames, scripts and semantic fields. In this case, the reference work under consideration will not be a dictionary of synonyms, but a bilingual onomasiological (ideographic) dictionary provided with an accompanying communicative grammar.

However, the discussion concerning translation lexicography should in the first place focus on the problems of *selection* of lexicographic (and grammaticographic) data and *adequate description* of appropriate linguistically and/or culturally significant facts enabling the user to *produce* L$_2$ texts. Matters of lexicographic presentation, i.e. whether lexicographic information should be accounted for in an alphabetically-organized word list, in contrasted synonyms groups, or within an onomasiological classificatory scheme seems to be a question of editorial decision. Of course, an onomasiological arrangement of lexical items seems very appealing for a bilingual dictionary designed for L$_2$ text production. Nevertheless, it seems feasible to present the same lexicographic data in a semasiologically organized word list.

The existing bilingual dictionaries are often criticized for an insufficient description of the syntagmatic potential of lexical items. It should be emphasized that this requirement primarily concerns translation-oriented reference works. Reception-oriented bilingual dictionaries are not to account for those linguistic phenomena.

Thus, one of the requirements for a translation dictionary should be extensive information concerning the syntagmatic properties of the lexical items of the target language, i.e. their syntactic and lexical-semantic valence and collocational range. As I have mentioned elsewhere, certain regularly co-occurring words may seem to be free word combinations from the standpoint of a native speaker of a language but clearly display the characteristics of set phrases from the non-native speaker's viewpoint. Nevertheless, they are language-specific, and as such, have to be noted in a learner's dictionary designed for a non-native speaker-user (Burkhanov 2003c: 110). For instance, in the Slavonic languages the lexical-semantic correlate of *be in prison* literally means 'sit in prison',

cf. *сидеть в тюрьме* (Russian) or *siedzieć w więzieniu* (Polish). There is even more reason to assume that those phrases should be featured in the word list of a bilingual translation-oriented work of reference.

A translation-oriented reference work should not only account for the syntagmatic potential of all the lexical items, i.e. their lexical-semantic and syntactic valence and collocability. It should also represent language-specific linguistic expressions that Mel'čuk (1995) designated as pragmatemes, i.e. pragmatically constrained expressions which, semantically, are free word combinations but cannot be substituted by synonymous linguistic expressions in certain sociocultural contexts; for instance the inscription on perishable groceries specifying the expiry date: *best before …* . Its equivalents vary from language to language, cf. *à consommer avant …* ['for consuming before'] (French), *haltbar bis …* ['keepable until'] (German), *sporzyć przed …* ['to consume before'], *срок годности до …* ['period of usability until'] (Russian), etc. It seems reasonable to assume that such expressions should be included into the word list of a bilingual $L_1 \rightarrow L_2$ translation-oriented dictionary.

In general, translation-oriented lexicography should be developing in the direction of accounting for complex units of translation rather than cross-linguistic correspondences between individual lexical items.

It should be clear from the considerations above that in the meantime a practicing translator has to compensate for the deficiency and have at his/her disposal a set of bilingual and monolingual lexicographic works of reference which not only provide target-language lexical-semantic equivalents, but also, and at least no less importantly, contain lexicographic data concerning target-language synonym discrimination, collocations and usage.

4.   Concluding Remarks

The majority of the existing bilingual dictionaries are compiled on the basis of the postulated cross-linguistic semantic-pragmatic correspondence between lexical items of an obligatory character. It has been demonstrated that the bilingual dictionary cannot, in principle, always be the source of immediately insertable translation equivalents.

The foregoing considerations are not intended to undermine the very *raison d'être* of bilingual dictionaries. They aim to emphasize that bilingual lexicography should be based on a clear-cut conception of the future reference work specified in terms of a number of lexicographic parameters, the principal ones being "intended user group" (foreign language learners vs. professional translators, etc.) and "the purpose of the dictionary" as a utility tool (text production in L_2 vs. text comprehension in L_1, etc.). Of primary importance in this context is the specification of the lexicographic function(s) the bilingual work of reference is intended to fulfil.

It has been claimed that a translation-oriented bilingual $L_1 \rightarrow L_2$ dictionary should be a reference work intended to serve the purpose of text production in

L$_2$. In the situation of professional translation, L$_2$ text production is subject to two types of constraints: (a) constraints as superimposed by the receptor language and culture; and (b) constraints as superimposed by the source text written in L$_1$. If constraints of the second type cannot be, in principle, envisaged, those of the first type can, and should, be accounted for in a bilingual dictionary designed to satisfy the needs of the professional translator. Certain requirements for an "ideal" bilingual translation-oriented reference work have been specified in this connection.

What has not been considered in this article is to what extent the requirements for a translation-oriented bilingual dictionary overlap with those for other types of bilingual reference works which are designed to perform different lexicographic functions, for instance: bilingual L$_1$→L$_2$ learner-oriented dictionaries or bilingual L$_1$→L$_2$ reception-oriented dictionaries. These are issues of primary significance and they still await principled solutions.

Endnotes

1. Gile (1991) also mentions the Source Language Reader in the case of communication in the written mode and the Client, i.e. a person or an organization, who or which commission translator's service, as an indispensable element of professional translation in both modes. Nevertheless, these two participants of the process of translating are of no importance for our present purposes.
2. The cases of intralingual and interlingual equivalence are underlined in the definitions.
3. For a detailed treatment of the notion "lexicographic function" see Bergenholtz and Tarp (2003) and Bergenholtz (2003).
4. Literal translations of lexical equivalents are provided in square brackets.
5. The analysis of the notion "lexicographic parameter" can be found in Burkhanov (1998).
6. They should be described with regard to other ways of lexicographic explication of meaning: exemplification, pictorial representations, etc.

Bibliography

Dictionaries

ALDCE3. **Hornby, A.S. (Ed.).** 1974. *Oxford Advanced Learner's Dictionary of Current English.* Oxford: Oxford University Press.

ALDCE6. **Wehmeier, Sally (Ed.).** 2000. *Oxford Advanced Learner's Dictionary of Current English.* Oxford: Oxford University Press.

COBUILD. **Sinclair, John M. (Ed.).** 1995. *Collins Cobuild English Dictionary*. London: HarperCollins.

COD4. **Fowler, H.W. and F.G. Fowler (rev. by F. McIntosh).** 1951. *The Concise Oxford Dictionary of Current English.* Oxford: Oxford University Press.

DRSW. **Rachmanow, I.W. et al. (Eds.).** 1983. *Deutsch–Russisches Synonymwörterbuch.* Moscow: Russkii Jazyk.

ERDS. **Apresjan, Juri. D. (Ed.).** 1979. *Anglo–russkii sinonimiceskii slovar'.* Moscow: Russkii Jazyk.

GEPD. **Stanislawski, Jan.** 1964. *The Great English–Polish Dictionary*. 2 volumes. Warsaw: Wiedza Powszechna.

GPED. **Stanislawski, Jan.** 1969. *The Great Polish–English Dictionary*. 2 volumes. Warsaw: Wiedza Powszechna.

NERD. **Galperin, Ilia R. (Ed.).** 1972. *New English–Russian Dictionary*. 2 volumes. Moscow: Sovetskaia Encyklopedia.

NHLD. **Shaw, John Robert and Sarah Janet Shaw.** 1990. *The New Horizon Ladder Dictionary*. New York: Signet.

Other literature

Bergenholtz, Henning. 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.

Bergenholtz, Henning and Sven Tarp. 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 171-196.

Bergenholtz, Henning et al. 1997. *Nordisk Leksikografisk Ordbok*. Oslo: Universitetsforlaget.

Birkenhauer, Klaus and Renate Birkenhauer. 1989. Shaping Tools for the Literary Translator's Trade. Snell-Hornby, Mary and Esther Pöhl (Eds.). 1989. *Translation and Lexicography*: 89-98. Amsterdam/Philadelphia: John Benjamins.

Burkhanov, Igor. 1993. Некоторые вопросы составления идеографического словаря при параллельном обучении русскому и английскому языкам. Jochym-Kuszlik, Ludwika (Ed.). 1993. *Historia i teraźniejszość Rosji w świetle faktów językowych*: 43-50. Kraków: WSP.

Burkhanov, Igor. 1997. Bilingual Dictionaries in Pedagogical Lexicography. Lewandowska-Tomaszczyk, Barbara and Marcel Thelen (Eds.). 1997. *Translation and Meaning. Part 4*: 443-450. Maastricht: Universitaire Pers Maastricht, Hogeschool Maastricht, School of Translation and Interpreting.

Burkhanov, Igor. 1998. *Lexicography: A Dictionary of Basic Terminology*. Rzeszów: University Press.

Burkhanov, Igor. 2002. Contrastive Lexicology, Semantic Typology of Languages, and Translation Theory. Lewandowska-Tomaszczyk, Barbara and Marcel Thelen (Eds.). 2002. *Translation and Meaning. Part 6*: 313-323. Maastricht: Universitaire Pers Maastricht, Hogeschool Maastricht, School of Translation and Interpreting.

Burkhanov, Igor. 2003a. *Translation: Theoretical Prerequisites*. Rzeszów: University Press.

Burkhanov, Igor. 2003b. Juri Apresjan and the Development of Semantics and Lexicography. *Lexikos* 13: 287-305.

Burkhanov, Igor. 2003c. Pragmatic Specifications: Usage Indications, Labels, Examples; Dictionaries of Style, Dictionaries of Collocations. Van Sterkenburg, Piet (Ed.). 2003: 102-113.

Catford, J.C. 1965. *A Linguistic Theory of Translation*. London: Oxford University Press.

Geeraerts, Dirk. 2003. Meaning and Definition. Van Sterkenburg, Piet (Ed.). 2003: 83-93.

Gentzler, Edwin. 1993. *Contemporary Translation Theories*. London/New York: Routledge.

Gile, Daniel. 1991. A Communication-Oriented Analysis of Quality in NonLiterary Translation and Interpretation. Larson, Mildred L. (Ed.). 1991: 188-200.

Gutknecht, Christoph. 2001. Translation. Aronoff, Mark and Janie Rees-Miller (Eds.). 2001. *The Handbook of Linguistics*: 692-703. Oxford: Blackwell.

Hanks, Patrick. 1987. Definitions and Explanations. Sinclair, John (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.

Hannay, Mike. 2003. Types of Bilingual Dictionaries. Van Sterkenburg, Piet (Ed.). 2003: 145-153.

Hartmann, Reinhart and Gregory James. 1998. *Dictionary of Lexicography.* London/New York: Routledge.

Heylen, Romy. 1993. *Translation, Poetics, and the Stage: Six French Hamlets*. London/New York: Routledge.

Kade, O. 1968. *Zufall und Gesetzmäßigkeit in der Übersetzung.* Leipzig: Enzyklopädie.

Larson, Mildred L. (Ed.). 1991. *Translation: Theory and Practice/Tension and Interdependence.* American Translators Association Scholarly Monograph Series. Volume 5. Binghamton: State University of New York at Binghamton (SUNY).

Mel'čuk, Igor A. 1995. Phrasemes in Language and Phraseology in Linguistics. Everaert, M. et al. (Eds.). 1995. *Idioms: Structural and Psychological Perspectives*: 167-232. Hillsdale: Lawrence Erlbaum Associates.

Nakamoto, K. 1998. From which Perspective does the Definer Define the Definiendum: Anthropocentric or Referent-based? *International Journal of Lexicography* 11(3): 205-218.

Newmark, Peter. 1988. *A Textbook of Translation.* New York/London: Prentice Hall.

Petrilli, Susan. 2001. Translation. Cobley, Paul (Ed.). 2001. *The Routledge Companion to Semiotics and Linguistics*: 278-279. London/New York: Routledge.

Piotrowski, Tadeusz. 1994. *Problems in Bilingual Lexicography.* Wrocław: University Press.

Rabadán, Rosa. 1991. The Unit of Translation Revisited. Larson, Mildred L. (Ed.). 1991: 38-48.

Reiss, Katarina and Hans J. Vermeer. 1984. *Grundlegung einer allgemeinen Translationstheorie.* Linguistische Arbeiten 147. Tübingen: Niemeyer.

Shveitser, Alexander D. 1993. Equivalence and Adequacy. Zlateva, Palma (Ed.). 1993. *Translation as Social Action: Russian and Bulgarian Perspectives*: 47-56. London/New York: Routledge.

Snell-Hornby, Mary. 1988. *Translation Studies: An Integrated Approach.* Amsterdam/Philadelphia: John Benjamins.

Snell-Hornby, Mary. 1990. Dynamics in Meaning as a Problem for Bilingual Lexicography. Lewandowska-Tomaszczyk, Barbara and Jerzy Tomaszczyk (Eds.). 1990. *Meaning and Lexicography*: 209-225. Linguistic and Literary Studies in Eastern Europe 28. Amsterdam/Philadelphia: John Benjamins.

Sundström, Mats-Peter. 1996. The Discrepancy between Dictionaries and Language Reality. Scattered Notes by a Dictionary Editor cum Translator. Thelen, Marcel and Barbara Lewandowska-Tomaszczyk (Eds.). 1996. *Translation and Meaning. Part 3*: 233-237. Maastricht: Universitaire Pers Maastricht, Hogeschool Maastricht, School of Translation and Interpreting.

Swanepoel, Piet. 2003. Dictionary Typologies: A Pragmatic Approach. Piet van Sterkenburg (Ed.). 2003: 44-69.

Tarp, Sven. 2002. Translation Dictionaries and Bilingual Dictionaries – Two Different Concepts. *Journal of Translation Studies* 7: 59-84.

Toury, Gideon. 1980. *In Search of a Theory of Translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics.

Van Sterkenburg, Piet (Ed.). 2003. *A Practical Guide to Lexicography.* Amsterdam/Philadelphia: John Benjamins.

Venutti, Lorence (Ed.). 2000. *The Translation Studies Reader.* London/New York: Routledge.

Wiegand, Herbert Ernst. 2003. On the Lexicographic Description of Equivalent Open Class Expressions. *Lexikos* 13: 38-64.

Zgusta, Ladislav. 1984. Translational Equivalence in the Bilingual Dictionary. Hartmann, Reinhard (Ed.). 1984. *LEXeter '83 Proceedings*: 147-154. Tübingen: Niemeyer.

The Lexicographic Treatment of the Demonstrative Copulative in Sesotho sa Leboa — An Exercise in Multiple Cross-referencing[*]

Gilles-Maurice de Schryver, *Department of African Languages and Cultures, Ghent University, Ghent, Belgium and Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (gillesmaurice.deschryver@UGent.be)*,
Elsabé Taljard, *Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (etaljard@postino.up.ac.za)*,
M.P. Mogodi, *Sesotho sa Leboa National Lexicography Unit, Pretoria Branch and Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (pmogodi@postino.up.ac.za)*, and
Salmina Maepa, *Department of Arts and Culture and Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (salmina.nong@dac.gov.za)*

Abstract: In this research article an in-depth investigation is presented of the lexicographic treatment of the demonstrative copulative (DC) in Sesotho sa Leboa. This one case study serves as an example to illustrate the so-called 'paradigmatic lemmatisation' of closed-class words in the African languages. The need for such an approach follows a discussion, in Sections 1 and 2 respectively, of the present and missing directions in African-language metalexicography. A theoretical conspectus of the DC in Sesotho sa Leboa is then offered in Section 3, while Section 4 examines the treatment of the DC in the four existing desktop dictionaries for this language. The outcomes from the two latter sections are then used in Section 5, which analyses the problems of and options for a sound lexicographic treatment of the DC in bilingual and monolingual dictionaries. The next two sections proceed with a review of the practical implementation of the DC lemmatisation suggestions in PyaSsaL, i.e. the *Pukuntšutlhaloši ya Sesotho sa Leboa* 'Explanatory Sesotho sa Leboa Dictionary' — with Section 6 focussing on the hardcopy and Section 7 on the online version. In the process, the very first fully monolingual African-language dictionary on the Internet is introduced. Section 8, finally, concludes briefly.

Keywords: LEXICOGRAPHY, PARADIGMATIC LEMMATISATION, AFRICAN LAN-GUAGES, SESOTHO SA LEBOA (NORTHERN SOTHO, SEPEDI), DEMONSTRATIVE COPULA-TIVE, CROSS-REFERENCING, CORPUS, MONOLINGUAL DICTIONARY, BILINGUAL DIC-

---

TIONARY, HARDCOPY, ONLINE, INTERNET, EXPLANATORY SESOTHO SA LEBOA DIC-TIONARY (PᴛᴀSsᴀL), SIMULTANEOUS FEEDBACK (SF)

Senaganwa: Tokelotlhalošo ya lešalašupi-leba ka mo pukuntšung ya Seso-tho sa Leboa — Tirišo ka go šupana go gontši. Ka go sengwalwana se sa nyakišišo, nyakišišo yeo e tseneletšego e laetšwa ka ga go lokelwa le go hlalošwa ga lešalašupi-leba ka mo pukuntšung ya Sesotho sa Leboa. Thutwana ya mohuta wo ya nyakišišo e šoma bjalo ka mohlala go laetša seo se bitšwago 'tokelo ya mantšu ka lenaneo' (paradigmatic lemmatisation) ya mantšu a legoro leo le tswaletšwego ka go maleme a Afrika. Tlhokego ya nyakišišo ya mohuta wo e tla ka morago ga therišano ya ditaetšo tša gonabjale le tšeo di sego gona ka go tlhamopukuntšu ya teori ya maleme a Afrika. Ditaba tše di hlalošwa ka go dikarolo 1 le 2. Tlhalošo ya teori ya lešalašupi-leba ka go Sesotho sa Leboa e fiwa ka go karolo 3, mola karolo 4 e hlahloba tokelo le tlhalošo ya lešalašupi-leba ka go dipukuntšu tše nne tšeo di lego gona mo polelong ye. Dipoelo tša dikarolo 3 le 4 di šomišwa karolong ya 5, yeo e sekasekago mathata le dikgonego tša tokelotlhalošo ya lešalašupi-leba ka go dipukuntšu tša polelopedi le tša polelotee. Dikarolo tše pedi tšeo di latelago di tšwela pele ka tekolo tirišong ya dikakanyetšo tša tšhomišo ya lešalašupi-leba ka go PyaSsaL, e lego *Pukuntšutlhaloši ya Sesotho sa Leboa*. Karolo 6 e lebane le taodišwana ya pampiri mola karolo 7 e lebane le taodišwana ya Inthanete. Ka go dira bjalo, pukuntšu ya mathomothomo ya polelotee ya maleme a Afrika e tsebagatšwa mo Inthaneteng. Mafelelong karolo 8 e fa kakaretšo ka bokopana.

Mantšu a bohlokwa: TLHAMOPUKUNTŠU, TOKELO YA MANTŠU KA LENANEO, MALEME A AFRIKA, SESOTHO SA LEBOA, LEŠALAŠUPI-LEBA, TŠHUPANO, KHOPHASE, PUKUNTŠU YA POLELOTEE, PUKUNTŠU YA POLELOPEDI, PUKUNTŠU YA PAMPIRI, KA GO INTHANETE, INTHANETE, PUKUNTŠUTLHALOŠI YA SESOTHO SA LEBOA (PᴛᴀSsᴀL), SIMULTANEOUS FEEDBACK (SF)

1.   Present directions in African-language metalexicography[1]

For over a decade now, African-language metalexicography has become in-creasingly popular in South Africa. No doubt, the new lexicographic dispensa-tion in the now officially eleven-lingual South Africa has been instrumental in boosting interest in this field. Three main directions may be observed. Firstly it is noticed that a substantial number of corpus-based lexicographical studies for the African languages are being produced, starting with Prinsloo's (1991) 'com-puter-assisted word frequency studies', and culminating a decade later in a string of suggestions for corpus-building as well as considerations for diction-ary-making on the macro- and microstructural levels (e.g. De Schryver and Prinsloo 2000b, 2000c, 2000d, 2001, 2003; Prinsloo and De Schryver 2001; De Schryver 2002). A second direction of research has been the development of concepts and tools for lexicography in the modern age. These include, inter alia, the concepts of *Simultaneous Feedback* or 'SF' (De Schryver 1999; De Schry-ver and Prinsloo 2000, 2000a) and *Fuzzy SF* (De Schryver and Prinsloo 2001a), as well as tools such as *Multidimensional Lexicographic Rulers* and *Block Systems* (Prinsloo and De Schryver 2002, 2003, 2004, 2004a; De Schryver 2003b) and the

dictionary compilation software *TshwaneLex* (Joffe et al. 2003, 2003a). On a third level lemmatisation studies proper can be grouped. Research articles in this field are typically entitled 'Lemmatisation of ...', and this 'formula' has been particularly successful for Sesotho sa Leboa. As with corpora for the South African languages, the formula was first set out by D.J. Prinsloo, with colleagues following suit. By way of example, Table 1 lists the most influential attempts for Sesotho sa Leboa in this regard.

Table 1: The 'lemmatisation of ...'-formula in the case of Sesotho sa Leboa

Topic: Lemmatisation of ...	Author(s)	Year	Journal/Proc.
Reflexives	Prinsloo	1992	*Lexikos* 2
Verbs	Prinsloo	1994	*SAJAL* 14(2)
Verbs (..ga/sa/se..~ convention)	Prinsloo and Gouws	1996	*SAJAL* 16(3)
Adjectives	Gouws and Prinsloo	1997	*Lexikos* 7
Nouns	Prinsloo and De Schryver	1999	*SAJAL* 19(4)
Days	De Schryver and Lepota	2001	*Lexikos* 11
Verbs (freq.-based tail slots)	De Schryver and Prinsloo	2001	*Kiswahili* 2000
Abbreviated nouns	Bosch and Prinsloo	2002	*SAJAL* 22(1)
Loan words	Nong, De Schryver and Prinsloo	2002	*Lexikos* 12
Copulatives	Prinsloo	2002	*Lexikos* 12
Adverbs	Prinsloo	2003	*Lexikos* 13

As can be deduced from Table 1, with this formula African source-language lexical items belonging to distinct word classes are analysed, and suggestions for lemmatisation are offered for each of them, viz. for different types of verbs (Prinsloo 1992, 1994; Prinsloo and Gouws 1996; De Schryver and Prinsloo 2001b), for different types of nouns (Prinsloo and De Schryver 1999; Bosch and Prinsloo 2002), for adjectives (Gouws and Prinsloo 1997), for adverbs (Prinsloo 2003), for copulatives (Prinsloo 2002), and even for specific lexical sets (De Schryver and Lepota 2001) and loan words (Nong et al. 2002).

The 'lemmatisation of ...'-formula is typically concerned with methods to 'enter' or 'list' African-language items in the macrostructure of a dictionary, or thus in ways to 'lemmatise' various related forms under a single 'dictionary citation form'. This process can also be seen as a search for the most suitable 'canonical form' of each lexical item. Given that lexical items can successfully be grouped into types of words, known as 'grammatical classes' or 'parts of speech' (POSs), it does make sense to suggest lemmatisation approaches for each main type of African-language POS. Implicitly, the 'lemmatisation of ...'-formula is most relevant to dictionaries with an African language as source language, typically a bilingual dictionary treating an African language in the macrostructure with translation equivalents in a language of wider diffusion,

or else an explanatory African-language dictionary. Note that even though this type of lemmatisation deals mainly with macrostructural aspects, certain suggestions also have representation repercussions on the microstructural level (such as, cf. Table 1, the introduction of the ..ga/sa/se..~ convention or of frequency-based tail slots).

2.   Missing directions in African-language metalexicography

There seem to be two crucial aspects that have received very little if any metalexicographical attention so far, viz. (a) the treatment of an African language in the reverse side of a bilingual dictionary, where it is thus used for translating and/or paraphrasing another language, and (b) the 'paradigmatic lemmatisation' of closed-class words in the African languages. A metalexicographical discussion of a combination of (a) and (b) is of course even harder to come by.

    These aspects can best be illustrated with an example from Rycroft's (1981) *Concise SiSwati Dictionary*. In the 100-million-word *British National Corpus* (BNC), 'its' — the third person singular possessive determiner (det-poss) — has a frequency of 163 081, which makes it the sixty-second most frequent word of the English language (Kilgarriff 1996). No dictionary with English as one of its (!) treated language pairs may thus omit this word, not even a junior dictionary (cf. De Schryver and Prinsloo 2003). However, the det-poss 'its' has not been entered into the *English–siSwati* side of Rycroft's dictionary. The (very) diligent dictionary user may however realise that 'its' can be derived from that dictionary's table in the front matter which summarises the concordial agreement system in siSwati. Table 2 shows the relevant (and simplified) section.

Table 2: Deriving the translation equivalents for 'its' in siSwati (patterned on Rycroft 1981: xxiv, with highlighting added)

3rd p: Class	Possessive Stem	Possessive Concord	
1	-âkhe	w-(e/a)-	
2	-âbo	b-(e/a)-	
3	-âwo	w-(e/a)-	
4	-âyo	y-(e/a)-	*POSSESSIVE STEMS:* These, denoting the 'pos-
5	-âlo	l-(e/a)-	sessor', occur only with a possessive concord
6	-âwo	ø-(e/a)-	agreeing with the class of the item possessed.
7	-âso	s-(e/a)-	The initial vowels shown here were relegated
8	-âto	t-(e/a)-	to the concord by previous analysts. But as
9	-âyo	y-(e/a)-	their tone (high or falling) is determined by
10	-âto	t-(e/a)-	the particular stem, it is better to include
11	-âlo	lw-(e/a)-	them with the stems.
14	-âbo	b-(e/a)-	
15	-âko	kw-(e/a)-	
16-18	-âko	kw-(e/a)-	

From Table 2 one may derive that the det-poss 'its' for a possessor from class 5 or 11 possessing something in classes 8 or 10 translates as *talo*, while it for instance becomes *bayo* for a possessor in class 9 possessing people (in class 2) or say abstract nouns (in class 14).[2]

With this brief example the very heart of the African-language system, in this case siSwati, is being touched. Indeed, it is not enough to lemmatise the items from the second column of Table 2 (-*âkhe* to -*âko*) in the siSwati–English side (as has been done in Rycroft's dictionary); when reversing the dictionary one must also provide the necessary clues under the relevant reversed entries. In this case, each highlighted possessive stem (-*âwo*, -*âlo*, -*âso*, -*âyo*, -*âlo*, -*âbo*, -*âko* and -*âko*) can combine with all the possessive concords (*w-(e/a)-* to *kw-(e/a)-*), which thus means that there are 8 times 14 or 112 possible ways to say 'its' in siSwati. In other words, in the reverse side of an African-language dictionary, the lemmatisation aspects/problems become microstructural design aspects/ problems. The dictionary compiler will have to decide whether or not to give all 112 forms in full, or only the frequent ones, and whether or not to provide examples for those treated. This decision will of course have to take the intended target user into account, and in today's dictionary landscape also whether the output is to paper (with severe space restrictions) or to an electronic format (with virtually no space restrictions).

In the previous discussion, the notion of 'paradigmatic lemmatisation' was implicit. Given that all nouns belong to classes in the African languages, and given that this membership drives the entire concordial agreement system, it is actually surprising that so little attention has been paid so far to this aspect in African-language metalexicography. The basic question is: If certain lexical items belong to a closed set or 'paradigm' driven by the class system, is it enough to randomly sample a few of its members for purposes of (a) lemmatisation, (b) microstructural representation in the source language, and, for a bilingual dictionary, (c) reversing the African-language source side? A random approach, even if based on intuition, does not seem appropriate in today's corpus-based/-driven lexicographical framework.

Paradigms of closed-class items, which by definition contain a *limited* number of members, are numerous in the African languages: object concords, subject concords, possessive concords, various types of demonstratives, various types of pronouns, etc. When there are moreover two dimensions, as in the case of possessives (cf. 'its' above), the paradigms may also become relatively large. Even though such paradigms are core blocks of the grammatical systems of the African languages — and thus important to both mother-tongue speakers learning another language (who e.g. have to be able to *map* 112 environments of 'its' on a single one) and learners of an African language (who have to do the reverse) — dictionaries notoriously fail to take the paradigms seriously.

Although paradigms are important when reversing a dictionary with an African language as the source language, whereby paradigms enter the microstructure on the reverse side, these will not be the focus of this article. Paradigms can further also play a role in the African-language source side, whether

in a bilingual or a monolingual dictionary, as illustrated for the various concord and pronoun paradigms introduced in Prinsloo and De Schryver (2002a: 81, 2002b: 173-174, 177). These will not be the focus of this article either. In its most narrow sense 'paradigmatic lemmatisation' refers to the 'lemmatisation of ...'-formula, but now applied to paradigms. For the African languages these for instance include various types of adjectives, but also numerous more traditional closed-class words. One instance of the latter, namely the demonstrative copulative in Sesotho sa Leboa, will now be presented. It should be clear from the outset, however, that even though just one single case study is presented for one African language, the implications are generic.

3. The demonstrative copulative (DC) in Sesotho sa Leboa: A theoretical conspectus

Demonstrative copulatives are primarily nominal determiners, appearing in either the pre-nominal or post-nominal position, the post-nominal position being the dominant one. Like all nominal determiners in Sesotho sa Leboa, they can also function as pronominal forms in cases where the nominal antecedent is deleted. Although the term 'demonstrative copulative' (DC) is a somewhat cumbersome one, it does give an apt description of the semantic nature of these forms. Its demonstrative meaning is vested in the fact that it specifies the locality of a person or object relative to the position of the speaker and addressee in terms of different positions. It also has a copulative meaning and is thus loosely translated as 'here is/are', 'there is/are', etc., representing a predicative form of the demonstrative. Compare the examples in (1).

(1)

Bašemane *šeba*	'*Here are* the boys'	[*šeba* = DC position I, class 2]
Kgoši *šeo*	'*There is* the chief'	[*šeo* = DC position II, class 9]
Mohlare wo mogolo *šola*	'*Over there is* the big tree'	[*šola* = DC position III, class 3]

Sesotho sa Leboa grammarians do not agree as to the number of positions that are to be distinguished, nor is there any consistency in the numbering/labelling of the positions which they do distinguish. It would seem that all grammarians recognise at least three basic positions, i.e. (a) a basic form consisting of a root *š(e)-* (*se-* for class 7), followed by a concordial morpheme, (b) a second form consisting of the basic form to which a raised *-ô* has been suffixed, and (c) a third form, consisting of the basic form plus the suffix *-la* or *-lê*.[3] Compare Ziervogel et al. (1969: 85), Ziervogel and Mokgokong (1975: 104-105, *Introduction*), Louwrens (1994: 49) and Poulos and Louwrens (1994: 87) in this regard. For the purposes of the current discussion, it has been decided to split the third position into two, since it would seem that both variants could (theoretically) occur in every class. Compare the examples for classes 1/2 and 5/6 in Table 3.

Table 3: Basic DC positions I, II, III and IIIa for classes 1/2 and 5/6

Class	I	II	III	IIIa
1	šo	šoô	šola	šolê
2	šeba	šebaô	šebala	šebalê
5	šele	šeleô	šelela	šelelê
6	šea	šeaô	šeala	šealê

The different positions are characterised by highly specific semantic distinctions, with each suffix carrying a particular semantic nuance, which distinguishes it from the other suffixes. As has been mentioned, these demonstratives are used to pinpoint the position of some person(s) or object(s) in relation to the position of both speaker and addressee. Thus, the DC of position I would be used to refer to some referent(s) close to both speaker and addressee, who are in turn in close proximity to one another. The translational equivalent of these forms would therefore be 'here (s)he/it is, close to us', or in the case of the plural 'here they are, close to us'. The DC of position II would be used in a situation where the speaker and the addressee are relatively far apart, while the person(s) or object(s) referred to is/are nearer to the addressee, but not right next to him/her. It would thus be translated as 'there (s)he/it is, close to you' or 'there they are, close to you'. Should the addressee and the speaker be in very close proximity to one another and the person(s) or object(s) being referred to is/are far away from the interlocutors, the DC of position III or IIIa would be used, carrying the meaning of 'there (s)he/it is, over yonder' or 'there they are, over yonder'.

Four additional forms are also mentioned in the literature, but these are generally regarded as dialectal forms. The first two of these four additional forms, the occurrence of which specifically excludes the eastern dialects (Kotzé 1985: 85), consist of the basic form (position I) to which the suffix *-nô* is added. The suffix *-khwi* occurs as a further dialectal variant of *-nô*. These two forms are labelled Ia and Ib respectively in Table 4.

Table 4: Basic DC positions I, II, III and IIIa, plus dialectal variants Ia and Ib, for classes 1/2 and 5/6

Class	I	**Ia**	**Ib**	II	III	IIIa
1	šo	**šonô**	**šokhwi**	šoô	šola	šolê
2	šeba	**šebanô**	**šebakhwi**	šebaô	šebala	šebalê
5	šele	**šelenô**	**šelekhwi**	šeleô	šelela	šelelê
6	šea	**šeanô**	**šeakhwi**	šeaô	šeala	šealê

According to Kotzé (1985: 85) and Louwrens (1991: 106), the DCs with suffixes *-nô* and *-khwi* would be used in a situation where the interlocutors are at a distance from one another and where the person(s) or object(s) being referred to

is/are right next to the speaker. The meaning expressed by these forms is thus 'here (s)he/it is, right next to me' or 'here they are, right next to me'.

The other two additional forms are also regarded as dialectal, with the suffix *-uwê* found in the dialects spoken in the vicinity of Polokwane, and its variant *-wê* found only in the dialects of Sekhukhuniland, particularly the Sepedi dialect (Ziervogel and Mokgokong 1975: 104, *Introduction*; Kotzé 1985: 86). These are labelled IIa and IIb respectively in Table 5.

Table 5: Basic DC positions I, II, III and IIIa, dialectal variants Ia and Ib, plus dialectal variants IIa and IIb, for classes 1/2 and 5/6

Class	I	Ia	Ib	II	**IIa**	**IIb**	III	IIIa
1	šo	šonô	šokhwi	šoô	**šouwê**	**šowê**	šola	šolê
2	šeba	šebanô	šebakhwi	šebaô	**šebauwê**	**šebawê**	šebala	šebalê
5	šele	šelenô	šelekhwi	šeleô	**šeleuwê**	**šelewê**	šelela	šelelê
6	šea	šeanô	šeakhwi	šeaô	**šeauwê**	**šeawê**	šeala	šealê

In a situation where the speaker and addressee are quite far apart from one another, these demonstratives would be used to refer to an object that is very close or directly next to the addressee. It can therefore be translated as 'there (s)he/it is, right next to you' or 'there they are, right next to you'.

Not all grammar books list the full paradigm for all the classes of Sesotho sa Leboa, and scrutiny of those that do, reveals that a number of differences exist with regard to especially the forms to be distinguished for classes 15 to 18. Ziervogel et al. (1969: 86), Lombard et al. (1985: 166), Nokaneng and Louwrens (1988: 221) and Poulos and Louwrens (1994: 88) indicate that the basic DC for class 15 is *šefa*, whereas Ziervogel and Mokgokong (1975: 104, *Introduction*) list *šego*. With regard to the basic DC for classes 16 to 18, Ziervogel and Mokgokong (1975: 104, *Introduction*) again differ from other scholars in that they are the only ones who distinguish the form *šego* for class 17. All other grammars indicate that the demonstrative for these classes is *šefa*. A search through a 6.1-million-word Sesotho sa Leboa corpus, henceforth 'the corpus', brought no results for a DC *šego*, thus the information provided by Ziervogel and Mokgokong with regard to both class 15 and class 17 seems to be unverified. Interestingly though, a basic demonstrative form not listed in any of the standard Sesotho sa Leboa grammars was thrown up by a corpus search. Indeed, seven instances were found of a demonstrative **šemo**, in all probability belonging to class 18, with a meaning similar to that of *šefa*, i.e. 'here (s)he/it is, here they are, here I am, etc.'

Another DC identified during a corpus search is **šetše**, containing the raised vowel [ẹ], i.e. [ʃẹtʃi]. This seems to be a variant of *šidi*, which is the so-called standard DC for classes 8 and 10. The two examples of *šetše* that were found in the corpus both have nouns in class 10 as antecedents, thus making it difficult to ascertain whether *šetše* could also have a noun from class 8 as antecedent. According to L.J. Louwrens (*personal communication*, 26 May 2004), this

particular variant is found as *sitši* in the Tlokwa dialect as a DC for both classes 8 and 10. Mother-tongue speakers also agree that this form may be used together with an antecedent from class 8. They furthermore indicate that the use of this form is widespread in the spoken language but, due to its non-standard status, is replaced in the written language with the standardised form.

In the last instance, it was noticeable that the variant **šedi** of the DC *šidi* for classes 8 and 10 is more often used than the so-called standard form. All grammar books list only *šidi* and both Ziervogel et al. (1969: 86) and Lombard et al. (1985: 166) clearly state that in the case of classes 8 and 10, assimilation between the vowel [e] of the root and the [i] of the concordial morpheme takes place, resulting in the assimilated form *šidi*. Despite the standard status of this form, only 36 concordance lines were found in the corpus in which *šidi* appears. Its distribution is furthermore relatively limited in that it is found in only 7 different sources. The 'non-standard' form *šedi*, however, appears in 209 concordance lines, spread across 56 different sources.

The full paradigm for the DC in Sesotho sa Leboa is therefore as shown in Table 6.

Table 6: Full paradigm for the DC in Sesotho sa Leboa (using the standard orthography) [**bold** = frequency of at least three in a 6.1-million-word corpus; *italics* = frequency of two and one in the corpus; (...) = only occurs in grammars]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1	**šo**	**šono**	**šokhwi**	**šoo**	(šouwe)	**šowe**	**šola**	**šole**
2	**šeba**	*šebano*	(šebakhwi)	**šebao**	(šebauwe)	*šebawe*	**šebala**	**šebale**
3	**šo**	(šono)	*šokhwi*	**šoo**	(šouwe)	*šowe*	**šola**	**šole**
4	**še**	*šeno*	(šekhwi)	**šeo**	(šeuwe)	(šewe)	(šela)	(šele)
5	**šele**	(šeleno)	(šelekhwi)	**šeleo**	(šeleuwe)	(šelewe)	(šelela)	*šelele*
6	**šea**	(šeano)	(šeakhwi)	**šeao**	(šeauwe)	(šeawe)	(šeala)	**šeale**
7	**sese**	(seseno)	(sesekhwi)	**seseo**	(seseuwe)	(sesewe)	*sesela*	(sesele)
8	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	*šedile*
8'	**šidi**	(šidino)	(šidikhwi)	(šidio)	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
8"	(šetše)	(šetšeno)	(šetšekhwi)	(šetšeo)	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
9	**še**	(šeno)	**šekhwi**	**šeo**	(šeuwe)	**šewe**	**šela**	**šele**
10	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	**šedile**
10'	**šidi**	(šidino)	(šidikhwi)	**šidio**	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
10"	*šetše*	(šetšeno)	(šetšekhwi)	*šetšeo*	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
14	**šebo**	(šebono)	(šebokhwi)	(šeboo)	(šebouwe)	*šebowe*	(šebola)	(šebole)
15	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
16	**šefa**	(šefano)	*šefakhwi*	**šefao**	(šefauwe)	(šefawe)	(šefala)	(šefale)
17	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
18	**šemo**	(šemono)	(šemokhwi)	(šemoo)	(šemouwe)	(šemowe)	(šemola)	(šemole)

Even though all forms enumerated in Table 6 are theoretically possible in Sesotho sa Leboa it is important to recall that, given that some positions are only found in particular dialects, one will not normally find all of them used by any single speaker. Actually, when comparing each DC against a 6.1-million-word corpus of Sesotho sa Leboa, only those forms marked in bold and italics do occur, with the bold items having a frequency of three or more, and the italicised ones a frequency of two or one only. Bracketed items have a zero frequency. Note further that each DC may be prefixed by the non-standard *a-* in the spoken language, effectively doubling the (theoretical) size of Table 6. In the corpus, however, not a single example was found of these *a-* forms. Given that dictionaries are mainly based on the standardised language, and that this is also the orthography represented in the corpus (since the corpus contains mainly written material), it is defendable not to include any of these *a-* forms in Sesotho sa Leboa dictionaries. If only a selection of members of the DC paradigm is to be treated in a dictionary, it further seems logical to focus on the truly frequent ones (i.e. the 42 bold items in Table 6).

4.   Treatment of the DC in the four existing desktop dictionaries for Sesotho sa Leboa

An investigation into the treatment of the DC in the four existing desktop dictionaries for Sesotho sa Leboa brings to light that these forms are dealt with in an inconsistent and sometimes even idiosyncratic manner. The four dictionaries, three of them bilingual and one trilingual, are (a) the third edition of the *Pukuntšu woordeboek* (Kriel 1983[3]), (b) the revision by Van Wyk of the *Pukuntšu woordeboek* (Kriel et al. 1989[4]) — being the latest edition of this dictionary, (c) the last edition of *The New English–Northern Sotho Dictionary* (Kriel 1976[4]), and (d) the trilingual *Comprehensive Northern Sotho Dictionary* (Ziervogel and Mokgokong 1975) — which saw only one edition.

4.1   Treatment of the DC in Kriel's (1983[3]) *Pukuntšu woordeboek, Noord-Sotho–Afrikaans*

One of the first problems one encounters when investigating the *Pukuntšu woordeboek* is the fact that indication of class membership is inconsistent and erratic, and the user is often left to his/her own devices to ascertain to which class a particular DC belongs, the only clue being provided by the translation equivalent (TE). The correct interpretation of the information contained in the TE furthermore often presupposes a thorough grammatical background. A case in point is the treatment of *šo*, which is the DC position I for both classes 1 and 3. The TE provided is *hier is hy/sy* 'here he/she is' which in all likelihood will be interpreted as referring to class 1, since class 1 is the one class containing nothing but nouns referring to humans. This implies that the DC position I of class 3 is not treated. In the case of *še*, which can be either class 4 or class 9, no indication of class membership is given. The TE *hier is dit* 'here it is', however, sug-

gests class 9, since it refers to the singular, class 4 being a plural class. Again, the implication is that the DC position I of class 4 is not treated in the dictionary. With regard to classes 8 and 10, the form *šidi* is listed, but with a label *ou spelling* 'old spelling' and with a cross-reference to *šedi*. This is contrary to the information uniformly provided by the Sesotho sa Leboa grammars that *šidi* is the standard form. Furthermore, class membership is incorrectly indicated as class 7 instead of class 8, while the DC of class 10 remains untreated. For the locative classes (16 to 18) only the form *šefa* is listed, with no class membership indicated and the TE given as *hier is dit* 'here it is'. For position Ia only one DC is treated, i.e. *šono*, which can belong to either class 1 or 3, the TE *hier is hy/sy* 'here he/she is' again suggesting class 1. None of the position Ib forms is found in the dictionary.

For position II the DCs of classes 5 (*šeleo*), 6 (*šeao*), 7 (*seseo*), 9 (*šeo*) and 16 (*šefao*) are treated. Under the entry *šeo* (class 9), mention is made of *šewe*, the DC position IIb of class 9. This is the only (implicit) cross-reference found to any of the DCs of position IIb in this particular dictionary; *šewe* itself, however, just as all other DCs from IIb, has not been entered. With regard to *šoo* the TE again suggests class 1 only, leaving the DC of class 3 untreated. It is further noticeable that although no DC position I *šego* is listed, the position II form *šegoo* is treated, albeit without any indication of class membership. No entries are found for any of the position IIa DCs. This is to a certain extent understandable, since the only grammar in which these variants are enumerated, is that of Nokaneng and Louwrens (1988: 221). These forms might have been regarded as (non-standard) dialectal forms and might therefore not have been included in the dictionary.

The DC position III class 1 listed is *šolaa*, a spelling not attested by any of the other sources.[4] Again, class membership is vested in the TE *daar is hy/sy* 'there he/she is'. The only other form of position III that is treated is *šegola*, but yet again without class indication, and this even though the basic form *šego* was not treated. For position IIIa, finally, only the DCs for classes 1 and 2 are listed, as well as *šefale* for the locative classes. A summary of these findings is presented in Table 7.

Table 7: Treatment of the DC in Kriel's (1983[3]) *Pukuntšu woordeboek, Noord-Sotho–Afrikaans* [dark shade = lemmatised; light shade = wrongly spelt]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1	**šo**	**šono**	**šokhwi**	**šoo**	(šouwe)	**šowe**	**šola**	**šole**
2	**šeba**	*šebano*	(šebakhwi)	**šebao**	(šebauwe)	*šebawe*	**šebala**	**šebale**
3	šo	(šono)	*šokhwi*	šoo	(šouwe)	*šowe*	šola	šole
4	še	*šeno*	(šekhwi)	šeo	(šeuwe)	(šewe)	(šela)	(šele)
5	**šele**	(šeleno)	(šelekhwi)	**šeleo**	(šeleuwe)	(šelewe)	(šelela)	*šelele*
6	**šea**	(šeano)	(šeakhwi)	**šeao**	(šeauwe)	(šeawe)	(šeala)	**šeale**
7	**sese**	(seseno)	(sesekhwi)	**seseo**	(seseuwe)	(sesewe)	*sesela*	(sesele)
8	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	*šedile*

8'	**šidi**	(šidino)	(šidikhwi)	(šidio)	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
8"	(šetše)	(šetšeno)	(šetšekhwi)	(šetšeo)	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
9	**še**	(šeno)	**šekhwi**	**šeo**	(šeuwe)	**šewe**	**šela**	**šele**
10	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	**šedile**
10'	**šidi**	(šidino)	(šidikhwi)	**šidio**	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
10"	*šetše*	(šetšeno)	(šetšekhwi)	*šetšeo*	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
14	**šebo**	(šebono)	(šebokhwi)	(šeboo)	(šebouwe)	*šebowe*	(šebola)	(šebole)
15	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
16	**šefa**	(šefano)	*šefakhwi*	**šefao**	(šefauwe)	(šefawe)	(šefala)	(šefale)
17	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
18	**šemo**	(šemono)	(šemokhwi)	(šemoo)	(šemouwe)	(šemowe)	(šemola)	(šemole)

From Table 7 it is clear that the DCs were lemmatised in a rather haphazard way in this dictionary. If corpus statistics are used as a guideline, then 42 forms (the bold ones) should be entered. In this dictionary, only 25 were lemmatised, including the wrongly spelt form. The overlap, however, is only 20 out of 42, or thus 48%. Conversely, 20 out of 25 lemmatised forms, or 80%, means that Kriel did a rather remarkable job on intuition alone.

4.2   Treatment of the DC in Kriel, Van Wyk and Makopo's (1989⁴) *Pukuntšu woordeboek, Noord-Sotho–Afrikaans*

This revision of the *Pukuntšu woordeboek* does represent some improvement in the lexicographical treatment of DCs. Class membership, for instance, is explicitly stated, thus easing the burden placed on the target user. Unfortunately, in a number of cases indication of class membership is incorrect. For position I, the DCs of classes 1 to 14 are treated consistently; however, the form *še* is indicated as belonging to classes 3 and 9, which is incorrect as it should be 4 and 9. Also, they claim that the DCs *šea* and *šeao* belong to class 8, whereas these forms are in fact the position I and II DCs for class 6. For classes 8 and 10, both *šedi* and *šidi* are given, but not *šetše*. No DC for class 15 is listed, and for the locative classes only *šefa* is treated. As far as the treatment of DCs of position Ia is concerned, only *šono* belonging to classes 1 and 3 is included. No DCs of position Ib have been lemmatised. With regard to DCs of position II, those of classes 2, 8, 10 and 14 have not been entered and, as was pointed out previously, *šeao* is incorrectly labelled with regard to class membership. No DC for class 15 is recorded and for the locative classes *šefao* and *šegoo* are treated. As was the case for the third edition of this dictionary, it is thus noticeable that the DC position II of class 17 is treated, but not the corresponding form for position I, i.e. *šego*. No DCs of positions IIa and IIb were found in the central lemma-sign list of this dictionary. Of all the position III DCs, only the form for class 17 is treated, i.e. *šegola*. With regard to position IIIa DCs, only those forms belonging to classes 1, 2, 3 and 16 have been entered into the dictionary. Compare Table 8 for a summary.

Table 8: Treatment of the DC in Kriel, Van Wyk and Makopo's (1989⁴) *Pu-kuntšu woordeboek* [dark shade = lemmatised; light shade = class wrongly assigned]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1	šo	šono	šokhwi	šoo	(šouwe)	šowe	šola	šole
2	šeba	*šebano*	(šebakhwi)	šebao	(šebauwe)	*šebawe*	šebala	šebale
3	šo	(šono)	*šokhwi*	šoo	(šouwe)	*šowe*	šola	šole
4	še	*šeno*	(šekhwi)	šeo	(šeuwe)	(šewe)	(šela)	(šele)
5	šele	(šeleno)	(šelekhwi)	šeleo	(šeleuwe)	(šelewe)	(šelela)	*šelele*
6	šea	(šeano)	(šeakhwi)	šeao	(šeauwe)	(šeawe)	(šeala)	šeale
7	sese	(seseno)	(sesekhwi)	seseo	(seseuwe)	(sesewe)	*sesela*	(sesele)
8	šedi	(šedino)	(šedikhwi)	šedio	(šediuwe)	(šediwe)	(šedila)	*šedile*
8'	šidi	(šidino)	(šidikhwi)	(šidio)	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
8"	(šetše)	(šetšeno)	(šetšekhwi)	(šetšeo)	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
9	še	(šeno)	šekhwi	šeo	(šeuwe)	šewe	šela	šele
10	šedi	(šedino)	(šedikhwi)	šedio	(šediuwe)	(šediwe)	(šedila)	šedile
10'	šidi	(šidino)	(šidikhwi)	šidio	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
10"	*šetše*	(šetšeno)	(šetšekhwi)	*šetšeo*	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
14	šebo	(šebono)	(šebokhwi)	(šeboo)	(šebouwe)	*šebowe*	(šebola)	(šebole)
15	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
16	šefa	(šefano)	*šefakhwi*	šefao	(šefauwe)	(šefawe)	(šefala)	(šefale)
17	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
18	šemo	(šemono)	(šemokhwi)	(šemoo)	(šemouwe)	(šemowe)	(šemola)	(šemole)

From Table 8 it is clear that as far as the lemmatisation proper is concerned, this fourth edition does not really improve much on the third edition. Including the three wrongly assigned forms, only 30 forms were lemmatised. The overlap with the frequent forms is 26 out of 42, or thus 54%. With 26 out of the 30 lemmatised forms also being frequent, the intuition score is however as high as 87%.

4.3   Treatment of the DC in Kriel's (1976⁴) *The New English–Northern Sotho Dictionary, Northern Sotho–English*

Whereas the two *Pukuntšu* editions above treated Afrikaans as second language pair, *The New English–Northern Sotho Dictionary* is currently the only bidirectional desktop dictionary with English and Sesotho sa Leboa as treated language pairs. Unfortunately, users of this dictionary are presented with an even more erratic treatment of the DCs. Table 9 summarises the lemmatisation status.

Table 9: Treatment of the DC in Kriel's (1976[4]) *The New English–Northern Sotho Dictionary, Northern Sotho–English* [dark shade = lemmatised; light shade = wrongly spelt]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1	šo	šono	šokhwi	šoo	(šouwe)	šowe	šola	šole
2	šeba	*šebano*	(šebakhwi)	šebao	(šebauwe)	*šebawe*	šebala	šebale
3	šo	(šono)	*šokhwi*	šoo	(šouwe)	*šowe*	šola	šole
4	še	*šeno*	(šekhwi)	šeo	(šeuwe)	(šewe)	(šela)	(šele)
5	šele	(šeleno)	(šelekhwi)	šeleo	(šeleuwe)	(šelewe)	(šelela)	*šelele*
6	šea	(šeano)	(šeakhwi)	šeao	(šeauwe)	(šeawe)	(šeala)	šeale
7	sese	(seseno)	(sesekhwi)	seseo	(seseuwe)	(sesewe)	*sesela*	(sesele)
8	šedi	(šedino)	(šedikhwi)	šedio	(šediuwe)	(šediwe)	(šedila)	*šedile*
8'	šidi	(šidino)	(šidikhwi)	(šidio)	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
8"	(šetše)	(šetšeno)	(šetšekhwi)	(šetšeo)	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
9	še	(šeno)	šekhwi	šeo	(šeuwe)	šewe	šela	šele
10	šedi	(šedino)	(šedikhwi)	šedio	(šediuwe)	(šediwe)	(šedila)	šedile
10'	šidi	(šidino)	(šidikhwi)	šidio	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
10"	*šetše*	(šetšeno)	(šetšekhwi)	*šetšeo*	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
14	šebo	(šebono)	(šebokhwi)	(šeboo)	(šebouwe)	*šebowe*	(šebola)	(šebole)
15	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
16	šefa	(šefano)	*šefakhwi*	šefao	(šefauwe)	(šefawe)	(šefala)	(šefale)
17	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
18	šemo	(šemono)	(šemokhwi)	(šemoo)	(šemouwe)	(šemowe)	(šemola)	(šemole)

Including two wrongly spelt forms, 19 DCs were entered in this dictionary, which, with an overlap of 15 items with the frequent forms, gives a coverage of 15 out of 42, or only 36%. The intuition score is still rather high, however, at 79% (15 out of 19). As was the case for the third edition of the *Pukuntšu*, there is virtually no class information, but more importantly, the labelling and data provided are anything but consistent. Apart from the fact that the POS of all forms is given as 'dem.', which means that these demonstrative *copulatives* are not distinguished from the real demonstratives, some are labelled 'pron.' and in addition even 'adj.' Other forms, although variants, are nonetheless treated very differently, e.g. '**še′di′le,** dem., pl., there they are, yonder.' versus '**šidi′le-e,** dem., there they are.' Or as a last example: '**šo,** dem., here he is, here she is.' versus '**šono,** dem., here he (she) is.'

4.4   Treatment of the DC in Ziervogel and Mokgokong's (1975) *Comprehensive Northern Sotho Dictionary, Northern Sotho–Afrikaans/English*

This trilingual dictionary is the only one of the desktop dictionaries that treats the prefixal element of the DCs. Both variants, i.e. *še-* and *ši-* are found in the

lemma-sign list: *še-* is fully treated and defined as a 'prefixal element in formation of cop. dem.', whereas *ši-* is defined as the 'assimilated form of **še-**', thus implicitly cross-referring the user to the canonical form of the prefix, i.e. *še-*. In this dictionary, the DCs of position I for classes 1 to 14 are consistently treated, with indication of class adherence. Only the dialectal variant *šetše* for classes 8 and 10 is not listed. Although both other forms of the DCs of classes 8 and 10 are found in the lemma-sign list, the 'non-standard' *šedi* is not treated, but cross-referred to the 'standard' form *šidi*, which is then fully treated. No DC is distinguished for class 15. For the locative classes, only *šefa* is listed. The DCs of positions Ia and Ib do not appear in this dictionary, and neither do those of positions IIa and IIb. The same DCs which are treated under position I, are also treated for position II, with the addition of *šegoo*, which is distinguished for the locative classes. It is labelled as a dialectal form, but it is (again) not clear why the position I form (*šego*) is not treated. The form *šedio* (classes 8 and 10) is listed, but cross-referred to *šidio*, which is then treated. With regard to position III, all DCs of classes 1 to 14 are treated, with the exception of those of classes 2 and 6. For the locative classes, only *šegola* is listed, again labelled as a dialectal form. For position IIIa, only classes 1 and 3 (*šole*), class 2 (*šebale*), class 6 (*šeale*) and class 16 (*šefale*) have been lemmatised. Compare Table 10 for an overview.

Table 10: Treatment of the DC in Ziervogel and Mokgokong's (1975) *Comprehensive Northern Sotho Dictionary, Northern Sotho–Afrikaans/English* [dark shade = lemmatised; light shade = wrongly spelt]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1	**šo**	**šono**	**šokhwi**	**šoo**	(šouwe)	**šowe**	**šola**	**šole**
2	**šeba**	*šebano*	(šebakhwi)	**šebao**	(šebauwe)	*šebawe*	**šebala**	**šebale**
3	**šo**	(šono)	*šokhwi*	**šoo**	(šouwe)	*šowe*	**šola**	**šole**
4	**še**	*šeno*	(šekhwi)	**šeo**	(šeuwe)	(šewe)	(šela)	(šele)
5	**šele**	(šeleno)	(šelekhwi)	**šeleo**	(šeleuwe)	(šelewe)	(šelela)	*šelele*
6	**šea**	(šeano)	(šeakhwi)	**šeao**	(šeauwe)	(šeawe)	(šeala)	**šeale**
7	**sese**	(seseno)	(sesekhwi)	**seseo**	(seseuwe)	(sesewe)	*sesela*	(sesele)
8	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	*šedile*
8'	**šidi**	(šidino)	(šidikhwi)	(šidio)	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
8"	(šetše)	(šetšeno)	(šetšekhwi)	(šetšeo)	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
9	**še**	(šeno)	**šekhwi**	**šeo**	(šeuwe)	**šewe**	**šela**	**šele**
10	**šedi**	(šedino)	(šedikhwi)	**šedio**	(šediuwe)	(šediwe)	(šedila)	**šedile**
10'	**šidi**	(šidino)	(šidikhwi)	**šidio**	(šidiuwe)	(šidiwe)	(šidila)	*šidile*
10"	*šetše*	(šetšeno)	(šetšekhwi)	*šetšeo*	(šetšeuwe)	(šetšewe)	(šetšela)	(šetšele)
14	**šebo**	(šebono)	(šebokhwi)	(šeboo)	(šebouwe)	*šebowe*	(šebola)	(šebole)
15	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
16	**šefa**	(šefano)	*šefakhwi*	**šefao**	(šefauwe)	(šefawe)	(šefala)	(šefale)
17	(šego)	(šegono)	(šegokhwi)	(šegoo)	(šegouwe)	(šegowe)	(šegola)	(šegole)
18	**šemo**	(šemono)	(šemokhwi)	(šemoo)	(šemouwe)	(šemowe)	(šemola)	(šemole)

As can be seen from Table 10, a total of 46 DCs (which includes two misspelled ones) were lemmatised in this dictionary, 33 of which belong to the frequent DCs. With 33 out of 42 frequent items, the coverage is rather satisfactory at 79%; the intuition ratio stands at slightly less with 72% (33 out of 46).

Of all the desktop dictionaries, this reference work is the only one that has a relatively extended front matter, including a mini-grammar of Sesotho sa Leboa. In this section, the DC is discussed, but the full paradigm is only provided for position I. For some other positions, information is provided on their morphological composition only. Even though the DCs of positions Ia (-*nô*) and Ib (-*khwi*) do not appear in the central lemma-sign list, reference is made in the front matter to these forms, but not to the DCs of positions IIa (-*uwê*) and IIb (-*wê*). Unfortunately, there is no system of cross-referencing from the central lemma-sign list to the information provided in the front matter. This is particularly relevant for the DCs of positions Ia and Ib, which are not treated in the central text, but which are discussed in the front matter. Even if these forms are not treated in the dictionary proper, they could simply have been entered with a cross-reference to the information provided in the front matter.

5.   Towards a sound lexicographic treatment of the DC: Problem analysis and options

From the above review of the treatment of the DC in the current desktop dictionaries for Sesotho sa Leboa, one actually realises that there are three main issues at stake. Firstly there is the problem of consistency, secondly there is the problem of which data categories to include and how, and thirdly there is the problem of which DCs to treat and where.

5.1   The problem of consistency

As far as consistency is concerned, a quick glance at an overview of the actual treatment of the DC in the four desktop dictionaries, reproduced *verbatim* in Addendum 1, makes the current erratic approach very evident. Even the treatment that clearly received the most detailed attention, viz. the one found in Ziervogel and Mokgokong's (1975) *Comprehensive Northern Sotho Dictionary*, is still full of inconsistencies. The very label for the POS, to begin with, is found 22 times as 'dem. kop.' versus 8 times as 'kop. dem.' (plus once as 'dem. kop ') in Afrikaans, and 22 times as 'dem. cop.' versus 8 times as 'cop. dem.' (plus once as 'dem. cop ') in English. On the semantic level one finds 'there are they over there' for *šebale* where it should be 'there they are over there', or 'doer is dit' at *sesela* where it should be 'dáár is dit'. On the cross-reference level, the abbreviation 'v.' is for instance used at *še-*, but 'cf.' at *ši-*. Lastly, as illustrations of general layout infelicities, in articles such as *sese* or *seseo* the Afrikaans and English sections are not correctly separated from one another, while in *šeleo* the brack-

ets do not come in pairs. The inconsistencies in the other three dictionaries are more abundant, as can be seen from Addendum 1.

With modern dictionary writing systems such as *TshwaneLex*, most of these problems are taken care of by the software. POSs, for example, are chosen from a list of options and ought not to be typed in anywhere by the lexicographers, cross-references are also chosen from a finite list of options and can only be inserted when the reference address physically exists, the data distribution structure is mainly an output aspect about which lexicographers should not have to worry during compilation, while basic typographical issues such as the pairing of brackets belong to standard error checks built into the software.

With specific reference to African-language paradigms, experience has shown that the only truly sensible way to treat them is to work through all forms consistently, going down the list of all the classes in a principled way. It is impossible to produce a coherent text if one member of a paradigm is treated today, and another a week later — as they happen to cross the compiler's way. In other words, there must not only be a strict lemmatisation approach to paradigms, paradigms must also be treated 'in block'.

5.2   The problem of which data categories to include and how

Apart from the inconsistent and erratic treatment of DCs in the dictionaries under discussion, an additional shortcoming is the insufficient attention that is paid to the semantic implication carried by each of the DCs. In all of these dictionaries, the meaning of the DC is defined in terms of the relative distance between two reference points only, i.e. the speaker and the person(s) or object(s) being referred to. As was indicated above, the crucial aspect in defining the exact semantic implication expressed by the DCs, is the position of the object in relation to the positions of both speaker and addressee. TEs and definitions — for bilingual and monolingual dictionaries respectively — of these forms should therefore be formulated in such a manner that the semantic nuances that distinguish the different DCs from one another are clearly indicated. In practical terms this actually means that there are *two* levels of data that need to be given due attention in the comment on semantics (CS) of the DCs; these are the *semantic content* and the *spatial relation*, as summarised in Table 11.

When dealing with a bilingual dictionary, the 'semantic content' column in Table 11 actually provides a rather precise set of TEs, but the full picture is only obtained when the information presented in the 'spatial relation' column is also available to the person consulting the reference work. For a monolingual dictionary, the TE-like column is close to irrelevant, precisely as a result of the structure of African languages, while it is the information from the 'spatial relation' column that should receive prime attention. While both a translation and an explanatory dictionary should thus try to incorporate as much as possi-

ble from the two main columns of Table 11, translation dictionaries will rather focus on the left column, while explanatory dictionaries will rather focus on the right column.

Table 11: Semantic content and spatial relation for each of the different positions of the DCs [with sg = singular, pl = plural, S = speaker, A = addressee, * = person(s) and/or object(s) being referred, ↔ and ↕ = relative distances]

DC	Semantic content		Spatial relation
I	sg here (s)he/it is, close to us pl here they are, close to us	SA *	used in a situation where speaker and addressee are in close proximity to one another, with the person(s) and/or object(s) being referred to also close to the interlocutors
Ia	sg here (s)he/it is, right next to me pl here they are, right next to me	S↔A *	used in a situation where speaker and addressee are at a distance from one another, while the person(s) and/or object(s) being referred to is/are right next to the speaker
Ib	sg here (s)he/it is, right next to me pl here they are, right next to me		
II	sg there (s)he/it is, close to you pl there they are, close to you	S↔A *	used in a situation where speaker and addressee are relatively far apart, while the person(s) and/or object(s) being referred to is/are nearer to the addressee but not right next to him/ her
IIa	sg there (s)he/it is, right next to you pl there they are, right next to you	S↔A *	used in a situation where speaker and addressee are quite far apart from one another, while the person(s) and/or object(s) being referred to is/are very close or directly next to the addressee
IIb	sg there (s)he/it is, right next to you pl there they are, right next to you		
III	sg there (s)he/it is, over yonder pl there they are, over yonder	SA ↕ *	used in a situation where speaker and addressee are in very close proximity to one another, while the person(s) and/or object(s) being referred to is/are far away from the interlocutors
IIIa	sg there (s)he/it is, over yonder pl there they are, over yonder		

Besides the comment on semantics (CS), one will also have to decide what to include under the comment on form (CF), selecting from aspects such as POS label, class affiliation, pronunciation, tone marking, etc.

5.3   The problem of which DCs to treat and where

The most important issue of all, however, remains how to decide which members of a paradigm like the one of the DCs to treat, and where to treat those members. In the case of the lemmatisation of a one-dimensional paradigm such as, say, the possessive concords, it seems rather trivial to just make sure all members are included in a dictionary's macrostructure. With a two-dimensional paradigm such as the one of the DCs, the choice is less obvious. One could argue that with two linked tables such as for example an adaptation of Tables 6 and 11, any dictionary user should have enough information regarding the DCs in Sesotho sa Leboa. Such tables could be presented in the dictionary's front or back matter, have the advantage that they display *all* forms, and there would be no need to lemmatise any DC. That this approach is actually not successful is shown by the above discussion of the absence of 'its' in Rycroft's dictionary: one either has to know the meaning and/or the orthographic form prior to looking up such members of a paradigm.

When dealing with complex paradigms, it is thus clear that one will have to lemmatise (a selection of) the members in the macrostructure. Apart from using intuition only, which is not advised, one could follow a principled approach in selecting certain forms. Principled approaches could for instance include: all members with certain characteristics only, all irregular forms, all forms known to be problematic or confusing, etc. Another approach could be to use frequency data in the selection process, and to only lemmatise those that are more frequent than a certain threshold. A further refinement could be to do the latter, and to also 'mention' all attested forms in the macrostructure, but without full treatment of these. If space is of no concern, one may even include and fully treat all forms in the central text. For each of these lemmatisation approaches one may of course in addition also incorporate a set of tables in the front or back matter, preferably with cross-references from the central text to these tables. A last feasible option could be to simply include all forms in the macrostructure, but to only provide a cross-reference to the tables instead of a full treatment of each. All these options are summarised in Table 12.

Table 12: Modern lemmatisation options for the members of a (complex) paradigm

Option		Lemmas	Tables
• only in table-form (*all* members of the paradigm by default)			✓
• only as lemma signs (*which ones?*:)		✓	
o principled selection of certain members	(+ tables)	✓	(✓)
o frequent members only	(+ tables)	✓	(✓)
o frequent members only + mention of attested ones	(+ tables)	✓	(✓)
o all (theoretically possible) members	(+ tables)	✓	(✓)
• in table-form + all members with cross-references to the tables		✓	✓

6.   A practical implementation: The treatment of the DC in the *Pukuntšu-tlhaloši ya Sesotho sa Leboa* (PyaSsaL)

In order to best see how the different options translate into a real dictionary project, the treatment of the DC in the PanSALB-sponsored PyaSsaL, i.e. the monolingual *Pukuntšutlhaloši ya Sesotho sa Leboa* 'Explanatory Sesotho sa Leboa Dictionary', will now be analysed.

From its inception the compilation of PyaSsaL has been fully corpus-based (cf. De Schryver and Lepota 2001: 3). More recently the results from fieldwork for especially cultural lexical items have supplemented the corpus data. The current policy is that each lexical item that occurs at least three times in a 6.1-million-word corpus be considered for inclusion.[5] Each defined lemma sign is further illustrated with at least one example *culled from the corpus*. The latter approach, which is considered of paramount importance as it ensures that one is dealing with *authentic* usage of *attested* forms, has actually prescribed the way in which to treat the DC in PyaSsaL. Each DC with a frequency of at least three received full treatment in the central text. In addition, it was decided to 'list' all members of the DC paradigm with lower frequencies as well and to cross-refer these to DC Tables, but *not* to lemmatise any forms not attested in the corpus. This thus means that one is effectively dealing with a three-tier structure in PyaSsaL, i.e. (a) frequent DCs are provided with both a CF and CS, (b) infrequent DCs are provided with a CF (but no CS) and a reference to DC Tables, and (c) non-attested DCs are tabulated outside the central lemma-sign list in DC Tables only. The last group includes those forms that are only mentioned in grammar books, with no examples in the other corpus sources.

Given that automatic POS taggers have as yet not been developed for Sesotho sa Leboa, both the meaning and the class affiliation of each DC 'form' had to be deduced from a meticulous scrutiny of concordance lines. That this is not a trivial process is shown by the detailed corpus data presented in Addendum 2. Firstly, note that some DC forms are *homonymous* with non-DC forms, such as for example the DC *šeba* which is homonymous with the verb *šeba* 'relish, eat as a titbit'. In this case the DC use is slightly more frequent, so this form carries the homonym number 1, while the verb is assigned homonym number 2. Even though some DCs are only attested in grammar books, their homonymous form might be relatively frequent, as in the case of *sesele* 'badger', which means that only this non-DC form is lemmatised. Secondly, note that corpus statistics also indicate in which order DCs that are morphologically similar (which is the case for classes 1 and 3, 4 and 9, and 8 and 10) ought to be ordered *within* an article. For classes 4 and 9, for example, the class 9 form is more frequent than the class 4 form for all positions except for position Ia. Thirdly, a combination of these first two observations also occurs, as for *šele*. This item occurs 817 times in all in the corpus, 719 times as an enumerative stem, 80 times as position I class 5 DC, 18 times as position IIIa class 9 DC, but not once as position IIIa class 4 DC. In a Sesotho sa Leboa–English dictionary, the mini-

mal treatment (i.e. without example sentences) could thus take the form of the articles shown in (2), where 'SEE <u>DC Tables</u>' is a cross-reference to a set of tables analogous to Tables 6 and 11 above, to be found in the front or back matter of the dictionary.[6]

(2)
še 1. *dem. cop. I cl. 9* here it is, close to us; **2.** *dem. cop. I cl. 4* here they are, close to us
šele[1] *enumerative stem* strange, foreign, different
šele[2] *dem. cop. I cl. 5* here it is, close to us
šele[3] (< **še**) **1.** *dem. cop. IIIa cl. 9* there it is, over yonder; **2.** *dem. cop. IIIa cl. 4* SEE <u>DC Tables</u>

Observe that in the case of the position IIIa class 4 DC, a CF is provided, even though there are no attestations of this form in the corpus. This is an example of a case where practical dictionary making overrides corpus data for the sake of consistency. It would indeed seem awkward to leave out the second sense of **šele[3]**, unlike leaving out fully unattested forms such as say *šelekhwi*, *šeawe* or *šemole* from the central dictionary text.

Frequency data thus enable one to decide in which order to present homonymous items, and also in which order to present the data within single articles. Such statistics further also enable one to cross-refer lesser frequent variants to more frequent ones, as can be seen from the cross-reference from the reference position in **šetše[4]** to the reference address **šedi[2]** in (3).

(3)
šedi[1] *n. cl. 9* care, attention
šedi[2] **1.** *dem. cop. I cl. 10* here they are, close to us; **2.** *dem. cop. I cl. 8* here they are, close to us
šetše[1] *aux.* already
šetše[2] *v.* **1.** remain (behind); **2.** follow (behind)
šetše[3] *v.* must pay attention; **..ga/sa/se..~** not pay attention
šetše[4] = šedi[2] 1. *dem. cop. I cl. 10* SEE <u>DC Tables</u>; **2.** *dem. cop. I cl. 8* SEE <u>DC Tables</u>

The treatment of the DC and its homonymous forms as shown in (2) and (3) is also how, mutatis mutandis, the DC is treated in PyaSsaL. If one carefully scrutinises these examples, one realises that one is actually dealing with a complex set of *multiple cross-references* that is partially driven by corpus data. One firstly sees from (2) that all higher-order positions (Ia, Ib, II, IIa, IIb, III and IIIa) may be linked to their base form (i.e. position I) through the use of the non-typographical structural marker '<'. Viewed from the DC Table angle, column-forms are thus cross-referenced. Secondly, row-forms may also be cross-referenced, such as in the case of the variant forms for which the non-typographical structural marker '=' is used, as can be seen in (3). The third type of cross-reference used to link some of the forms of the full DC paradigm are the links between the central text and the DC Tables in the front or back matter. Properly treating a complex African-language paradigm, in casu that of the DC in Sesotho sa Leboa, thus leads to an exercise in multiple cross-referencing.

7.   *PyaSsaL ka Inthanete*: A pioneer in the untrodden forest

In the previous section, it was pointed out that a corpus has been consulted since compilation of PyaSsaL began. Also present since its inception has been the theoretical framework of *Simultaneous Feedback* or 'SF', which can be understood as entailing a method in terms of which the release of several small-scale parallel dictionaries triggers off feedback that is instantly channelled back into the compilation process of a main dictionary. To date, three parallel dictionaries have been released in hardcopy format.

The electronic adaptation of SF is known as *Fuzzy SF*, and PyaSsaL indeed went electronic on 10 June 2004. This new online dictionary is freely available from http://africanlanguages.com/psl/ and is known as *PyaSsaL ka Inthanete* 'Online PyaSsaL' (Mojela et al. 2004). The online data show work in progress — a selection of around 7 500 articles out of the 10 000 currently in preparation — and the main purpose is to retrieve various types of feedback, both of the implicit type through a study of the dictionary-use log files, and of the explicit type through receiving comments. An online feedback form can be filled in to that effect. The Online PyaSsaL is a pioneer in that it is the very first monolingual African-language dictionary on the Internet for which also the interface and the metalanguage of all macro- and microstructural elements are presented in the African language (cf. De Schryver 2003a: 9). A screenshot of the start page of this online dictionary can be seen in Addendum 3.

The multifarious advantages of an electronic medium for dictionaries are well known (cf. De Schryver 2003), and this is no different when it comes to the treatment of the DC. It is important to realise that even though the Online PyaSsaL is presented on the Internet, the work remains primarily compiled with a printed dictionary in mind. It would thus be naive for example to decide to treat *all* DCs online, as this would make the hardcopy version unnecessarily bulky, apart from the fact that one cannot find a single corpus example for six out of each ten theoretically possible DCs. Nonetheless, even though one is effectively dealing with the same data, the electronic environment *does* provide an array of useful extra lexicographical devices. The focus will be on one of them: cross-referencing. In line with the concepts of SF and Fuzzy SF, some of the options discussed below are already being implemented, while others are only experimented with — the idea exactly being to attempt to tailor the approach to the users of the dictionary.

Whereas the various reference relations are, to save space, typically established by a set of symbols in printed dictionaries, such a symbol set can easily be replaced with a set of user-friendlier text segments in an electronic environment. In the dictionary compilation software, one only needs to create a 'parallel set of cross-references', without touching any of the dictionary data. As such, the reference marker '<' can for example be replaced throughout with *Go tšwa go* 'Derived from'. This thus means that any *textual* information of any length can be placed within the reference marker itself, greatly enhancing the

readability. On a second level one can point out that whereas paging *to* a reference address in a paper dictionary is cumbersome and time-consuming, that process is sped up with what is known as *hyperlinking* in a computational environment. A further user-friendly enhancement in an electronic medium is to display all related cross-references of an item. Here cross-references *from*, but also cross-references *to* an article can simply be culled out of the database automatically and be presented together with the looked-up item. If one looks up *šidi* in the Online PyaSsaL, for example, the article *šedi* (which is referred to from *šidi*) as well as the articles *šidio* and *šidile* (which contain cross-references to *šidi*) are shown on the same output page, enabling the dictionary user to quickly obtain a good overview of the preferred variant and derived forms. A simpler case can be seen in Addendum 4, where a search for *šokhwi* also returns the article for the base form *šo*.

At least as powerful is the possibility to call up tables at any point, so clicking on BONA Lenaneo la mašalašupi-leba 'SEE Table of demonstrative copulatives' in the article of *šokhwi* (cf. Addendum 4), for example, will currently display a table similar to Table 6 above. At present this DC Table is generated as a static table within the same dictionary page, but plans include experimenting with a dynamically generated page, opening in a pop-up window and with the particular DC being looked up immediately highlighted. Such a cross-reference (hyperlink) will instantly allow the dictionary user to frame the DC at hand within the full paradigm of all DCs — an exploit that cannot be achieved in the paper dimension. A further option could be to make each item from the DC Table that is also lemmatised clickable, so that dictionary users could even more easily browse through the dictionary.

8.   In conclusion

In 2004, the year South Africa celebrates and looks back on 10 years of democracy, it seemed appropriate to also assess some of the achievements in African-language metalexicography. It was shown that great strides have been made during the past decade, in particular when it comes to Sesotho sa Leboa. New directions of research were then suggested for the future, and one of them, namely the 'paradigmatic lemmatisation' of closed-class words, was singled out for this article. As a case study within this field, the lexicographic treatment of the demonstrative copulative (DC) in Sesotho sa Leboa was discussed in depth. To that end, the DC as lemmatised in the current desktop dictionaries for Sesotho sa Leboa was first analysed, and following an identification of the problem areas, options for a sound treatment in both bilingual and monolingual dictionaries were then suggested. It was indicated that a sound approach to 'paradigmatic lemmatisation' should preferably (a) make use of professional dictionary software such as *TshwaneLex* to ensure consistency, (b) take both the semantic content and the spatial relation of each member of the paradigm into account, and (c) put corpus frequency data to good use when deciding on

which members to lemmatise and how to order homonyms and senses. It was also pointed out how the use of multiple cross-referencing (hyperlinking) can successfully link the alphabetically dispersed members of a paradigm. In addition, the benefit was emphasised of also making such 'links' with overview tables that list all theoretically possible forms.

An actual implementation of a modern treatment of the DC was then presented for PyaSsaL, i.e. the *Pukuntšutlhaloši ya Sesotho sa Leboa* 'Explanatory Sesotho sa Leboa Dictionary'. The compilation of PyaSsaL is both corpus-based and undertaken within the theoretical framework of *Simultaneous Feedback* (SF). In order to speed up the process of retrieving feedback, a selection of the PyaSsaL data is currently presented online as work in progress at http://africanlanguages.com/psl/, making *PyaSsaL ka Inthanete* 'Online PyaSsaL' the only truly monolingual dictionary on the Internet for any African language at present. Various presentation options for the DC are currently experimented with online, including multiple ways of cross-referencing and hyperlinking to and from static as well as dynamically generated tables, with and without highlighting, etc. Given that the developed approaches are generic, and given that there are many more paradigms in Sesotho sa Leboa — as well as in all other African languages — the considerable multiplication factor of this study can hardly be underestimated.

Endnotes

1. Since this article is being submitted for publication in a South African journal, necessary sensitivity with regard to the term 'Bantu' languages is exercised in our choice rather to use the term *African* languages. Bear in mind, however, that the latter includes more than just the 'Bantu Language Family'.

2. Observe that tone is not normally marked in the siSwati orthography, which is why it is left out in the running text of this article.

3. In the standard Sesotho sa Leboa orthography the 'raised ê' and 'e' are collapsed to 'e', and likewise the 'raised ô' and 'o' are collapsed to 'o'. This practice is also followed in the running text of this article.

4. A long final vowel is however suggested for positions III and IIIa by Ziervogel and Mokgokong (1975: 104-105, *Introduction*), but only in their grammatical outline, not in their dictionary proper.

5. This 6.1-million-word corpus for Sesotho sa Leboa is an organic corpus that is continuously being expanded and revised. It is built by staff members in the Department of African Languages at the University of Pretoria, in cooperation with lexicographers from the Sesotho sa Leboa National Lexicography Unit (NLU) and the authors of this article.

6. Depending on the intended target user group, the type and amount of information provided in such tables, as well as the way in which this information is presented, will differ. The different members of the DC paradigm that are also treated in the central lemma-sign list could for example be presented in a different colour, typeface or even be highlighted, to differentiate these forms from those that are not attested in the corpus. It could even be considered in-

cluding the actual corpus frequencies at each paradigm member, as has been done in Addendum 2. The minimal treatment, conversely, is as a simple enumeration of all forms in a two-dimensional table.

References

Bosch, Sonja E. and D.J. Prinsloo. 2002. 'Abbreviated Nouns' in African Languages: A Morphological, Semantic and Lexicographic Perspective. *South African Journal of African Languages* 22(1): 92-104.

De Schryver, Gilles-Maurice. 1999. Bantu Lexicography and the Concept of *Simultaneous Feedback*, Some Preliminary Observations on the Introduction of a New Methodology for the Compilation of Dictionaries with Special Reference to a Bilingual Learner's Dictionary *Cilubà–Dutch.* Unpublished MA thesis. Ghent: Ghent University.

De Schryver, Gilles-Maurice. 2002. Web for/as Corpus: A Perspective for the African Languages. *Nordic Journal of African Studies* 11(2): 266-282.

De Schryver, Gilles-Maurice. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

De Schryver, Gilles-Maurice. 2003a. Online Dictionaries on the Internet: An Overview for the African Languages. *Lexikos* 13: 1-20.

De Schryver, Gilles-Maurice. 2003b. Drawing up the Macrostructure of a Nguni Dictionary, with Special Reference to isiNdebele. *South African Journal of African Languages* 23.

De Schryver, Gilles-Maurice and B. Lepota. 2001. The Lexicographic Treatment of Days in Sepedi, or When Mother-Tongue Intuition Fails. *Lexikos* 11: 1-37.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000. Dictionary-Making Process with 'Simultaneous Feedback' from the Target Users to the Compilers. Heid, U., S. Evert, E. Lehmann and C. Rohrer (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, August 8th–12th, 2000*: 197-209. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Stuttgart University.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000a. The Concept of 'Simultaneous Feedback': Towards a New Methodology for Compiling Dictionaries. *Lexikos* 10: 1-31.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000b. The Compilation of Electronic Corpora, with Special Reference to the African Languages. *Southern African Linguistics and Applied Language Studies* 18(1-4): 89-106.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000c. Electronic Corpora as a Basis for the Compilation of African-language Dictionaries, Part 1: The *Macrostructure*. *South African Journal of African Languages* 20(4): 291-309.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000d. Electronic Corpora as a Basis for the Compilation of African-language Dictionaries, Part 2: The *Microstructure*. *South African Journal of African Languages* 20(4): 310-330.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2001. Corpus-based Activities versus Intuition-based Compilations by Lexicographers, the Sepedi Lemma-Sign List as a Case in Point. *Nordic Journal of African Studies* 10(3): 374-398.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2001a. Fuzzy SF: Towards the Ultimate Customised Dictionary. *Studies in Lexicography* 11(1): 97-111.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2001b. Towards a Sound Lemmatisation Strategy for the Bantu Verb through the Use of *Frequency-based Tail Slots* — with Special Reference to Cilubà, Sepedi and Kiswahili. Mdee, J.S. and H.J.M. Mwansoko (Eds.). 2001. *Makala ya kongamano la kimataifa Kiswahili 2000. Proceedings*: 216-242, 372. Dar es Salaam: TUKI, Chuo Kikuu cha Dar es Salaam.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2003. Compiling a Lemma-sign List for a Specific Target User Group: The Junior Dictionary as a Case in Point. *Dictionaries: Journal of The Dictionary Society of North America* 24: 28-58.

Gouws, Rufus H. and D.J. Prinsloo. 1997. Lemmatisation of Adjectives in Sepedi. *Lexikos* 7: 45-57.

Joffe, David, Gilles-Maurice de Schryver and D.J. Prinsloo. 2003. Introducing TshwaneLex — A New Computer Program for the Compilation of Dictionaries. De Schryver, G-M (Ed.). 2003. *TAMA 2003 South Africa: CONFERENCE PROCEEDINGS*: 97-104. Pretoria: (SF)[2] Press.

Joffe, David, Gilles-Maurice de Schryver and D.J. Prinsloo. 2003a. Computational Features of the Dictionary Application "TshwaneLex". *Southern African Linguistics and Applied Language Studies* 21(4) [Special issue on 'Human Language Technology in South Africa: Resources and Applications']: 239-250.

Kilgarriff, Adam. 1996. *BNC Database and Word Frequency Lists* [online]. Available: <http://www.itri.bton.ac.uk/~Adam.Kilgarriff/bnc-readme.html>.

Kotzé, Albert E. 1985. Herinterpretasie van die betekenis van demonstratiewe in Noord-Sotho. *South African Journal of African Languages* 5(3): 82-87.

Kriel, Theunis J. 1976[4]. *The New English–Northern Sotho Dictionary, English–Northern Sotho, Northern Sotho–English.* Johannesburg: Educum Publishers.

Kriel, Theunis J. 1983[3]. *Pukuntšu woordeboek, Noord-Sotho–Afrikaans, Afrikaans–Noord-Sotho*. Pretoria: J.L. van Schaik.

Kriel, Theunis J., Egidius B. van Wyk and Staupitz A. Makopo. 1989[4]. *Pukuntšu woordeboek, Noord-Sotho–Afrikaans, Afrikaans–Noord-Sotho*. Pretoria: J.L. van Schaik.

Lombard, Daniël P., Egidius B. van Wyk and Pothinus C. Mokgokong. 1985. *Inleiding tot die grammatika van Noord-Sotho*. Pretoria: J.L. van Schaik.

Louwrens, Louis J. 1991. *Aspects of Northern Sotho Grammar.* Pretoria: Via Afrika Ltd.

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

Mojela, M.V. (Editor-in-Chief), M.P. Mogodi, M.C. Mphahlele and M.R. Selokela (Compilers). 2004. *Pukuntšutlhaloši ya Sesotho sa Leboa ka Inthanete (Explanatory Sesotho sa Leboa Dictionary on the Internet)* [online]. Available: <http://africanlanguages.com/psl/>.

Nokaneng, Mogobo B. and Louis J. Louwrens. 1988. *Segagešo Mphato* 9. Pretoria: Via Afrika Ltd.

Nong, Salmina, Gilles-Maurice de Schryver and D.J. Prinsloo. 2002. Loan Words versus Indigenous Words in Northern Sotho — A Lexicographic Perspective. *Lexikos* 12: 1-20.

Poulos, George and Louis J. Louwrens. 1994. *A Linguistic Analysis of Northern Sotho.* Pretoria: Via Afrika Ltd.

Prinsloo, D.J. 1991. Towards Computer-assisted Word Frequency Studies in Northern Sotho. *South African Journal of African Languages* 11(2): 54-60.

Prinsloo, D.J. 1992. Lemmatization of Reflexives in Northern Sotho. *Lexikos* 2: 178-191.

Prinsloo, D.J. 1994. Lemmatization of Verbs in Northern Sotho. *South African Journal of African Languages* 14(2): 93-102.

Prinsloo, D.J. 2002. The Lemmatization of Copulatives in Northern Sotho. *Lexikos* 12: 21-43.

Prinsloo, D.J. 2003. The Lemmatisation of Adverbs in Northern Sotho. *Lexikos* 13: 21-37.

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 Gilles-Maurice de Schryver. 2001. Monitoring the Stability of a Growing Organic Corpus, with Special Reference to Sepedi and Xitsonga. *Dictionaries: Journal of The Dictionary Society of North America* 22: 85-129.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2002. Designing a Measurement Instrument for the Relative Length of Alphabetical Stretches in Dictionaries, with Special Reference to Afrikaans and English. Braasch, A. and C. Povlsen (Eds.). 2002. *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 13-17, 2002*: 483-494. Copenhagen: Center for Sprogteknologi, University of Copenhagen.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2002a. The Use of Slashes as a Lexicographic Device, with Special Reference to the African Languages. *South African Journal of African Languages* 22(1): 70-91.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2002b. Reversing an African-language Lexicon: The *Northern Sotho Terminology and Orthography No. 4* as a Case in Point. *South African Journal of African Languages* 22(2): 161-185.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2003. Effektiewe vordering met die *Woordeboek van die Afrikaanse Taal* soos gemeet in terme van 'n multidimensionele Liniaal. Botha, W. (Ed.). 2003. *'n Man wat beur. Huldigingsbundel vir Dirk van Schalkwyk*: 106-126. Stellenbosch: Bureau of the WAT.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2004. Crafting a Multidimensional Ruler for the Compilation of Sesotho sa Leboa Dictionaries. Mojalefa, J. (Ed.). 2004. *Rabadia Ratšhatšha: Indepth Literature, Linguistics, Translation and Lexicography Studies in African Languages. Festschrift in Honour of P.S. Groenewald*. Pretoria: J.L. van Schaik.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 2004a. Managing Eleven Parallel Corpora and the Extraction of Data in all Official South African Languages. Daelemans, W. and T. du Plessis (Eds.). 2004. *Multilingualism and Electronic Language Management*. Pretoria: J.L. van Schaik.

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.

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

Ziervogel, Dirk, Daniël P. Lombard and Pothinus C. Mokgokong. 1969. *Handboek van Noord-Sotho*. Pretoria: J.L. van Schaik.

Ziervogel, Dirk and Pothinus C. Mokgokong. 1975. *Pukuntšu ye kgolo ya Sesotho sa Leboa, Sesotho sa Leboa–Seburu/Seisimane / Groot Noord-Sotho-woordeboek, Noord-Sotho–Afrikaans/Engels / Comprehensive Northern Sotho Dictionary, Northern Sotho–Afrikaans/English.* Pretoria: J.L. van Schaik.

Addendum 1:   Treatment of the DC in the four existing desktop dictionaries for Sesotho sa Leboa [reproduced verbatim, including all inconsistencies and errors]

Pukuntšu woordeboek, Noord-Sotho–Afrikaans
— Kriel (1983³)

še, hier is dit.
šea, *še.a*, hier is hulle.
šeao, *še a.o*, daar is hulle(ma-klas).
šeba, *še'ba*, hier is hulle.
šebale, *še ba.le*, daar is hulle.
šebo, *še.'bo*, hier is dit (bo-klas).
šedi, *še.'di*, hier is hulle (*se* – klas).
šefa, *še.'fa*, hier is dit.
šefale, *še fa.lê*, daar is dit.
šefao, *še fa.o*, daar is dit.
šegola, *še go.la*, daar is dit.
šegoo, *še go.o*, daar is dit (go-klas).
šele, *še.'le*, hier is dit (le-klas); *leeba* -**,** hier is die duif.
šeleo, *še'le.o*, (le-klas), daar is dit.
šeo, *še.o, (šewe)*, daar is dit/hy, (die *n-, di-*klas).
sese, *se'se*, hier is dit.
seseo, *se se.o*, daar is dit.
šidi, *šidi*, ou spelling, kyk: *šedi*, hier is hulle.
šo, hier is hy/sy.
šolaa, *šo'la.a*, daar is hy/sy.
šole, *šo.'le*, daar is hy/sy.
šono, *'šo.no*, hier is hy/sy.
šoo, *šo.o*, daar is hy/sy.

Pukuntšu woordeboek, Noord-Sotho–Afrikaans
— Kriel, Van Wyk and Makopo (1989⁴)

še, kop. dem. 1, kl 3/9, H: hier is (dit).
šea, kop, dem. 1, kl 8, HH: hier is (hulle).
šeao, kop. dem 2, kl 8, HHL: daar is (hulle).
šeba, kop. dem. 1, kl 2, H: hier is (hulle).
šebalê, kop. dem. 3, kl 2, HHL: dáár is (hulle).
šebo, kop. dem. 1, kl 14, HH: hier is (dit).
šedi, kop. dem. 1, kl 8/10, HH: hier is (hulle).
šefa, kop. dem. 1, kl 16, HH: hier is (dit).

šefalê, kop. dem. 3, kl 16, HHL: dáár is (dit).
šefao, kop. dem. 3, kl 16, HHL: daar is (dit).
šegola, kop. dem. 3, kl 17, HHL: dáár is (dit).
šegoo, kop. dem. 2, kl 17, HHL: daar is (dit).
šele, kop. dem. 1, kl 5, HH: hier is (dit).
šeleo, kop. dem. 2, kl 5, HHL: daar is (dit).
šeo, kop. dem. 2, kl 4/9, HL: daar is (dit/hulle).
sese, kop. dem. 1, kl 7, HH: hier is (dit).
seseo, kop. dem. 2, kl 7, HHL: daar is (dit).
šidi, kop. dem. 1, kl 8/10, HH: hier is (hulle).
šo, kop. dem. 1, kl 1/3, H: hier is (hy/sy/dit.
šolê, kop. dem. 3, kl 1/3, HL: dáár is (hy/sy/dit).
šono, kop. dem. 1a, kl 1/3, HL: hier(so) is (hy/sy/dit).
šoo, kop. dem. 2, kl 1/3, HL: daar is (hy/sy/dit).

The New English–Northern Sotho Dictionary, Northern Sotho–English
— Kriel (1976⁴)

'še, dem., pron., here it is.
še'a, dem., here they are.
še'ba, dem., here they are.
'še'di, dem., **di–**class, here they are.
še'di'le, dem., pl., there they are, yonder.
še'di'o, dem., there they are.
še'fa, dem., here it is, here you are.
še'le, dem., adj., here it is.
še'leo, dem., adj., there it is.
šidi'le-e, dem., there they are.
ši'dio, – se'dio, dem., there they are (animal or things).
'šo, dem., here he is, here she is.
'šole, dem., there he (she) is.
'šono, dem., here he (she) is.

Comprehensive Northern Sotho Dictionary, Northern Sotho–Afrikaans/English — Ziervogel and Mokgokong (1975)

ŠÉ- [prefigale element by vorming van kop. dem.] // [prefixal element in formation of cop. dem.] v. **ŠÉBÁ, ŠÉO,** etc.

ŠÉ [dem. kop. I kl. **n-, me-**] hier is hy/sy/dit/hulle // [dem. cop. I cl. **n-, me-**] here he/she/it is, there they are

ŠÉÁ [dem. kop. I kl. **ma-**] hier is hulle // [dem. cop. I cl. **ma-**] here they are

ŠÉÁLE (**šealê**) [dem. kop. III kl. **ma-**] dáár is hulle // [dem. cop. III cl. **ma-**] there they are over there

ŠÉÁO (**šeaô**) [dem. kop. II kl. **ma-**] daar is hulle // [dem. cop. II cl. **ma-**] there they are

ŠÉBÁ [dem. kop. I kl. **ba-**] hier is hulle // [dem. cop. I cl. **ba-**] here they are

ŠÉBÁLE (**šebalê**) [kop. dem. III kl. **ba-**] dáár is hulle // [cop. dem. III cl. **ba-**] there are they over there

ŠÉBÁO (**sebaô**) [kop. dem. II kl. **ba-**] daar is hulle // [cop. dem. II cl. **ba-**] there they are

ŠÉBÓ [kop. dem. I kl. **bo-**] hier is dit // [cop. dem. I cl. **bo-**] here it is

ŠÉBÓLA [dem. kop. III kl. **bo-**] dáár is dit // [dem. cop. III cl. **bo-**] there it is over there

ŠÉBÓO (**šeboô**) [dem. kop. II kl. **bo-**] daar is dit // [dem. cop. II cl. **bo-**] there it is

ŠÉDÍ v. **ŠÍDÍ**

SÉDÍLA v. **ŠÍDÍLA**

ŠÉDÍO v. **ŠÍDÍO**

ŠÉFÁ [dem. kop. I vir lo. klasse] hier is dit // [dem. cop. I for lo. classes] here it is

ŠÉFÁLE (**šefalê**) [dem. kop. III vir lo. klasse] dáár is dit // [dem. cop. III for lo. classes] there it is over there

ŠÉFÁO (**šefaô**) [dem. kop. II vir lo. klasse] daar is dit // [dem. cop. II for lo. classes] there it is

ŠÉGÓLA [dem. kop. III vir lo. klasse (dia.)] dáár is dit // [dem. cop. III for lo. classes (dia.)] there it is over there

ŠÉGÓO (**šegoô**) [dem. kop. II vir lo. klasse (dia.)] daar is dit // [dem. cop. II for lo. classes (dia.)] there it is

ŠÉLA [dem. kop. III kl. **n-, me-**] dáár is hy/sy/dit/hulle // [dem. cop. III cl. **n-, me-**] there he/she/it is over there, there they are over there

ŠÉLE [dem. kop. I kl. **le-**] hier is hy/sy/dit // [dem. cop. I cl. **le-**] here he/she/it is

ŠÉLÉLA [dem. kop. III kl. **le-**] dáár is hy/sy/dit // [dem. cop. III cl. **le-**] there he/she/it is over there

ŠÉLÉO (**šeleô**) [dem. kop. II kl. **le-**] daar is hy/sy/dit // (dem. cop. II cl. **le-**] there he/she/it is

ŠÉO (**seô**) [kop. dem. II kl. **n-, me-**] daar is hy/sy/dit/hulle // [cop. dem. II cl. **n-, me-**] there he/she/it is, there they are

SÉSE [kop. dem. I, kl. **se-**] // [cop. dem. I, cl. **se-**] hier is dit // here it is

SÉSÉLA [kop. dem. III, kl. **se-**] // [cop. dem. III, cl. **se-**] doer is dit // there it is over there

SÉSÉO (**seseô**) [kop. dem. II, kl. **se-**] // [cop. dem. II, cl. **se-**] daar is dit // there it is

ŠÍ- [geassimileerde vorm van **še-**] // assimilated form of **še-**], cf. **ŠIDI**

ŠÍDÍ (< **šedi**) [dem. kop. I kl. **di-, din-**] hier is hulle // [dem. cop I cl. **di-, din-**] here they are

ŠÍDÍLA [dem. kop. III kl. **di-, din-**] dáár is hulle // [dem. cop. III cl. **di-, din-**] there they are over there

ŠÍDÍO (**šidiô**) [dem. kop. II kl. **di-, din-**] daar is hulle // [dem. cop. II cl. **di-, din-**] there they are

ŠO [dem. kop. I kl. **mo-**] hier is hy/sy/dit // [dem. cop. I cl. **mo-**] here he/she/it is

ŠÓLA [dem. kop. III kl. **mo-**] dáár is hy/sy/dit // [dem. cop. III cl. **mo-**] there he/she/it is over there

ŠÓLE (**šolê**) [dem. kop. III kl. **mo-**] dáár is hy/sy/dit // [dem. cop. III cl. **mo-**] there he/she/it is over there

ŠOO (**šoô**) [dem. kop II kl. **mo-**] daar is hy/sy/dit // [dem. cop. II cl. **mo-**] there he/she/it is

Addendum 2:   Frequencies of all DCs and homonymous items in a 6.1-million-word Sesotho sa Leboa corpus [∑ = total frequency; ? = not possible to see class affiliation; G = occurs in grammar book(s) only; *v.* = verb(s); *aux.* = auxiliary verb; *n.* = noun(s); *adj.* = adjective; *e.* = enumerative stem]

	I	Ia	Ib	II	IIa	IIb	III	IIIa
1&3	**šo**	**šono**	**šokhwi**	**šoo**	**šouwe**	**šowe**	**šola**	**šole**
	∑ 564	∑ 3	∑ 18	∑ 83	∑ 3	∑ 19	∑ 20	∑ 102
	I.1 348	Ia.1 3	Ib.1 16	II.1 76	IIa.1 2G	IIb.1 17	*v.* 9	IIIa.1 97
	I.3 198	Ia.3 0	Ib.3 2	II.3 7	IIa.3 1G	IIb.3 2	III.1 8	IIIa.3 5
	? 18						III.3 3	
2	**šeba**	**šebano**	**šebakhwi**	**šebao**	**šebauwe**	**šebawe**	**šebala**	**šebale**
	∑ 248	∑ 2	∑ 1	∑ 22	∑ 2	∑ 1	∑ 5	∑ 25
	I.2 125	Ia.2 2	Ib.2 1G	II.2 22	IIa.2 2G	IIb.2 1	III.2 5	IIIa.2 25
	v. 123							
4&9	**še**	**šeno**	**šekhwi**	**šeo**	**šeuwe**	**šewe**	**šela**	**šele**
	∑ 412	∑ 4	∑ 7	∑ 45	∑ 2	∑ 4	∑ 12	∑ 817
	I.9 244	*aux.* 3	Ib.9 7	II.9 41	IIa.9 1G	IIb.9 4	III.9 7	*e.* 719
	I.4 153	Ia.4 1	Ib.4 0	II.4 4	IIa.4 1G	IIb.4 0	*n.* 5	I.5 80
	? 15	Ia.9 0					III.4 0	IIIa.9 18
								IIIa.4 0
5	**šele**	**šeleno**	**šelekhwi**	**šeleo**	**šeleuwe**	**šelewe**	**šelela**	**šelele**
	∑ 817	∑ 0	∑ 0	∑ 7	∑ 1	∑ 0	∑ 0	∑ 2
	e. 719			II.5 7	IIa.5 1G			IIIa.5 2
	I.5 80							
	IIIa.9 18							
6	**šea**	**šeano**	**šeakhwi**	**šeao**	**šeauwe**	**šeawe**	**šeala**	**šeale**
	∑ 135	∑ 0	∑ 0	∑ 17	∑ 1	∑ 0	∑ 1	∑ 5
	I.6 133			II.6 17	IIa.6 1G		III.6 1G	III.6 5
	v. 2							
7	**sese**	**seseno**	**sesekhwi**	**seseo**	**seseuwe**	**sesewe**	**sesela**	**sesele**
	∑ 100	∑ 0	∑ 0	∑ 15	∑ 1	∑ 0	∑ 34	∑ 20
	I.7 97			II.7 15	IIa.7 1G		*n.* 32	*n.* 19
	adj. 3						III.7 2	IIIa.7 1G
8&10	**šedi**	**šedino**	**šedikhwi**	**šedio**	**šediuwe**	**šediwe**	**šedila**	**šedile**
	∑ 502	∑ 0	∑ 0	∑ 75	∑ 0	∑ 0	∑ 0	∑ 15
	n. 293			II.10 70				IIIa.10 13
	I.10 185			II.8 5				IIIa.8 2
	I.8 24							
8'&10'	**šidi**	**šidino**	**šidikhwi**	**šidio**	**šidiuwe**	**šidiwe**	**šidila**	**šidile**
	∑ 36	∑ 0	∑ 0	∑ 8	∑ 2	∑ 0	∑ 8	∑ 4
	I.10 25			II.10 8	IIa.10 1G		*v.* 6	IIIa.10 2
	I.8 7			II.8 0	IIa.8 1G		III.10 1G	IIIa.8 2
	? 4						III.8 1G	

8"&10"	**šetše** ∑ 6846 *aux.* 6069 *v.* 775 I.10 2 I.8 0	**šetšeno** ∑ 0	**šetšekhwi** ∑ 0	**šetšeo** ∑ 1 II.10 1 II.8 0	**šetšeuwe** ∑ 0	**šetšewe** ∑ 0	**šetšela** ∑ 0	**šetšele** ∑ 0
14	**šebo** ∑ 24 I.14 24	**šebono** ∑ 0	**šebokhwi** ∑ 0	**šeboo** ∑ 1 II.14 1G	**šebouwe** ∑ 1 IIa.14 1G	**šebowe** ∑ 1 IIb.14 1	**šebola** ∑ 1 III.14 1G	**šebole** ∑ 0
15&17	**šego** ∑ 6 *n.* 6	**šegono** ∑ 0	**šegokhwi** ∑ 0	**šegoo** ∑ 0	**šegouwe** ∑ 0	**šegowe** ∑ 0	**šegola** ∑ 0	**šegole** ∑ 0
16	**šefa** ∑ 106 I.loc 106	**šefano** ∑ 0	**šefakhwi** ∑ 2 Ib.loc 2	**šefao** ∑ 16 II.loc 16	**šefauwe** ∑ 1 IIa.loc 1G	**šefawe** ∑ 0	**šefala** ∑ 1 III.loc 1G	**šefale** ∑ 0
18	**šemo** ∑ 7 I.loc 7	**šemono** ∑ 0	**šemokhwi** ∑ 0	**šemoo** ∑ 0	**šemouwe** ∑ 0	**šemowe** ∑ 0	**šemola** ∑ 0	**šemole** ∑ 0

*šea → sea *v.* examine – 2
šeba *v.* relish, eat as a titbit – 123
šedi *n. cl. 9* care, attention – 293
*šego → lešego *n. cl. 5* blessing – 3;
    → bošego *n. cl. 14* night – 2;
    → sešego *n. cl. 7* granary – 1
*šela → lešela *n. cl. 5* cloth – 5
šele *enumerative stem* strange, foreign, different – 719
*šeno → seno *aux.* as soon as – 3

-sese *adj.* thin, small – 3
Sesela (person's name) – 32
sesele *n. cl. 7* badger – 19
šetše[1] *aux.* already – 6069
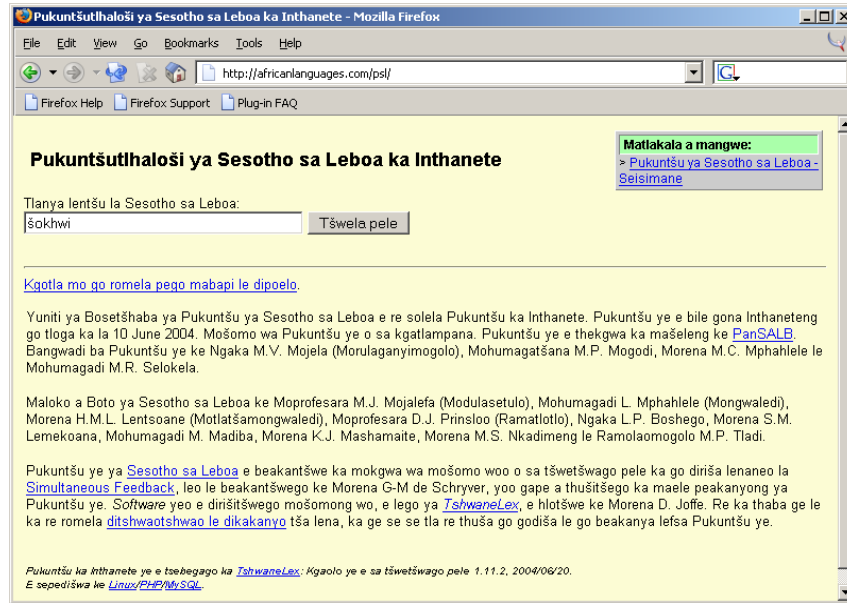**šetše[2]** *v.* **1.** remain (behind) – 520; **2.** follow (behind) – 160
šetše[3] *v.* must pay attention; **..ga/sa/se..~** not pay attention – 95
šidila *v.* **1.** iron – 5; **2.** massage – 1
šola *v.* bring bad luck – 9

Addendum 3:   Screenshot of a search for *šokhwi* in the Online PyaSsaL – input

Addendum 4:   Screenshot of a search for *šokhwi* in the Online PyaSsaL – output

Outer Texts in Bilingual Dictionaries[*]

Rufus H. Gouws, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa (rhg@sun.ac.za)*

Abstract: Dictionaries often display a central list bias with little or no attention to the use of outer texts. This article focuses on dictionaries as text compounds and carriers of different text types. Utilising either a partial or a complete frame structure, a variety of outer text types can be used to enhance the data distribution structure of a dictionary and to ensure a better information retrieval by the intended target user. A distinction is made between primary frame structures and secondary frame structures and attention is drawn to the use of complex outer texts and the need of an extended complex outer text with its own table of contents to guide the user to the relevant texts in the complex outer text. It is emphasised that outer texts need to be planned in a meticulous way and that they should participate in the lexicographic functions of the specific dictionary, both knowledge-orientated and communication-orientated functions, to ensure a transtextual functional approach.

Keywords: BACK MATTER, CENTRAL LIST, COMMUNICATION-ORIENTATED FUNC-TIONS, COMPLEX TEXT, CULTURAL DATA, EXTENDED COMPLEX TEXT, EXTENDED TEXTS, FRONT MATTER, FRAME STRUCTURE, KNOWLEDGE-ORIENTATED FUNCTIONS, LEXICO-GRAPHIC FUNCTIONS, OUTER TEXTS, PRIMARY FRAME, SECONDARY FRAME

Opsomming: Buitetekste in tweetalige woordeboeke. Woordeboeke vertoon dikwels 'n partydigheid ten gunste van die sentrale lys met min of geen aandag aan die buitetekste nie. Hierdie artikel fokus op woordeboeke as tekssamestellings en draers van verskillende teks-soorte. Met die benutting van óf 'n gedeeltelike óf 'n volledige raamstruktuur kan 'n verskeiden-heid buitetekste aangewend word om die dataverspreidingstruktuur van 'n woordeboek te verbe-ter en om 'n beter herwinning van inligting deur die teikengebruiker te verseker. 'n Onderskeid word gemaak tussen primêre en sekondêre raamstrukture en die aandag word gevestig op kom-plekse buitetekste en die behoefte aan 'n uitgebreide komplekse buiteteks met sy eie inhoudsop-gawe wat die gebruiker kan lei na die relevante tekste binne die komplekse buiteteks. Dit word benadruk dat buitetekste sorgvuldig beplan moet word en dat hulle deel moet hê aan die leksiko-grafiese funksies van die betrokke woordeboek, sowel die kennis- as die kommunikasiegerigte funksies, om 'n transtekstuele funksionele benadering te verseker.

Sleutelwoorde: AGTERTEKSTE, BUITETEKSTE, KENNISGERIGTE FUNKSIES, KOMMU-NIKASIEGERIGTE FUNKSIES, KOMPLEKSE TEKS, KULTURELE DATA, LEKSIKOGRAFIESE FUNKSIES, PRIMÊRE RAAM, RAAMSTRUKTUUR, SEKONDÊRE RAAM, SENTRALE LYS, UIT-GEBREIDE KOMPLEKSE TEKS, UITGEBREIDE TEKSTE, VOORTEKSTE

---

[*]   This article was presented as a keynote address at the Eighth International Conference of the African Association for Lexicography, organised by the Department of Germanic and Romance Languages, University of Namibia, Windhoek, Namibia, 7–9 July 2003.

Introduction

The average dictionary user does not judge a book, even a dictionary, by its cover. To these users the success of a dictionary is not determined primarily by its extent, the lay-out or the quality of the printing but rather by the way in which they manage to achieve a successful procedure of dictionary consultation by retrieving from the dictionary the kind of information they need and have been looking for. The lack of a dictionary culture among the majority of dictionary users leads to a situation where these users do not realise that the success of a dictionary consultation procedure does not only depend on the relevant data being included in the dictionary. Access to the data is a prerequisite for a successful retrieval of information. Not only the data included in a dictionary but also the structuring of the dictionary and the positioning of the data as entries determine the quality of the dictionary as a utility instrument. When evaluating a dictionary as a container of knowledge one has to focus on both the container and the knowledge. In the past, lexicographers, dictionary users and dictionary critics too often only looked at the contents of the dictionary. This bias was detrimental to the study, the development and the use of the structure of a dictionary and the different structural components. A stronger structure-directed approach puts the focus on both the container and the contained data.

The compilation of a dictionary is only part of the overall lexicographic process and is preceded by the planning phase during which the dictionary conceptualisation plan is formulated and a sound theoretical basis is established. The compilation process should adhere to the dictionary conceptualisation plan. Dictionaries are compiled for real users and intended to be used as utility instruments. A user-driven lexicography compels lexicographers to plan their dictionaries in such a way that their target users can achieve an optimal retrieval of information. Consequently the dictionary plan should involve more than an account of the data types to be included in the treatment of the lemmata. It has to reflect on the dictionary as a carrier of text types and the positioning of these texts in the different textual venues.

The frame structure and the structure of outer texts

The traditional bias towards the contents of dictionary articles as the main field of interest for dictionary users, lexicographers and metalexicographers has resulted in the narrowing of the focus on dictionaries to a mere study of the central list. Recent metalexicographic research rather switched the interest to the structure of dictionaries. This deviates the attention from the central list towards a more inclusive structural approach with the assumption that the central list can be complemented by front and/or back matter texts functioning as outer texts. In a dictionary the central list as well as any text occurring in the front or back matter can be extended and this extension can be complete or par-

tial. A partial extension leads to a given text being complemented by either front or back matter texts whereas a complete extension results in a text complemented by both front and back matter texts.

Partially extended texts

```
              Text 1                              Text 2
             /      \                            /      \
       FM text    Main text              Main text    BM text
```
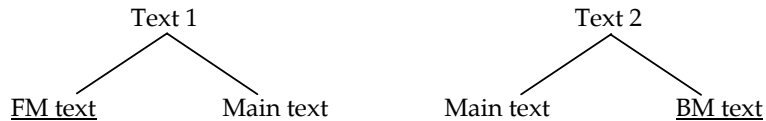
Completely extended text

```
                        Text 3
                    /     |     \
              FM text   Main text   BM text
```
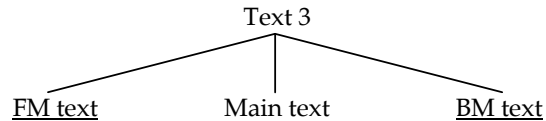
According to Kammerer and Wiegand (1998), a dictionary with a central text, and front and back matter texts constitutes a frame structure. The employment of a frame structure gives the lexicographer a much wider range of possibilities when it comes to decisions regarding the distribution of the data to be included in the dictionary and, if used with the necessary care, enhances the access of the user to a more comprehensive data selection. A user-dominated approach should have a distinct influence on the frame structure of modern dictionaries. This applies to the contents and the structure of the texts presented in the frame. As a carrier of text types a dictionary can be regarded as a so-called "big text" or a text compound. Hausmann and Wiegand (1989: 330) argue that the front matter as a whole is not a functional part of a dictionary but rather an arbitrary set of functional text parts. Kammerer and Wiegand (1998: 228) take this line of argument further by saying that the front and back matter are textual parts but not textual constituents because textual constituents should always be functional textual parts. To illustrate this: the user's guide text included in the front matter of a dictionary is not a user's guide to the front matter but a user's guide to the dictionary as a whole. It is a functional component of the dictionary and not of the front matter and therefore it is a textual constituent.

```
                          Dictionary
            /       /        |        \         \        \
       Title  Imprint  Users' guide  Central list  Addendum 1  Addendum 2
```
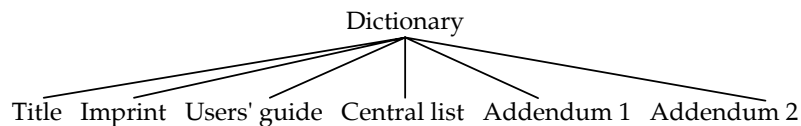
Gouws (2001: 526) introduces the notion of an extended outer text not only to reflect the situation where a central list is complemented by an outer text but also to make provision for an outer text with its own outer text(s). A typical example of an extended outer text could be found in a dictionary with one or more complex outer texts, i.e. an outer text consisting of a number of subtexts, where this complex text is extended to include its own table of contents as a secondary outer text which enhances the access to the different subtexts. The German learner's dictionary *Wörterbuch Deutsch als Fremdsprache* has a frame structure which hosts a variety of outer texts. In the back matter section of this dictionary there is a complex outer text, indicated by the table of contents in the front matter functioning as a primary outer text as *Wortfelder* (Word fields). This complex outer text comprises a number of subtexts representing different words included in the central list which refers to specific semantic fields, i.e. man, environment, artefacts, religion, time, etc. No less than twelve subtexts are included in this complex outer text.

Complex outer text

Word fields

Man   Environment   Artefacts   Religion   Time
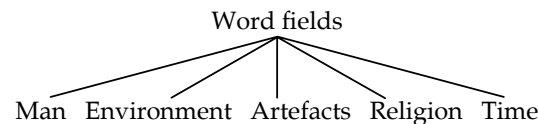
To ensure rapid access to the different subtexts this outer text has been extended to display its own table of contents. This table of contents only has the outer text *Wortfelder* in its scope. It is a functional textual constituent but directed at a single outer text.

Extended complex outer text

Word fields

Table of contents   Man   Environment   Artefacts
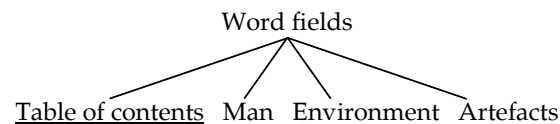
A text which is partially extended constitutes a partial frame, whereas a complete extension results in a text being the core of a frame structure with its own front and back matter texts. Gouws (2001: 527) makes a distinction between a primary and a secondary frame structure. Outer texts functioning as functional constituents of the dictionary as a text compound, i.e. those texts framing the central list of the dictionary, form the primary outer texts and in combination with the central list they constitute the primary frame structure.

Primary frame

Primary frame
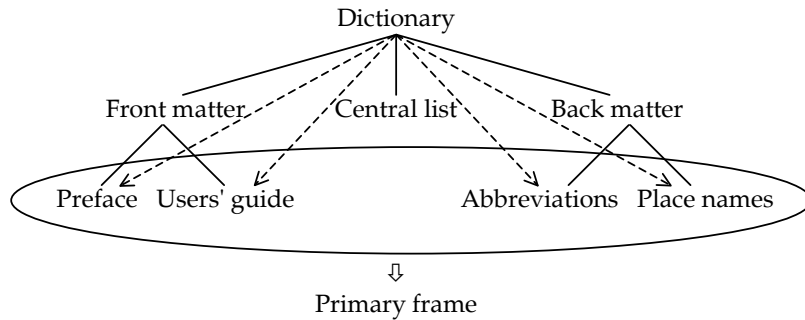
Where an outer text displays its own complete extension resulting in a frame structure it represents a secondary frame structure. A secondary frame structure typically applies to outer texts but it can also prevail in dictionaries with more than one word list, e.g. a bilingual dictionary with a biscopal character (cf. Hausmann and Werner 1991), where more than one of the word lists could display a frame structure with these outer texts directed at the specific word list and not at the dictionary as a whole. An example of this kind of outer texts can be found in the Afrikaans–English bilingual dictionary *Tweetalige Woordeboek/Bilingual Dictionary* (henceforth abbreviated as BD). It is a biscopal dictionary with the central list containing two major components, i.e. the section Afrikaans–English and the section English–Afrikaans to form a word list series.

Secondary frame

Secondary frame

```
                            Dictionary

        Front matter      Central list      Back matter

              Word list 1              Word list 2

          FM texts   Main text   BM texts

                          ⇩
                    Secondary frame
```
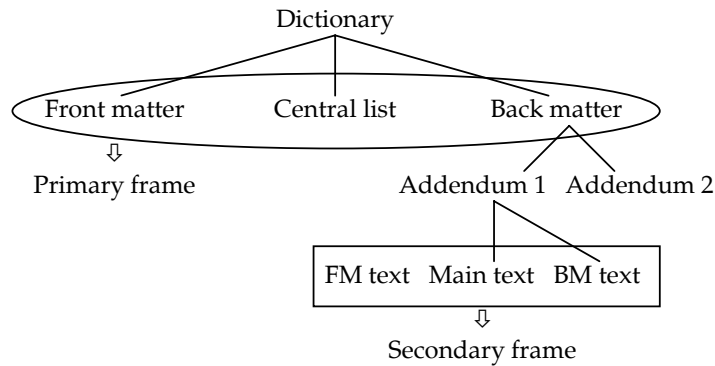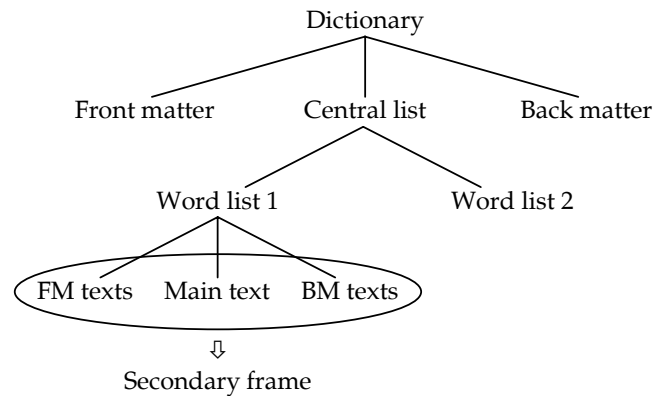
In the BD each component in the word list series displays a frame structure which functions as secondary frame. As a textual compound this dictionary contains primary front matter texts followed by the word list series. The final text in the BD is a back matter text but it is not a primary back matter text. It is rather directed at the last member of the word list series. The BD is an example of a dictionary with secondary frame structures but without a primary frame structure. On the first level of text extension, this dictionary displays a partial primary frame. This reveals that a primary frame structure is not a prerequisite for the use of a secondary frame structure.

The planning of outer texts

The planning of the outer texts and the selection of data to be included in these texts should form an integral part of the dictionary conceptualisation plan. Therefore the data distribution programme should be worked out with meticulous care to ensure a functional data distribution structure. Bergenholtz, Tarp and Wiegand (1999) make a distinction between integrated and unintegrated outer texts. This is based on the relation between an outer text and the genuine purpose of a specific dictionary. Unintegrated outer texts complement the central list and are not needed to retrieve the information presented in the articles of the central list, neither do they add to the treatment of the subject matter of the dictionary. Integrated outer texts function in co-ordination with the central list and are aids in ensuring an optimal and full retrieval of information and an accomplishment of the genuine purpose.

The front matter of a dictionary could contain a text in which the lexicographer expresses his/her gratitude to people or institutions for their assistance with the dictionary project. Although this text is important it does not play a pivotal role in conveying data regarding the treatment units or the subject matter of the dictionary to its potential users. This is an example of an unintegrated outer text.

Grammatical data entries can be allocated to different texts. One possibility is that each article could include an extensive account of the grammatical aspects of the lemma sign. As an alternative the lexicographer could restrict the presentation of grammatical data in an article and include an outer text with a comprehensive and systematic discussion of grammatical aspects or a clearly marked table with a cryptic indication of some relevant grammatical categories. Users can be referred from the article to a specific section in the relevant outer text. The *Longman Dictionary of Contemporary English* offers scant grammatical data in the articles of the central list. The verb *refine* is marked as a verb and this entry is followed by an entry giving the code [T1]. The knowledgeable dictionary user will interpret this code as a cross-reference entry directing the user to a back matter text. The final back matter text in this dictionary is the *Table of codes* which gives all the grammatical codes used in the dictionary along with an explanation. The code [T] refers to a transitive verb with one object and the code [1] indicates that the specific word is followed by one or more nouns or pronouns. The user can then deduce that [T1] indicates that *refine* is a transitive verb which is followed by one or more nouns or pronouns.

An outer text like the *Table of codes* does not only complement the central list but constitutes a functional part of the lexicographic treatment of the lemmata. This would be an example of an integrated outer text.

A lexicographer may decide to exclude all idioms from the articles in the central list and phase them out to a separate outer text in the back matter of the dictionary. In this text, the idioms are included as treatment units and their treatment constitute part of the overall lexicographic treatment given as part of the assignment of the lexicographer towards the user. Yet again, such a text does not merely complement the central list as a nice-to-have addition but it falls directly within the scope of the lexicographic presentation a user expects from the given dictionary. Consequently it has to be regarded as an integrated outer text. Gouws (2002) argues:

> The different texts included in a dictionary can individually contribute to the successful achievement of the genuine purpose of a dictionary. These texts belong either to the class of so-called "help texts" with the function to assist the user with a more successful dictionary consultation procedure or, in as far as they contain a part of the lexicographic data presented in the dictionary, they are integrated texts from which the user can retrieve information regarding the subject matter of the dictionary. Both these types of outer texts stand in a relation with the genuine purpose of the dictionary. They either serve (as help texts) to explain the use of the dictionary to the user or they are, as a result of an integration into the dictionary internal data distribution, directly integrated into the genuine purpose of the dictionary.

The planning and selection of outer texts in any dictionary should not be done in a random way but it should be determined by the typological nature of the dictionary, the user and usage situation and the consequent functions of the dictionary. The data distribution structure must determine the way in which

outer texts function as either integrated or unintegrated outer texts and the nature and extent of the data to be accommodated in these texts. The frame structure is not only a means to describe the distribution of texts in a dictionary but it is primarily an instrument to assist the lexicographer in a comprehensive use of the dictionary as a carrier of different text types and to ensure the user access to the data on offer in the dictionary. Lexicographers and dictionary users should be made aware of the fact that an outer text is not only an ornamental part of a dictionary or a text included as a page-filler. Outer texts are functional textual constituents and they need to be utilised accordingly. For the lexicographer the challenge is there to use outer texts in an innovative way to help with the transfer of data relevant to the needs of the target user of the specific dictionary. Realising the possibilities of outer texts demands a wholly new approach to the data distribution programme.

Bilingual dictionaries

When considering the use of outer texts during the planning phase of any dictionary, a lexicographer should work with the suggestion made by Hausmann and Wiegand (1989) that a dictionary should include two compulsory texts, i.e. the central list and a text, usually presented in the front matter, which includes the users' guidelines to ensure a successful use of the dictionary. Besides this compulsory outer text, a dictionary would typically also have front matter texts like the title page and the imprint. Another front matter text which is frequently used but often not to its full potential is the table of contents. The role of this type of outer text will be discussed in more detail in a subsequent section of this article.

Where front matter texts have a limited occurrence in dictionaries, back matter texts are employed even less frequently. This is also the situation in bilingual dictionaries. Many bilingual dictionaries have been compiled without a prior consideration of the function of the dictionary or a knowledge of the use of outer texts. Within a theory of lexicographic functions different functions are identified and distinguished to fulfil the needs of different dictionary users in different situations of dictionary use. On practical grounds the compilation of a series of four to eight monofunctional dictionaries for a given language pair (cf. Hausmann 1977, Kromann et al. 1984), could be a too expensive endeavour. Consequently Wiegand (1996) argues in favour of the compilation of multifunctional dictionaries. The nature and extent of the functions of any given dictionary should play a decisive role in the selection and presentation of its outer texts. In this regard lexicographers should pay the necessary attention to the needs of their users in terms of communication-orientated functions like text production and text reception as well as knowledge-orientated functions (cf. Tarp 2000, Gouws and Tarp 2004 and Bergenholtz and Tarp 2003). Both these main categories of dictionary functions should have an influence on the outer texts presented in a bilingual dictionary.

Although researchers in the field of metalexicography strive towards the

formulation of a general theory of lexicography and although such a theory plays an important role in both dictionary research and the lexicographic practice, one should never underestimate the unique character of any given dictionary. Because dictionaries are compiled for very specific target users the needs of that user group will often result in criteria formulated in a general theory of lexicography being adapted or even altered. When planning a new dictionary project lexicographic theory should provide invaluable guidelines to the lexicographer but the user situation and the usage situation should eventually determine the way in which the functions of the dictionary are formulated in order to identify the genuine purpose of the dictionary. Yet again, when one looks at the outer texts given in some existing dictionaries it is hard to see how they contribute to the functions and the genuine purpose of the dictionary. This remark does not imply that unintegrated outer texts, on account of the fact that they are not integrated into the genuine purpose of the dictionary, do not contribute to achieving the functions of the dictionary. A carefully selected and well-constructed outer text, although not integrated into the genuine purpose of a dictionary, could still form part of the functions allocated to the specific dictionary. Unfortunately, although outer texts have been used in many dictionaries for many years one has the impression that the selection of these texts had often been done in an arbitrary way without negotiating the genuine purpose or the functions of the specific dictionary or the needs of the intended target user.

Functions and outer texts

Knowledge-orientated and communication-orientated functions

When consulting a dictionary a user may need data on a specific topic or on the language treated in the dictionary. To help with the provision of assistance for these situations Bergenholtz and Tarp (2003) distinguish between knowledge- and communication-orientated functions. Knowledge-orientated functions assist the user by providing

> — general cultural and encyclopaedic data,
> — special data about the subject field, and
> — data about the language.

Communication-orientated functions assist the user to solve problems related to

> — text production in the native language,
> — text production in the foreign language,
> — text reception in the native language,
> — text reception in the foreign language,

— translation of texts from the foreign to the native language, and
— translation of texts from the native to the foreign language.

The function of a dictionary is not only determined by the users but also by the usage situation. The situation in which a dictionary is used should have a definite influence on the data distribution programme and on the function of that dictionary.

Outer texts with a knowledge-orientated function

The *Everyman's Dictionary of First Names* primarily has a knowledge-orientated function which prevails in the presentation of e.g. data regarding the origin and development of first names. The back matter section of this dictionary contains four outer texts which clearly share the knowledge-orientated function of the central list. These texts are: *The top fifty first names for boys, England and Wales; The top fifty first names for girls, England and Wales; The top fifty first names for boys, USA;* and *The top fifty first names for girls, USA.* Between the central list and these back matter texts a relation of thematic progression, i.e. coherence, is established (cf. Wiegand 1998: 8). This functional coherence is ascertained by the knowledge-orientated contents of the outer texts. Looking at the subject matter of this dictionary, one can understand the motivation for the inclusion of these specific back matter texts. From a data distribution perspective it would have been difficult to present the data offered by these outer texts by means of the central list. The words occurring in these outer texts could have been marked in the central list as belonging to the respective top fifty names but the user would not have had access to these names as a group. These outer texts are therefore functional constituents which add to the thematic progression and to an efficient transfer of data.

In an environment where multifunctional dictionaries represent the default approach, bilingual dictionaries are directed not only at the source and target languages but also at the needs of at least two user groups. These users often have different native languages and different cultures. Traditionally, and too often this tradition still prevails, bilingual dictionaries are primarily dominated by a semantic bias, i.e. a treatment directed at achieving semantic equivalence although this semantic equivalence does not necessarily lead to communicative equivalence (cf. Gouws 1989; 2000). The typical treatment presented in bilingual dictionaries focuses on the presentation of translation equivalents, sometimes accompanied by co-text and context entries, to assist the user in choosing the proper target language form for a given source language item. This illustrates a typical communication-orientated function.

Looking at the history and development of the lexicographic practice, Gallardo (1980) clearly indicates that in non-standardised language situations bilingual dictionaries but not monolingual dictionaries will exist. It often is the case in a multilingual environment that speakers rely on bilingual dictionaries as their only lexicographic resource. Such a dictionary should then not only have

a text reception and a text production function, both communication-orientated functions, but it should also display knowledge-orientated functions because users rely on such a bilingual dictionary for some general cultural and encyclopaedic data. The bilingual dictionaries by Van Dale *Groot woordenboek Engels–Nederlands* and *Groot woordenboek Nederlands–Engels* are compiled for Dutch users. The treatment allocated to the lemmata in the central list typically represents communication-orientated functions. However, the back matter contains a number of outer texts with both communication- and knowledge-orientated functions. The British and American terms used for weights and measures differ from the metric system terms used in the Netherlands. From a cultural and encyclopaedic perspective the Dutch user of this dictionary is in need of this kind of knowledge. Consequently the lexicographers have included back matter texts presenting this type of data.

Maten en gewichten

B — UK-stelsel

minim	(min)			0,059	ml
fluid drachm	(fl dr)	60	minims	3,55	ml
fluid ounce	(fl oz)	8	fluid drachms	28,41	ml
gill	(gi)	5	fluid ounces	0,142	l
pint	(pt)	4	gills	0,568	l
quart	(qt)	2	pints	1,136	l
gallon	(gal)	4	quarts	4,546	l
peck	(pk)	2	gallons	9,092	l
bushel	(bu)	4	pecks	36,369	l
hogshead	(hhd)	52½	of	238,5 of	
		54	gallons	245,5	l
quarter		8	bushels	290,94	l

C — US-stelsel

I *eenheden voor vloeistoffen*

minim	(min			0,062	ml
fluid dram	(fl dr)	60	minims	3,70	ml
fluid ounce	(fl oz)	8	fluid drams	29,57	ml
gill	(gi)	4	fluid ounces	0,118	l
pint	(liq pt)	4	gills	0,473	l
quart	(liq qt)	2	pints	0,946	l
gallon	(gal)	4	quarts	3,785	l
barrel (petroleum)					
	(oil bll)	42	gallons	158,97	l
hogshead		63	gallons	238,46	l

II *eenheden voor droge waren*

pint	(dry pt)			0,550	l
quart	(dry qt)	2	pints	1,101	l
gallon	(dry gal)	4	quarts	4,405	l
peck		2	gallons	8,809	l
bushel		4	pecks	35,238	l
barrel	(dry bll)	105	quarts	115,6	l

If it had been a bilingual dictionary with Dutch and, say, German or French as treated language pair the situation of usage would have been different. The same Dutch target user would have been confronted with a situation where the speakers of both the source and the target language use the same system of weights and measures. Such a back matter text would then not have been necessary for the needs of the specific dictionary.

A knowledge-orientated approach may also lead to the inclusion of back matter texts that transforms the dictionary from mono-accessible to poly-accessible. The *Nuovo Dizionario,* a bilingual dictionary with Italian and English as treated language pair, includes as one of its back matter texts an alphabetically ordered list of Italian words from the semantic field of motor-car terms. In this back matter text, these Italian terms are co-ordinated with their English equivalents. All these terms are also included in the central list as lemma signs and their articles contain the same equivalents provided in the back matter text along with the other microstructural data. However, the knowledge-orientated function motivates the inclusion of a back matter text with a set of terms belonging to the same semantic field and the occurrence of these terms in close proximity assists the user in retrieving the lexicon of a semantic field without being impeded by the remote appearance of these terms in the central list of the dictionary.

The *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Dictionary* (henceforth abbreviated as RD) is a bilingual dictionary which is based on the same central list as the BD but its central list has been increased by the inclusion of two additional columns on each page. These columns accommodate a more comprehensive treatment of some of the words included as lemma signs on the relevant page but also some inserted text boxes with pragmatic data. This aspect is discussed in Gouws (1996) and will not receive further attention here. Of relevance to this article is the use of outer texts in the RD.

The structure of the *Reader's Digest Afrikaans–Engelse Woordeboek/ English–Afrikaans Dictionary*

The RD displays both a primary and a secondary frame structure. In this regard it differs from the BD with its complete secondary frame structure but only a partial primary frame. Like the BD the central text of the RD consists of two separate word lists to provide for the Afrikaans–English and English–Afrikaans sections respectively.

The primary frame structure of this dictionary is formed by a number of front and back matter texts, featuring both integrated and unintegrated outer texts. Following Herberg (1998: 339), Gouws (2002: 473) distinguishes between two types of outer texts, i.e. the help texts, which assist the user to achieve a more successful dictionary consultation procedure, and those outer texts containing a part of the lexicographic data to be presented to the user. The front matter texts in the RD primarily belong to the category of help texts, e.g. the

text *A new kind of bilingual dictionary.* The table of contents is a primary outer text because it guides the user to both the word lists in the central list and different texts included in the front and back matter sections. As part of the rapid outer access structure of the dictionary (cf. Hausmann and Wiegand 1989: 333), this is an integrated outer text because it assists the user in realising the genuine purpose of the dictionary by presenting an early milestone on the external search route to help the user in the proper use of the dictionary and the needed access to the different functional textual constituents. The front matter texts featuring in the secondary frame structure of the RD also show a strong help tendency, e.g. the text *How to use this dictionary.*

The structure of the RD

```
                                RD
               _____|_____
              |                |                |
        Front matter      Central list     Back matter
                        _____|_____
                       |             |
                 Word list A–E    Word list E–A
                               _____|_____
                              |      |      |
                             FM  Central list  BM
              _____|                |
             |  |      |       |       |       |
         [Title Use  Suppl. text  Dict. Abbrev.  Pron.]   [Gen. Abbrev.]
```
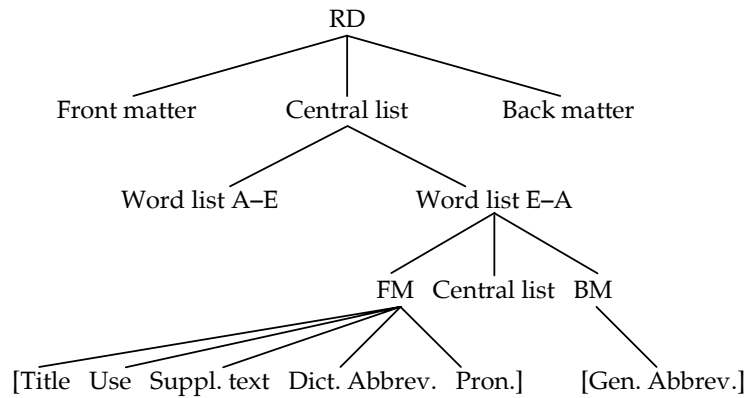
Primary frame of the BD

[Front matter	Central list	Back matter]
<u>Front matter</u>		<u>Back matter</u>
Pronunciation		Writing a letter
Title		Forms of address
Title +		Place names
Imprint		People's names
New kind of dictionary		See and say
Contents		Numbers and time

The lack of a dictionary culture which makes users aware of the spectrum of data on offer in all the different texts of a dictionary should compel a lexicographer to make the access to data as clear and unambiguous as possible. This assignment of the lexicographer falls within the scope of what Gouws (2003) refers to as the data exposure structure. Users should realise that outer texts are used to achieve a more comprehensive and more user-friendly data distribu-

tion. Consequently the nature and contents of the different outer texts need to be explicated. In the RD the back matter text *See and say* is a complex outer text with seven subfields. The RD is a multifunctional dictionary. In its communication-orientated approach it makes provision for text production, text reception and bidirectional translation functions. The expanded data load, compared to the BD, and the nature of this load makes this a so-called family dictionary. A wide-ranging group of users benefit from the lexicographic treatment presented in this dictionary. These users are not expecting the type of data on offer in the different outer texts. The subtexts presented in the outer text *See and say* cover different themes and consist of line-drawings in which a variety of objects have been numbered. The themes introduced in this complex outer text are: *The human body, Where you work, Where you live, Cuts of meat, The car, Sport* and *The world.*

See and say   **Sien en sê**

Where you live   **Waar jy woon**

		In the lounge		**In die sitkamer**
	curtains	1	gordyne	
	curtainrod	2	gordynstok	
	display cabinet	3	toonkas, glaskas	
	grandfather clock	4	staanhorlosie	
	(loud)speaker	5	luidspreker	
	television set	6	televisiestel	
	aerial	7	lugdraad, (opvangdraad)	
	record player	8	platespeler	
	tuner/amplifier	9	instemmer/versterker	
	tape deck	10	bandspeler	
	video recorder	11	video-opnemer	
	fireplace	12	kaggel	
	firescreen		kaggelskerm	
	firedog		(h)erdyster, vuuryster	
	slow-combustion stove		smeulstoof	
	firegrate	13	herdrooster	
	mantelpiece	14	mantelrak, kaggelrak	
	settee	15	rusbank, sitbank	
	tea trolley	16	teewaentjie	
	Oriental rug	17	Oosterse mat	
	coffee table	18	koffietafel(tjie)	
	vase	19	blompot	
	easy chair	20	leun(ing)stoel, armstoel	
	reclining chair	21	lêstoel	
	cushion	22	(stoel)kussing	
	rocking chair	23	rystoel, skommelstoel	
	wall-to-wall carpet	24	volvloertapyt	
	picture	25	prent	
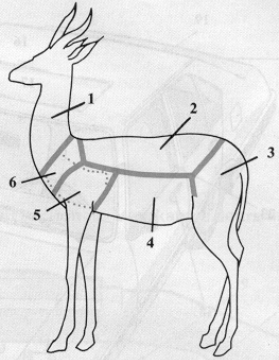
The drawings are complemented by a list of English words, referring to the numbered objects in the drawing, along with an Afrikaans translation equivalent. The subtext *Where you live* is a second level complex text with its own subtexts, i.e. *Your house, In the lounge, In the dining room, In the bedroom, In the bathroom, In the kitchen, In the garden.* The majority of the subtexts of the outer text *See and say* are both subtexts and second level complex texts, cf. the text *Cuts of meat* with its subtexts *Venison, Beef, Mutton, Pork.*

Sien en sê
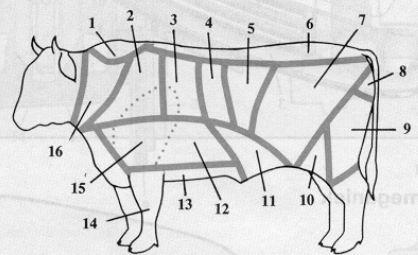
Cuts of meat

Vleissnitte

Venison (Game)   Wildsvleis

neck	1	nek
loin and saddle	2	lende en saal
leg	3	boud
flank	4	lies
shoulder	5	blad
breast	6	bors

venison chops	wildstjop
venison steak	wildskyf
jugged venison	fyngestoofde wildsvleis
roast venison	gebraaide wildsvleis/ wildsbraad
venison pie	wildspastei
venison stew	wildsbredie
jointed guinea fowl	litgedeelde tarentaal
ripen game	wildsvleis laat ryp word
game biltong	wildsbiltong

Beef   Beesvleis

hump	1	blad
chuck	2	dikrib
prime rib	3	primarib
wing rib	4	voorrib
sirloin	5	beeslende
fillet	6	beeshaas, beesfilet
rump	7	kruisstuk
aitchbone	8	stertstuk, ysbeen
silverside	9	dystuk
thick flank	10	diklies
thin flank	11	dunlies
flat rib	12	platrib
brisket	13	beesbors
shin	14	skenkel
bolo	15	bolo
neck	16	nek

mince	maalvleis, gemaalde vleis (fynvleis)
rump steak	kruisskyf
topside	binneboud

The table of contents in the primary frame of the RD does give access to the back matter texts but only to the first level of a complex text. In the table of contents an entry *See and say* guides the user to this complex back matter text but no indication is given of the fact that this text contains a variety of subtexts. Looking at the table of contents, the novice users of the RD have no idea that a systematic exposition of the lexicon of thematic fields like *Cuts of meat* or the lower-level text *Pork* are accessible to them. If they do not find these texts by chance this knowledge-enriching lexicographic bonus is lost to them. A preferred additional lexicographic procedure would have been the use of the mediostructure to include cross-reference entries in the respective central-list articles guiding the user to the outer text as reference address for additional data

and the positioning of a given word within a broader semantic field. In the previously mentioned *Wörterbuch Deutsch als Fremdsprache* (henceforth abbreviated as WDF), the back matter contains two complex texts, i.e. *Wortfelder* and *Tafeln.* Like the RD the table of contents in the front matter of the primary frame only indicates the title of these complex texts.

<div style="border:1px solid">

Inhaltsübersicht

</div>

It would have been much more user-friendly to indicate the subfields of these complex texts. The WDF does, however, improve on the system followed in the RD. In the back matter section, both these complex texts have been extended to include a table of contents.

<div style="border:1px solid">

Tafeln
(Übersicht)

</div>

These table of contents texts function as secondary outer texts and constitute a partial secondary frame structure. The subtexts of the complex text *Wortfelder* include a variety of lower level subtexts and the table of contents also list these texts as second level subtexts of the relevant first level subtexts.

<table>
<tr><td colspan="2" align="center">**Wortfelder**
(Auswahl)
auf der Basis der im Wörterbuch erscheinenden Stichwörter

Übersicht</td></tr>
<tr>
<td valign="top">

Wortfelder I: Der Mensch

1.   Körperteile/Organe (↗ *auch* 4, 5, 6)
2.   Charaktereigenschaften
3.   Fähigkeiten der Sinneswahrnehmung
3.1 Gesichtssinn
3.2 Gehörsinn
3.3 Geruchssinn
3.4 Gesmackssinn
3.5 Tastsinn
4.   Geistige Tätigkeiten
4.1 Denken
4.2 Urteilen/Diskutieren
4.3 Beweisen/Widerlegen/Zweifeln
4.4 Wissensdrang/Aufmerksamkeit/Sorgfalt
5.   Intellektuelle Fähigkeiten
6.   Psychische Prozesse
7.   Körperliche Tätigkeiten
7.1 Körperliche Ruhelage
7.2 Körperliche Bewegung/Fortbewegung
    (↗ *auch* VIII, VI.1.2)
7.3 Das Bewegen eines Objekts
7.4 Sportliche Tätigkeiten/Disziplinen/Sportgeräte

</td>
<td valign="top">

Wortfelder III: Zustandsformen

1.   Oberflächenform
2.   Nässe/Trockenheit
3.   Glätte/Stumpfheit
4.   Festigkeit/Mangelnde Festigkeit
5.   Intaktheit/Defekt

Wortfelder IV: Lageverhältnisse

1.   hoch/tief
2.   waagerecht/schräg/senkrecht
3.   seitlich/hinten/vorn

Wortfelder V: Artefakte

1.   Kleidung
2.   Gebäude
3.   Architektur (↗ *auch* II.5)
4.   Möbel/Einrichtung
5.   Geräte/Maschinen/Werkzeuge
6.   Waffen
7.   Gefäße/Behälter
8.   Produktion/Erzeugung/Ware

</td>
</tr>
</table>

The complex texts *See and say* and *Wortfelder* in the RD and the WDF respectively fulfil a knowledge-orientated function by presenting the data in such a way that the user can retrieve the lexicon of a specific semantic field. The RD goes one step further with the line-drawings representing a form of ostensive definition (cf. Gouws 1994). The lexicographic treatment performed in a complex text like the RD's *See and say* and the subject matter being treated put this text within the scope of the genuine purpose of the dictionary (cf. Gouws 2001a). Consequently *See and say* has to be regarded as an integrated outer text.

Cultural and encyclopaedic data

The frame structure plays an important role in the accommodation of cultural and encyclopaedic data (cf. Gouws 2003). It is part of the knowledge-orientated

function of a dictionary to provide this kind of data and consequently it should also prevail in the outer texts. In this regard *The Greater Dictionary of Xhosa* sets a good example with its back-matter addenda containing anthropological material. Yet again it should be emphasised that the users of a bilingual dictionary often come from different cultural and educational backgrounds. The central list should treat some of these knowledge-orientated issues by means of complex articles which contain more data than the default single article (cf. Gouws Forthcoming). The back matter texts may and should be employed to give a more comprehensive account of this material. Putting the data exposure structure to proper use could result in the back matter containing texts which guide the users to some of the complex articles where an enhanced transfer of cultural or encyclopaedic data is possible.

Cultural data in a bilingual dictionary also includes everyday lexical items like the names of the days of the week and the months of the year. Even in dictionaries where these items are not included in the central list a lexicographer will do well to include them in back matter texts dealing with this type of thematic presentation. In the biscopal *Oxford-Duden German Dictionary* with German and English as the treated language pair, a so-called cultural calendar has been included as a front matter text. This complex text has two subtexts to reflect on the traditions, festivals and holidays of the English and German communities respectively. In the central text, data of this nature can hardly be presented in a coherent way.

....

Calendar of traditions, festivals, and holidays

....

Januar

1	8	15	22	29
2	9	16	23	30
3	10	17	24	31
4	11	18	25	
5	12	19	26	
6	13	20	27	
7	14	21	28	

1 Jan – **Neujahr** New Year's Day
6 Jan – **Heilige Drei Könige, Epiphanias**
   Epiphany. *Public holiday in Austria and in parts of Germany*

Februar

1	8	15	22
2	9	16	23
3	10	17	24
4	11	18	25
5	12	19	26
6	13	20	27
7	14	21	28

2 Feb. – **Mariä Lichtmess, Lichtmess** Candlemas. In the Catholic Church, a festival marked by the blessing of candles and by candlelight processions. According to weather lore, if the bear sees its shadow when emerging from hibernation on this day – i.e. if the weather is sunny it will creep back into its den, indicating six more weeks of wintry weather
14 Feb – **Valentinstag** St. Valentine's Day

März

1	8	15	22	29
2	9	16	23	30
3	10	17	24	31
4	11	18	25	
5	12	19	26	
6	13	20	27	
7	14	21	28	

19 March – **Josephstag**
   Holiday in parts of S
25 March – **Mariä Verk**
   Annunciation

By giving this data in an outer text, the lexicographer ensures the user of a knowledge-enriching experience which falls within the scope of the kind of information a user would like to retrieve from a source of reference co-ordinating his/her native language with a foreign language.

Communication-orientated functions

The average user of a dictionary regards a dictionary as a source of linguistic knowledge. If he/she needs to know something about language this average member of a speech community will not go to the 400 shelf in a library but to THE dictionary. In a bilingual dictionary, the users want to find a target language form for a given source language item but they also want to know how and where to use that item. Not only text reception but also text production is at the core of the communication-orientated functions of a bilingual dictionary. The central list can assist the user in various ways with text production. This is primarily done by means of co-text entries and by referring the user to an integrated outer text in which grammatical patterns and typical word order issues are discussed. Here an interactive relation between central list and outer text is of extreme importance.

However, as a wide-ranging aid in text production a dictionary, especially a multifunctional bilingual dictionary, should endeavour to assist the target users in a more comprehensive way than the mere giving of translation equivalents or even the contextualisation and co-textualisation of these equivalents. Unfortunately such comprehensive assistance can hardly be given with regard to the use of individual lexical items. Looking not only at the user situation but also at the usage situation, lexicographers could try to help with some typical text production efforts relevant to everyday situations of usage. One such a situation of usage is the writing of letters, especially formal correspondence.

One of the primary back matter texts in the RD is a bilingual text *Writing a letter in Afrikaans/Die skryf van 'n brief in Engels.* Dictionary users get advice regarding the writing of letters in both the treated languages. A text in which some conventions are explained is complemented by further texts in which examples of different types of letters are given. This unintegrated complex back matter text has an undisputed communication-orientated function and complements the central text in its text production goals. Likewise the back matter text *Forms of address* assists the user and complements the communication-orientated function of the central text. Unfortunately the RD does not extend the complex text dealing with the writing of letters to include a table of contents in which the ten types of letters given in the subtexts are listed.

The *Oxford-Duden German Dictionary* also has outer texts dealing with text production. A section *German correspondence* and its equivalent section *Englische Musterbriefe* contain texts such as an invitation to a wedding, good wishes for the new year, renting a holiday house, applying for a job as au pair, resigning from a post, and cancelling a hotel booking. This complex text is extended by

the inclusion of a secondary front matter text giving a table of contents for this complex text. Another extended complex text with its own table of contents is the *Useful phrases according to function.*

Useful phrases according to function / Formulierungshilfen für verschiedene Situationen

In English and German, users are helped with phrases to express greetings, congratulations, condolences, advice, approval and disapproval, etc. Once again it is illustrated that cultural differences between different language groups imply that a phrase in one language cannot merely be translated on a word by word basis to produce a valid target language phrase.

For the user to utilise the outer texts of a dictionary to the full he/she should be familiar with the lexicographic functions decided on by the lexicographer. The same functions should also prevail in the outer texts. A transtextual functional approach is needed.

In conclusion

The proper use of outer texts in bilingual dictionaries can improve the quality of the information transfer. However, the construction of outer texts must be done in a scientific way and the lexicographers should be familiar with the

underlying theoretical motivations. A dictionary culture should be established where dictionary users know that outer texts are just as much part of the dictionary as the translation equivalents given for a source language lemma. That may lead to a situation where the dictionary user while consulting the central list will realise: the truth is out there.

Bibliography

Dictionaries

Bosman, D.B. et al. (Eds.). 1984[8]. *Tweetalige Woordeboek/Bilingual Dictionary.* Cape Town: Tafelberg.

Dunkling, L. and W. Gosling. 1983. *Everyman's Dictionary of First Names.* London: J.M. Dent and Sons.

Grobbelaar, Peter et al. (Eds.). 1987. *Reader's Digest Afrikaans–Engelse Woordeboek/English–Afrikaans Dictionary.* Cape Town: The Reader's Digest Association, South Africa (Pty) Ltd.

Kempcke, Günter et al. 2000. *Wörterbuch Deutsch als Fremdsprache.* Berlin/New York: Walter de Gruyter.

Lysle, A. de R. 1928. *Nuovo Dizionario.* Torino: Libreria Editrice F. Casanova and C.

Martin, W. et al. (Eds.). 1989. *Van Dale Groot woordenboek Engels–Nederlands.* Utrecht: Van Dale.

Martin, W. et al. (Eds.). 1989. *Van Dale Groot woordenboek Nederlands–Engels.* Utrecht: Van Dale.

Pahl, H.W. (Ed.). 1989. *The Greater Dictionary of Xhosa.* Alice: University of Fort Hare.

Procter, P. (Ed.). 1978. *Longman Dictionary of Contemporary English.* Harlow: Longman.

Scholze-Stubenrecht, W. and J.B. Sykes (Eds.). 1999[2]. *Oxford-Duden German Dictionary.* Oxford: Oxford University Press.

Other literature

Bergenholtz, H. and S. Tarp. 2003. Two Opposing Theories: On Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 171-196.

Bergenholtz, H., S. Tarp and H.E. Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Hoffmann, L. et al. (Eds.). 1999. *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research*: 1762–1832. Berlin: De Gruyter.

Gallardo, A. 1980. Dictionaries and the Standardization Process. Zgusta, L. (Ed.). 1980. *Theory and Method in Lexicography*: 59-69. Columbia: Hornbeam Press.

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

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

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

Gouws, R.H. 2000. Strategies in Equivalent Discrimination. Mogensen, J.E. et al. (Eds.). 2000. *Symposium on Lexicography IX*: 99-111. Tübingen: Max Niemeyer.

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

Gouws, R.H. 2001a. Lexicographic Training: Approaches and Topics. Emejelu, J.D. (Ed.). *Éléments de lexicographie gabonaise. Tome I*: 58-94. New York: Jimacs-Hillman.

Gouws, R.H. 2002. The Outer Texts in the *De Gruyter Wörterbuch Deutsch als Fremdsprache*. Wiegand, H.E. (Ed.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II*: 471-484. Tübingen: Max Niemeyer.

Gouws, R.H. 2003. Using a Frame Structure to Accommodate Cultural Data. Emejulu, J.D. (Ed.). *Éléments de lexicographie gabonaise. Tome II*: 54-69. New York: Jimacs-Hillman.

Gouws, R.H. Forthcoming. Issues Regarding the Comment on Semantics in Bilingual Dictionaries Dealing with Closely Related Languages. Petkov, P. (Ed.). Forthcoming. *Proceedings: Internationales Kolloquium zur Wörterbuchforschung*.

Gouws, R.H. and S. Tarp. 2004. Wie leer wat uit Afrikaanse (aan)leerderwoordeboeke? *Tydskrif vir Geesteswetenskappe*.

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

Hausmann, F.J. and R.O. Werner. 1991. Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher. Hausmann, F.J. et al. (Eds.). 1989–1991: 2729-2769.

Hausmann, F.J. 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Max Niemeyer.

Hausmann, F.J. et al. (Eds.). 1989–1991. *Wörterbücher: Ein internationales Handbuch zur Lexikographie/Dictionaries: An International Encyclopedia of Lexicography/Dictionnaires: Encyclopédie internationale de lexicographie*. Berlin/New York: Walter de Gruyter.

Herberg, Dieter. 1998. Die Außentexte in *Langenscheidts Grosswörterbuch Deutsch als Fremdsprache*. Wiegand, H.E. (Ed.). 1998. *Perspektiven der pädagogischen Lexikographie des Deutschen I*: 331-341. Tübingen: Max Niemeyer.

Kammerer, Matthias and Herbert Ernst Wiegand. 1998. Über die textuelle Rahmenstruktur von Printwörterbüchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-238.

Kromann, H.-P., T. Riiber and P. Rosbach. 1984. 'Active' and 'Passive' Bilingual Dictionaries: The Ščerba Concept Reconsidered. Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings*: 207-215. Tübingen: Max Niemeyer.

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

Wiegand, H.E. 1998. Lexikographische Textverdichtung. Entwurf zu einer vollständigen Konzeption. Zettersten, A., V.H. Pedersen and J.E. Mogensen (Eds.). 1998. *Symposium on Lexicography VIII*: 1-35. Tübingen: Max Niemeyer.

Wiegand, H.E. 1996. Das Konzept der semiintegrierten Mikrostrukturen. Ein Beitrag zur Theorie zweisprachiger Printwörterbücher. Wiegand, H.E. (Ed.). 1996. *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger lexikographischen Kolloquium*: 1-82. Tübingen: Max Niemeyer.

Improving Dictionary Skills in Ndebele

Samukele Hadebe, *Department of African Languages and Literature, University of Zimbabwe, Harare, Zimbabwe (samukeleh@yahoo.co.uk)*

Abstract: This article proposes ways of improving dictionary skills amongst the Ndebele. One way of accomplishing this is incorporating the teaching of dictionary skills into teacher training syllabi. Teachers can impart their knowledge to students and a dictionary culture can develop for enhancing effective use of current dictionaries and helping in the compilation of future ones. In general lexicographers are aiming at producing dictionaries that are 'user-friendly', that is, accessible to their users. Current and future Ndebele dictionaries are likely to remain underutilised unless users develop requisite skills for dictionary use. This article outlines some aspects of dictionary studies that could be incorporated in the training syllabi for potential Ndebele language teachers. Anticipated problems of implementation and possible solutions to these problems are also discussed.

Keywords: DICTIONARY, DICTIONARY CULTURE, DICTIONARY SKILLS, REFERENCE NEEDS, REFERENCE SKILLS, USERS, USER-FRIENDLY, NDEBELE, LEXICOGRAPHY, CHANGE OF ATTITUDE

Opsomming: Verbetering van woordeboekvaardighede in Ndebele. Hierdie artikel stel maniere voor vir die verbetering van woordeboekvaardighede onder die Ndebeles. Een manier om dit te bewerkstellig is om die onderrig van woordeboekvaardighede in die opleidingsillabusse vir onderwysrs in te sluit. Onderwysers kan hul kennis aan studente oordra, en 'n woordeboekkultuur kan ontwikkel om die doeltreffende gebruik van huidige woordeboeke te verbeter en by die samestelling van toekomstiges te help. In die algemeen streef leksikograwe daarna om woordeboeke te lewer wat "gebruikersvriendelik", d.w.s., toeganklik vir hul gebruikers is. Huidige en toekomstige Ndebelewoordeboeke sal waarskynlik onderbenut bly mits gebruikers die nodige vaardighede vir woordeboekgebruik ontwikkel. Hierdie artikel skets 'n aantal aspekte van woordeboekstudie wat in die opleidingsillabusse vir toekomstige Ndebeletaalonderwysers ingesluit kan word. Verwagte probleme by die instelling en moontlike oplossings vir hierdie probleme word ook bespreek.

Sleutelwoorde: WOORDEBOEK, WOORDEBOEKKULTUUR, WOORDEBOEKVAARDIG-HEDE, NASLAANBEHOEFTES, NASLAANVAARDIGHEDE, GEBRUIKERS, GEBRUIKERS-VRIENDELIK, NDEBELE, LEKSIKOGRAFIE, VERANDERING VAN HOUDING

1.   Introduction

This article deals with the teaching of dictionary-using skills with the long-term aim of developing a dictionary culture amongst the Ndebele[1] of Zimbabwe. By *dictionary culture* is meant, according to Hartmann and James (1998: 41): 'The

critical awareness of the value and limitations of dictionaries and other refer-ence works in a particular community.' The focus is on *user perspective* which investigates the reference skills and reference needs of dictionary users (Hart-mann 2000: 11, 2001: 81, Béjoint 2000: 140, Svensén 1993: 13). User perspective in lexicography is a user-driven approach to dictionary-making and dictionary research. *Reference skills* are 'the abilities required on the part of the dictionary user to find the information being sought' (Hartmann and James 1998: 117), while *reference needs* are 'the circumstances that drive individuals to seek infor-mation in reference works such as dictionaries' (Hartmann and James 1998: 116).

Research on lexicography in the Ndebele language is still in its early stages. Actually, no research on Ndebele lexicography had been done until the establishment of the African Languages Lexical Project (ALLEX[2]), now the African Languages Research Institute (ALRI), at the University of Zimbabwe. Researchers in Ndebele lexicography are taking into cognisance the needs of dictionary users and the user perspective in dictionary studies.

Lexicographers in general are attempting to produce dictionaries that are 'user-friendly', that is, accessible to their users. Similarly, publishers of diction-aries also want products that are attractive and appealing to their markets. Thus the designation 'user-friendly' has become a catchword for lexicogra-phers, publishers and critics alike. Béjoint (2000: 140) states:

> Lexicographers in many countries have recently felt the need to go beyond empirical observations on the use of the general-purpose dictionary, seeking to find out what the users really do, as opposed to what they are believed to do, in order to make sure that the dictionary really corresponds to the needs of the public.

Similarly, lexicographers compiling Ndebele-language dictionaries, especially at ALRI, are concerned with fulfilling the users' expectations and needs. What has now been described as user needs in Ndebele, are in fact assumptions by dictionary editors. User-friendliness is difficult to attain, let alone to measure in a community where reference skills are relatively low and dictionary culture is non-existent.

As long as Ndebele dictionary users lack the requisite skills, current and possible future dictionaries in the language are likely to be underutilised until users have developed appropriate reference skills. There is a need to develop dictionary skills among the Ndebele to higher levels than those existing at pre-sent. Only after the community has a certain level of competence in dictionary use, can there be talk of user-friendliness and user needs. It is important, there-fore, to establish user needs and user perspectives in Ndebele if the production of dictionaries were going to be of use. Such skills cannot be developed fully through informal learning and individual experience. They have to be taught formally. Dictionary makers have to know, and not only assume what they think users know.

Writing on the significance of teaching dictionary skills in pedagogical lexicography, Dolezal and McCreary (1999: IX) state: 'As we look to the future development of this area, the mutual concerns of the teacher, the lexicographer and the learner/dictionary user could be the focus of pedagogical lexicography.'

The production of dictionaries is expensive, time-consuming and labour-intensive. When these factors are taken into consideration, it would be uneconomic to continue producing reference works if such works will never be fully utilised beyond their function as cultural artefacts. Therefore, there is a need for user education in Ndebele, especially for teacher training. *User education* is defined by Hartmann and James (1998: 152) as 'the training of users in the reference skills in response to reference needs'. By training teachers to be good users themselves, it is envisaged that they could pass these skills on to students and eventually to society in general.

2.   The importance of teaching dictionary skills

In Zimbabwe, books in indigenous languages are produced for the school market, and dictionaries are no exception. The development of a dictionary culture would be enhanced in the Ndebele community, especially in pupils through targeting teachers first. However, students cannot acquire dictionary skills by themselves without assistance from teachers. The initial focus should then be on teacher training. Hartmann (2001: 26) says: 'Teachers have rarely been offered the training needed to judge the benefits and limitations of particular dictionaries, or to instruct their students in the reference skills appropriate to guarantee successful consultation. Hardly ever are they in touch with lexicographers or dictionary researchers.' The problem of the lack in the teaching of reference skills is not peculiar to Ndebele, but is common in most languages. Landau (2001: 26) notes: 'The habit of using a dictionary is formed early in life, and if the skills to use it are neglected, the student may never be comfortable using dictionaries.' This observation is true for Ndebele dictionary users as well.

Writing on the possible solutions to this problem in general, Dolezal and McCreary (1999: XVIII) identified the following:

(a)   make a more so-called 'user-friendly' dictionary (which in positive connotations must mean to develop a more readable format, better examples, better organization, etc., and in negative connotations must mean to dumb-it-down); and

(b)   teach students reference skills.

As already mentioned above, the effectiveness of a user-friendly dictionary is partly dependent on the reference skills of the user (Svensén 1993: 16). Therefore, point (a) in the above quotation becomes secondary to point (b), which

advocates the teaching of reference skills to students. The need to teach students about dictionaries is also noted by McKean (2000: 82): 'School dictionary lexicographers have been trying to encourage the teaching of dictionary activities and dictionary use in the classroom.'

To improve the lexicographic situation in Zimbabwe in general and Ndebele in particular, the teaching of reference skills should first target the teachers themselves. Teachers should be taught reference skills as part of their training in order to enable them to make maximum use of dictionary and reference books and impart this knowledge to their students. Teachers are in a better position than other researchers or even lexicographers to assess the students' lexicographic needs. If teachers gain competence in dictionary skills and pass on the skills to students, a dictionary culture would grow in the community.

Teachers and pupils, being the main users of dictionaries in the Ndebele community, can in the long run give a better evaluation of existing dictionaries and present views on shortcomings and needs in Ndebele lexicography. McKean (2000: 87) contends: 'The biggest area of improvement does not seem to be in the dictionary, but teacher training.' It is through improving the teachers' competence that dictionary culture could be taught to students and in turn produce more informed dictionary users in general. McKean further states: 'Better teacher training would, in time, lead to better dictionary consumers — a boon, not only for makers of school dictionaries, but for makers of commercial adult dictionaries as well.' Therefore, the choice of focusing on teacher training is justified.

The majority of Ndebele teachers live and work in the countryside and the few in urban areas are mostly in the poor schools in high-density suburbs. The common condition of schools in these circumstances is the lack of textbooks and other essential teaching aids. There are schools with no library and thousands others far from acquiring their first computer. In Zimbabwe there is still a partiality for science subjects, which are considered more important than arts subjects. This partiality has implications for resource allocation. The little resources a school is allocated firstly goes to purchasing science books and materials considered as a priority by both parents and educationists. When it comes to the arts subjects, English is given top priority both in terms of staffing, time and resource allocation. Therefore, in this general scarcity of resources in schools, the situation of the teacher of Ndebele is even more critical.

Teacher training in Zimbabwean colleges can be considered good by Third World standards. Teacher training colleges are satisfactorily equipped and staffed to produce on the average well-trained teachers. This is, with reference to the conventional colleges for secondary school teachers, the focus of this discussion. If he/she is equipped with the necessary reference skills and knowledge of the school conditions in the country, the teacher can use the little resources at his/her disposal in the best possible way. These could be a few textbooks and a handful of copies of the Ndebele dictionary.

Murphy (2000: 80) poses the questions: 'How can dictionary skills be

applied and encouraged in grades in which "dictionary skills" are not listed in the official curriculum? … How can we lexicographers reach out to teachers, teachers-in-training, and teacher trainers?' The same questions can be asked about Ndebele and the education system in Zimbabwe. The language syllabi of the two main teacher training colleges in the Ndebele-speaking provinces of Zimbabwe do not include dictionary skills among the skills taught. The Ndebele syllabi of the Hillside Teachers' College and United College of Education for 2002 both do not mention the teaching of reference skills[3]. Similarly, the school curricula are silent upon dictionary skills, probably assuming that these are acquired along the way.

In this article, it will be shown that there is no need specifically for separate courses on dictionary skills. The teaching of dictionary skills can be incorporated into courses on research skills and the use of learning aids, which are already a core part of the teacher training syllabi for Ndebele language teachers. As part of the research skills taught to student teachers in general, several aspects of lexicography could be included to complement these.

3.   Teaching aspects of lexicography

The following areas can be proposed: general lexicography, history of the lexicography of Nguni[4] and the Ndebele language, dictionary typology, lemmatisation in Nguni and Ndebele and dictionary structure. Each of these areas will now be discussed, with an elaboration of what each could cover.

3.1   Lexicography in general

There is a need to initiate teachers into the use of reference books of various types: dictionaries, thesauri, encyclopaedias and others. As there are not many types of reference works in Ndebele or, for that matter, in any African language, studying some aspects of English lexicography as reference point is inevitable. Dictionaries in the Ndebele language are, in any case, largely relying on general trends and traditions acquired from English language lexicography. Some problems of dictionary use in English could be highlighted in order to prepare the ground for a discussion of problems likely to face the Ndebele language student.

As stated earlier, teaching dictionary skills is part of language teaching. Alphabetic ordering in dictionaries should be discussed as well as its purported advantages and disadvantages. It is an important element of most dictionaries. This would also furnish a background for the discussion on lemmatisation in Nguni touched upon below. Teachers should be taught that, apart from the alphabetic system, other methods of ordering, for instance thematic ordering, are also used.

3.2   History of the lexicography of Nguni and the Ndebele language

The first dictionaries and wordlists in Ndebele and other Nguni languages, especially Zulu, were bilingual, mainly targeted at missionaries who wanted to learn African languages. In these, there was a partiality towards Christian religion: 'Pelling's *A Practical Ndebele Dictionary*, has most of the terms relating to Christian religion and hardly any on African religion' (Hadebe 2000: 226). A similar bias is observed in Shona in Zimbabwe (Mpofu 2001: 244) and in Fan in Gabon (Mavoungou 2001: 123). In course of time, the bilingual dictionaries were also meant to assist Ndebele people learning English.

Compared to bilingual dictionaries, monolingual dictionaries have different issues to address. As far as the Ndebele language is concerned, various aspects can be learnt from such dictionaries. The role of a general-purpose monolingual dictionary for the native speaker is significant. The compilation of monolingual dictionaries in African languages is itself an important indicator of the status of these particular languages in society. The compilation of the first monolingual Ndebele dictionary had a number of interesting implications for the Ndebele language as far as its standardisation was concerned. For instance, a number of orthographic problems were addressed in the process of making this dictionary. Some of these include the division of compound words and the spelling of loan-words.

A historical background of Nguni lexicography in general and of Ndebele in particular will be useful for teachers. It would enhance their understanding of different types of dictionaries and why a particular type dominates the Ndebele and the Nguni languages. It could further be useful in their understanding of dictionary typology and dictionary structure. It could also give a background to the understanding of particular problems in Ndebele and Nguni lexicography, for instance the lack of consistency in lemmatisation and alphabetic ordering (see 3.5 below).

3.3   Dictionary typology

Teachers have to know the differences between types of dictionaries. This is important not only to guide students in using them appropriately, but also to advise their schools in purchasing the right kind of dictionaries. Knowledge of dictionary types enables the user to know what to expect in a particular type of dictionary. Knowledge of dictionary typology may assist in the understanding of dictionary structure.

Unless this knowledge is conveyed to teachers to pass on to students, dictionary users will continue to think that all dictionaries are the same and that anything can be looked up in any dictionary. Dolezal and McCreary (1999: XIII) state:

> Results of empirical studies suggest that dictionary users do not distinguish among types of dictionaries; moreover, no matter how lexicographers might

> classify a 'thesaurus', students do not necessarily consider a thesaurus as a text belonging to a separate category of reference book.

As is shown in this quotation, many users of dictionaries, even of established languages like English, still have problems with distinguishing types of dictionaries. Acquisition of knowledge about basic distinctions in dictionaries could greatly enhance the use of dictionaries both as teaching and learning resources and as general reference works. Beyond establishing that there are dictionaries of different types, students in Ndebele are unlikely to master these basic distinctions in dictionary types by themselves. They will have to be taught about these differences.

3.4   Dictionary structure

The ability to use any dictionary, no matter how casually, presupposes and requires some basic knowledge of the structure of dictionaries. One such basic aspect, for instance, is the alphabetic ordering. Gouws (2001: 110) stresses the user-perspective in lexicography:

> Modern-day lexicography is dominated by the user-perspective and the access structure of any new dictionary has to be user-driven, i.e. it has to be planned in accordance with the user profile of the dictionary and the needs and reference skills of a well-defined target user.

A comparative study of Ndebele and other Nguni language dictionaries focusing on dictionary structure could be relevant here. Although there is a generally expected structure in every dictionary, each dictionary has its own peculiar structure. Through the study of dictionary structure, teachers can learn how to use the dictionary as a tool for teaching in the classroom. The following areas could be helpful to a Ndebele language teacher:

3.4.1   Grammatical information

For Ndebele, having no comprehensive grammar book, the dictionary, for instance the recently published *Isichazamazwi SesiNdebele* (2001), could be useful for studying grammar. Information on aspects of grammar is usually given in the front matter of the dictionary. In the *Isichazamazwi SesiNdebele,* for instance, apart from a brief grammatical outline of the language in the front matter, the word category is given for each entry. Entries are marked as verbs, nouns, adjectives, adverbs, ideophones, pronouns, copulatives and conjunctives. Noun classes[5] are also indicated while for verbs it is stated that they are either transitive or intransitive. Different types of morphemes, like tense morphemes, aspect morphemes, suffixes and reflexive morphemes are marked and defined. This is very useful information for language teachers.

3.4.2 Spelling

Any general-purpose standard dictionary is assumed to reflect standard spelling. Teachers have to insist on their students spelling correctly, and the use of dictionaries is important for checking spelling. Ndebele still has unsolved spelling problems that might call for adjustment and reform. It is crucial for language users to be familiar with spelling conventions in the language, especially the division of compound words and the spelling of loan-words.

3.4.3 Pronunciation

Along with spelling goes the importance of pronunciation in dictionaries. Speakers need to spell words correctly and also pronounce them correctly. Most dictionaries in African languages mark tone because varying the tone changes the meaning of a particular word.

3.4.4 Meaning

Studies in general have shown that meaning is the first information users seek in a dictionary. The same assumedly applies to Ndebele users. Teachers have to be familiarised with types of meanings and the various possible senses a word may have. To save space, most senses are usually excluded from a general-purpose dictionary. The use of synonyms in dictionaries should be taught as well.

Meaning is usually found through the context in which the particular word is used. Dictionary users have to be familiar with the description of meanings and how to use dictionaries in finding the meanings of words and expressions as they occur in texts. Commenting on dictionary meanings in general, Moon (1987: 86) states: 'Dictionaries traditionally record vocabulary as if meaning is something independent, inherent and unique to an item, and serving to distinguish it from all others.' Teaching some basics about types of meanings would be adequate for dictionary users to enable them to make full use of dictionaries as guides in deriving meanings of words and expressions in their various contexts.

3.4.5 Vocabulary

Vocabulary acquisition is important in all language teaching. Teachers have to help their students acquire as wide a vocabulary as possible. Students with a good command of their language would, for instance, reflect it in their use of synonyms and antonyms. The dictionary is one of the more reliable sources for enhancing vocabulary growth in students. For teaching Ndebele to non mother tongue speakers the appropriate vocabulary is essential. A significant number of both teachers and students come from minority language communities and learn and speak Ndebele at school[6].

In addition to teaching aspects like meaning or vocabulary, teachers need to cultivate in their students the ability to use the dictionary for more than looking up meaning, spelling and pronunciation. Teachers have to learn and teach the importance of the front matter and the back matter. It is a commonly acknowledged fact that users do not read instructions given in the front and back matter of dictionaries. This habit should be changed if users are to find dictionaries fully informative. Different ways of accessing information in dictionaries should be taught to teachers so that such knowledge could be conveyed to students at an early age.

As noted above, the objective to produce more user-friendly dictionaries has led lexicographers to make various innovations which often go unnoticed except perhaps by reviewers. Dolezal and McCreary (1999: XVIII) write:

> Lexicographers painstakingly introduce more discrete definitions, examples, and grammatical information: many critiques of their work call for yet more nuanced information …, while the learner, according to some studies, doggedly resists being a user of any type of dictionary.

Therefore, the user perspective in lexicography cannot attain its objective unless the user is efficient in dictionary use. For instance, children's dictionaries usually have a large font, coloured pictorial illustrations and all techniques designed for easy access of information. All this could, however, be a wasted effort on the part of the lexicographer and the publisher unless children have basic reference skills.

3.5   Lemmatisation and alphabetic ordering in Nguni and Ndebele

It is important to prepare teachers for potential problems students in Ndebele might face in using Nguni language dictionaries. There is the lemmatisation problem, inconsistencies in the alphabetic ordering of nouns by stem, prefix proper or initial vowel. For the lemmatisation of verbs and adjectives, the alphabetic ordering is based on the first letter of their respective stems.

The particular emphasis on lemmatisation problems in Ndebele should complement the study of dictionary structure. Lemmatisation by noun stem obviously makes access to entries more complicated than lemmatisation based on the initial letter of the noun, which in Ndebele is always a vowel. Traditionally, nouns in Nguni have largely been lemmatised according to noun stems. This presupposes users to have a certain level of understanding of grammar and to know the morphology of every word to be able to identify the stems.

3.6   A case study of a monolingual Ndebele dictionary

Teachers, as appliers of dictionaries, should be able to evaluate the effectiveness of dictionaries in fulfilling the objectives they purport to achieve. This is

the essential part teachers have to play in assessing the usefulness of a dictionary for students and the best way it could be improved. By understanding the dictionary structure, teachers could easily assess students' needs and students' inadequacies in dictionary skills. This knowledge would be valuable for lexicographers and other researchers in language teaching, textbook writing and teaching methods.

A monolingual Ndebele dictionary, for instance *Isichazamazwi SesiNdebele*, could be used for a case study in dictionary criticism. This would enable the students to learn what is contained in the dictionary and what else they think should have been included. The dictionary criticism should broadly cover the text structure, macrostructure, microstructure, mediostructure and access structure.

4.   Anticipated problems in implementation

Notwithstanding the expected advantages of teaching reference skills, there are some foreseeable problems. Some of the anticipated problems are societal attitudes, lack of competence and lack of training of lecturers.

4.1   Attitude problems

Attitude problems have a long history. Some sections of the population, especially the middle class, have negative attitudes towards all African languages in general. These people send their children to schools where African languages are not taught or are taught as optional subjects. Their negative attitude towards their mother tongues is likely to be transferred to dictionaries in these languages as well. Although this is a relatively small minority, they are the most influential people in the community. They tend to be the better educated and the well employed. Due to the colonially designed education system, these people still believe, falsely of course, that English or other languages of wider communication are the only languages worth learning. They want their children to start primary school in English, which is their second language.

Such negative attitudes held by people in strategic positions in society like teachers, doctors, education officers, and the like have implications for the money spent on dictionaries. In Zimbabwean schools, there is a tendency to give Ndebele as a subject fewer hours than subjects like English, Mathematics or Science. Similarly, the money schools spend on books is unequal, with Ndebele usually getting the least. This may mean that less reference books and dictionaries in Ndebele are bought for schools. In fact, even the staffing of Ndebele as a subject is at a disadvantage. Teachers trained to teach Ndebele are expected to teach other subjects like Mathematics and English when there is a general shortage of staff in the country. Ndebele is then given to untrained

teachers because of the false belief that any mother tongue speaker having passed the school leaving examinations is competent enough to teach it. There are cases of teachers trained and qualified as language teachers who have never taught the subject for the whole of their careers. Therefore, some people with a knowledge of Ndebele lexicography might never pass on that knowledge to students, as they might be teaching other subjects like Mathematics for instance.

These negative attitudes ultimately lead to less investment of resources in Ndebele-language dictionaries by both researchers and publishers. Parents would buy dictionaries for their children if they think the language is an important school subject. The negative attitudes towards the language might cause few parents to purchase dictionaries. This has double implications. Firstly, there would be few users with access to dictionaries and that affects negatively on enhancing dictionary skills. Secondly, if less dictionaries are bought, publishers would in future shun publishing dictionaries. This also has a negative impact on the growth of Ndebele lexicography and consequently dictionary culture.

4.2   Problems of incompetence

There is a general dislike of Ndebele grammar as a subject amongst many students. Students at both secondary and tertiary colleges prefer literature and poetry to grammar. The problem of incompetence is further compounded by some teachers who even avoid teaching grammar at all. A number of reasons could be advanced for this situation, which of course is not peculiar to Ndebele. Firstly, Ndebele lacks grammar textbooks. Zulu textbooks like *Uhlelo lwesiZulu* are used, and some students have problems with the Zulu terminology. Secondly, most teachers are not confident in handling grammar because they were not properly taught themselves.

4.3   Lack of teacher trainers

Lastly, the question could be posed: Who will teach the teachers? There is a big possibility that even lecturers in teachers' colleges and universities in Zimbabwe have not themselves acquired these reference skills adequately.

Change, even if it is for the good, is sometimes resisted because people are not keen to abandon what they are used to. As the syllabi in teachers' colleges are an internal college or departmental affair, there can be no way of enforcing or ensuring that the appropriate skills are taught at all. One lecturer might teach it one year, and the following year another might abandon it. It therefore wholly depends on people on the ground realising the present deficiency and taking responsibility for improving it.

5.   Possible solutions to problems of implementation

5.1   Changing attitudes

Changing of attitude should be focused on specific groups within the community, that is, politicians, teachers, parents and students. Each group has to be approached with a different strategy for changing its attitude and the priority groups need attention first. First the politicians who are responsible for language and educational policies of the country should be won over. Once the policies are favourable to indigenous languages in education generally, those charged with implementation should be convinced. The desired policy changes should be to give Ndebele a meaningful role in education, preferably as a language of instruction for some subjects. It must be made mandatory that only appropriately qualified teachers should teach the language. Also, it should be allotted adequate time in the school timetables just like English or Mathematics.

Teachers and other professional people need evidence based on research to convince them of the importance of mother tongue instruction and consequently the acquisition of dictionary skills in Ndebele. A pilot study could perhaps be made with two or three different secondary schools. In each, two classes on the same level could be used for experimentation. One class is taught in Ndebele and also instructed to acquire dictionary skills, while the other is left with the current teaching methods and content. Then, at the end of a term, a test should be given to assess whether those who received instruction in dictionary skills would perform better than the control group. If the results of the experiment show (which it presumably will) that those students who received instruction in Ndebele and dictionary skills tend to perform better, then teachers will be motivated to adopt this method too. The good results obtained from adopting the new strategy may engender positive attitudes in teachers towards Ndebele in education. Parents generally want the best for their children and the best is what enables them to pass in school. With teachers convinced and their attitudes changed, parents would easily follow, especially if the government policy is already favourable. Students would consequently acquire a positive attitude towards Ndebele from their teachers, parents and from society in general.

5.2   Addressing problems of incompetence

The level of competence in grammar has generally to be raised in teachers as well as in students if they are to find using dictionaries easy and profitable. As noted earlier, the problem is also found elsewhere, for instance amongst English-language students as noted by Dolezal and McCreary (1999: XVII):

> The unfortunate lack of language analysis skills among the current student population can only hinder their ability to comprehend and read dictionaries as

reference texts: students who do not know the basic terminology of traditional or descriptive grammars will be hard pressed to decipher even the newest and best formatted information that is founded upon the latest theories of reading comprehension, grammatical theory and lexicology.

As mentioned earlier, most dictionaries teachers and students have been using prior to the publication of the monolingual Ndebele dictionary have noun entries entered by the first letter of the noun stem. To access a word implies to know the morphological structure of that word. Furthermore, there is no consistency in the dictionaries on the structure of certain nouns, especially loanwords. The metalanguage used in the Ndebele and other Nguni dictionaries is a language peculiar to grammar studies. If students cannot identify word categories correctly, the dictionary would be of little use as a language-learning tool. Therefore, the general lack of competence in grammar has to be changed, especially of language teachers.

5.3   Addressing the problem of teacher trainers

In colleges and universities, lecturers in Ndebele are still only a small number, not exceeding thirty at most. These could take part in a workshop conducted by lexicographers from ALRI who have received training and have experience in practical dictionary-making. That could be the starting point. In fact, these college and university lecturers have been invited periodically to ALRI seminars. There is a good working relationship already between the colleges and ALRI. Publishers could jointly fund the seminar as they too have always participated in these ALRI annual seminars.

6.   Advantages for Ndebele lexicography

If the above proposals could be incorporated in the training syllabus for language teachers, a number of possible outcomes can be expected. To start with, teachers would be more knowledgeable and confident in their use of reference works. This ability would enable them not only to use dictionaries effectively in their teaching but also to teach reference skills to students. Once students and teachers have acquired reference skills and developed positive attitudes towards dictionaries, a dictionary culture would be growing amongst the Ndebele. As noted earlier, the Ndebele community, like most African communities, is generally poor and schools have very limited resources. Those limited resources should be used to the maximum. At present, with poorly developed research skills even amongst teachers, students are not putting the available resources to full use.

For the foreseeable future, dictionaries in Ndebele will be mainly for students or would rather be mainly used by students. In Zimbabwe, the general

trend in publishing in African languages is that the school is the targeted market. This does not mean that dictionaries in Ndebele are aimed at a limited section of society. Actually, the potential market for dictionaries could be considered from primary school pupils to university students. Teachers and lecturers of all these various levels of pupils and students are similarly potential regular users of dictionaries. Once pupils and students develop the habit of using dictionaries, parents would have to acquire dictionary skills too, as they would be required to assist their children with homework.

Creative writers, who target the school market, would likewise have to rely on dictionaries for standard vocabulary and spelling, and so would textbook writers and publishers in general. Should there be such developments, interest in research on dictionaries is bound to grow.

Teachers would be interested in better dictionaries for teaching purposes, and textbook writers would be compelled to produce better works too. Publishers generally want to keep their market satisfied with the best that can possibly be produced. All these factors would combine to enhance research on dictionary-making in Ndebele.

At the moment, lexicographers and publishers are guessing what Ndebele dictionary users want. Ndebele dictionaries compiled in the past and being compiled at present, have not been and are not based on prior research on user needs and reference skills. Such research to establish user needs would not yield much, considering the low standard of reference skills amongst speakers. Once teachers have a basic knowledge of dictionaries and how they can best be utilised, there would be users able to give reliable feedback about particular dictionaries and user needs. At the moment, there is no point in finding out the extent of the development of reference skills in Ndebele because these skills are not being taught at all. Once they begin to be taught, research to evaluate the appropriateness of what is taught would be relevant.

The publication of the first monolingual Ndebele dictionary in 2001 was greeted with great enthusiasm, especially in Bulawayo. For many Ndebele-speaking people, the publication of the dictionary marked the revival of their language and culture from what is perceived as cultural oppression. The dictionary is considered with national pride. However, dictionaries must go beyond being just artefacts and symbols of national pride; they must become tools to improve communication.

Teaching of the relevant dictionary skills should develop a dictionary culture amongst the Ndebele. The conditions for such a development are favourable, judging from the enthusiasm people showed for the dictionary in newspaper articles, letters to editors and phone-ins on the radio after its launch. The linguistic climate[7] is now conducive to decisive changes in the teaching and promotion of Ndebele in Zimbabwe. Among potential users the motivation is already there; what now remains is to give them the requisite skills to make full use of dictionaries.

7.   Conclusion

It is possible to improve the reference skills of Ndebele dictionary users through teaching the appropriate dictionary skills. By focusing firstly on teacher-training, this knowledge would easily reach the school pupils and eventually the general Ndebele community. In this way, a dictionary culture could be enhanced which in turn would be instrumental in directing future lexicographic work based on the actual reference needs of dictionary users. Once teachers have acquired the appropriate skills and the teaching of reference skills is incorporated in the school syllabus, other additional means of enhancing dictionary skills could be employed as well. Some of these could be seminars and workshops on dictionaries for teachers, writers and other dictionary users. Even policy decisions could be made from an informed position, as most stakeholders would be more knowledgeable on dictionaries than they are currently. User-friendly dictionaries in Ndebele can be compiled when reference needs and reference skills of dictionary users are known.

Acknowledgements

Endnotes

1.   This is the Ndebele language spoken in Zimbabwe as opposed to that spoken in the Republic of South Africa.

2.   The ALLEX Project is a lexical project of cooperation involving the University of Zimbabwe, the University of Oslo and the University of Gothenburg, funded by NUFU (the Norwegian Universities' Committee for Development, Research and Education).

3.   Hillside Teachers' College trains secondary school teachers, while the United College of Education is responsible for the training of primary school teachers.

4.   Nguni refers to the sister languages Ndebele, Zulu, Xhosa and Swati. As these languages share a high percentage of their lexicon, speakers can refer to any of the language dictionaries, Ndebele for instance having used Zulu dictionaries.

5.   Ndebele nouns fall roughly within classes 1 to 18, following the noun classification of Bantu languages by Carl Meinhof.
6.   Ndebele is the officially sanctioned national language in the Matabeleland provinces of Zimbabwe where about five minority languages are also spoken. Speakers of these minorities learn Ndebele at school as if it were their first language and have difficulties with the vocabulary and pronunciation of Ndebele.
7.   According to Vikør (1993: 179), *linguistic climate* refers to 'the ideological climate with regard to language use, language correctness and language planning'.

Bibliography

Béjoint, H. 2000. *Modern Lexicography*. New York: Oxford University Press.

Dolezal, F.T. and D.R. McCreary. 1999. *Pedagogical Lexicography Today*. Tübingen: Max Niemeyer.

Gouws, R.H. 2001. The Use of an Improved Access Structure in Dictionaries. *Lexikos* 11: 101-111.

Hadebe, S. 2000. Developing Terminology in African Languages with Special Reference to Ndebele. Chiwome, E.M. et al. (Eds.). 2000. *Indigenous Knowledge and Technology in African and Diasporan Communities*: 225-231. Harare: Mond Publishers.

Hadebe, S. et al. (Eds.). 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.

Hartmann, R.R.K. 2000. European Lexicography: Perspectives on Dictionary Research, with Special Reference to the Countries of the European Union. *Dictionaries* 21: 1-21.

Hartmann, R.R.K. 2001. *Teaching and Researching Lexicography*. Essex: Pearson Education Ltd.

Hartmann, R.R.K. and G. James. 1998. *Dictionary of Lexicography*. London: Routledge.

Landau, S.I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Second edition. Cambridge: Cambridge University Press.

Mavoungou, P.A. 2001. Macro- and Microstructural Issues in Mazuna Lexicography. *Lexikos* 11: 122-138.

McKean, E. 2000. Dictionary Activities in the Elementary Classroom: News for Lexicographers. *Dictionaries* 21: 81-89.

Moon, R. 1987. The Analysis of Meaning. Sinclair, J.M. (Ed.). 1987. *Looking Up: An Account of the Cobuild Project in Lexical Computing*: 86-103. London: Collins ELT.

Mpofu, N. 2001. Problems of Equivalence in Shona–English Bilingual Dictionaries. *Lexikos* 11: 242-251.

Murphy, M.L. 2000. Dictionaries in the Classroom. *Dictionaries* 21: 78-80.

Nyembezi, S. 1991. *Uhlelo lwesiZulu*. Fifth revised edition. Pietermaritzburg: Shuter and Shooter.

Pelling, J. 1971. *A Practical Ndebele Dictionary*. Harare: Longman Zimbabwe.

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

Vikør, L.S. 1993. *The Nordic Languages*. Oslo: Novus Press.

Other References

IsiNdebele Academic Study Syllabus. 2002. United College of Education.

IsiNdebele Syllabus. 2002. Hillside Teachers' College.

`

Language Contact and Lexical Change: A Lexicographical Terminographical Interface in Zimbabwean Ndebele*

Langa Khumalo, *Ndebele Lexicography Unit, African Languages Research Institute (ALRI), University of Zimbabwe, Harare, Zimbabwe (langa@arts.uz.ac.zw)*

Abstract: In this article linguistic changes and/or developments in the Ndebele language spoken in Zimbabwe are investigated and analyzed and the implications these have for term creation and standardization through lexicography. This is followed by a brief historical analysis of the nature and causes of language change in Ndebele. The first part of the article gives a brief background of the Ndebele language, highlighting the movement of the Ndebele people from South Africa which provided a fertile ground for the process of language contact and therefore language change. The second part of the article contrasts terminography with lexicography. The Ndebele corpus of both spoken and written material demonstrates a large extent of borrowing and also loss of certain lexical items. Ndebele lexicograpers were expected to introduce some terms for the purposes of popularizing their use and consequently their acceptance.

Keywords: NDEBELE, LANGUAGE CONTACT, LEXICAL CHANGE, VOCABULARY CHANGE, LEXICAL ENGINEERING, LEXICOGRAPHY, TERMINOGRAPHY

Opsomming: Taalkontak en leksikale verandering: 'n Leksikografiese terminografiese koppelvlak in Zimbabwiese Ndebele. In hiedie artikel word linguistiese veranderinge en/of ontwikkelinge in die Ndebeletaal wat in Zimbabwe gepraat word, ondersoek en ontleed, en die implikasies wat dit vir taalskepping en standaardisering deur die leksikografie het. Die eerste deel van die artikel gee 'n kort agtergrond van die Ndebeletaal, met beklemtoning van die beweging van die Ndebelevolk vanaf Suid-Afrika wat 'n vrugbare grond vir die proses van taalkontak en gevolglik taalverandering verskaf het. Dit word gevolg deur 'n kort historiese ontleding van die aard en oorsake van taalverandering in Ndebele. Die tweede deel van die artikel stel terminografie teenoor leksikografie. Die Ndebelekorpus van sowel gesproke as geskrewe materiaal laat blyk 'n groot omvang van ontlening en ook 'n verlies van sekere leksikale items. Van Ndebeleleksikograwe is verwag om sekere terme op te neem met die doel om hul gebruik te populariseer en gevolglik hul aanvaarding te bewerkstellig.

---

*   This article was presented as a paper at the Seventh International Conference of the African Association for Lexicography, organised by the Dictionary Unit for South African English, Rhodes University, Grahamstown, 8–10 July 2002. It owes its scope to the following works: Hadebe (2001), Hadebe (2002), from where the article draws most of its examples, and Khumalo (2001).

Sleutelwoorde: NDEBELE, TAALKONTAK, LEKSIKALE VERANDERING, WOORDE-
SKATVERANDERING, LEKSIKALE AANPASSING, LEKSIKOGRAFIE, TERMINOGRAFIE

1.   Introduction

The article investigates and analyses linguistic changes and/or developments
that the Ndebele language, spoken in Zimbabwe, might have undergone from
its earliest attested form to its present-day form and the implications these have
in term creation and standardization through lexicography. Language change
is oftentimes viewed as obvious and at the same time mysterious. The Ndebele
of the past is very different from the modern Ndebele. The existence of such
differences between early and later variants of the same language raises ques-
tions about how and why languages change over time. This article will make a
very brief historical analysis of the nature and causes of language change in
Ndebele. Ndebele has seen many modifications to its lexicon. The article will
therefore focus on lexical change in Ndebele.

The first part of this article will give a brief background of the Ndebele
language. It will highlight the movement of the Ndebele people from South
Africa where they had linguistic contact with, among other language groups,
the Zulus, Xhosas, Swatis, Sothos, and the Afrikaners. Later in Zimbabwe the
Ndebeles had further contact with the Kalangas, Shonas, Vendas, Nambyas,
Tswanas, and Tongas whom they subjugated and incorporated into their politi-
cal system. Even later, they had further contact with English and the techno-
logical advancement it brought with it. From a linguistic point of view, the
above scenario provides a fertile ground for the process of language contact
and therefore language change.

The second part of the article will contrast terminography with lexicogra-
phy. Although they seemingly have much in common, they also differ greatly.
The specialized nature of the lexical items studied in terminography gives the
discipline its own distinguishing features. In 2001, Ndebele lexicographers
produced their first corpus-based monolingual dictionary *Isichazamazwi Sesi-
Ndebele*, henceforth referred to as the ISN. The Ndebele corpus of both spoken
and written material demonstrates a large extent of borrowing and also loss of
certain lexical items in the Ndebele lexical inventory as a result of reasons
stated above. The article will finally demonstrate that Ndebele lexicograpers
were expected to introduce terms for the purpose of popularizing their use and
consequently effecting their acceptance, since some of them can no longer be
excluded from the Ndebele vocabulary.

2.   Historical Background of the Ndebele People and Language

The history of the Ndebele language is closely intertwined with the history of
the Ndebele people themselves. It is difficult to discuss one without discussing

the other. The Ndebeles are a Bantu people found mainly in the western parts of Zimbabwe. Their language belongs to the Nguni sub-group of the Bantu language family. Some Nguni languages are Zulu, Xhosa, Swati and South African Ndebele. The history of the Ndebele people and their language isiNdebele dates back to the 1820s when the people who are today known as the Ndebele, broke away from the then mighty Zulu kingdom, in the present-day province of KwaZulu-Natal in South Africa. The tale of Mzilikazi's fight with Tshaka and the subsequent migration of him and his followers from Zululand is not only well documented but is also well known and treasured by both the Zulu and the Ndebele people. The original group that left Zululand was initially called the Khumalo, as their leader, Mzilikazi, was from the Khumalo clan. When and why the group was later renamed Ndebele is a subject that has been debated without convincing conclusions, partly because the subject is full of half-truths and myths. The issue is further complicated by the fact that by then the Ndebele people had not developed or adopted any writing system from which the various claims being made can be verified (Khumalo 2004).

Mzilikazi and his Khumalos, as they were known, then moved northwards from Zululand into Sotho territory (in the present-day province of Gauteng in South Africa). Mzilikazi assimilated a number of Sotho people either through persuasion or coercion or by employing both means. What can be ascertained now is that the Sotho people soon outnumbered the original Khumalos who were of Nguni origin and consequently the group ceased to be referred to as the Khumalos. The Sotho referred to all raiding Nguni tribes as *kiMatebele* (meaning 'warriors with long shields'), and hence the name Matebeles for the Ndebeles, as well as the Ndebele provinces of Matabeleland. How the name Ndebele was derived from Matebele is still a controversy in historical and linguistic studies on Ndebele, that is, if the name derives from it at all (see Hadebe 2002).

The Ndebele people, as they are known today, came to settle in the southern part of present-day Zimbabwe at Ntabazinduna near Bulawayo around 1838–1840. The southwestern parts of Zimbabwe are today referred to as the Matabeleland provinces. Mzilikazi continued to expand his nation by raiding and incorporating many tribes into his political system. It is in Zimbabwe that the Ndebeles assimilated by far the largest number of peoples. These included mainly the Nyubis, Kalangas, Nanzwas, Nambyas, and to some extent the Tongas and some Shona tribes. Most of these were previously under rulers known as the Mambo and hence in Ndebele they were referred to as *Abaka-Mambo*, that is, 'Mambo's people' (see Hadebe 2002).

The name Ndebele therefore did not refer to a single tribe but to a multi-ethnic nation. Within the newly-found Ndebele nation which was then a kingdom, the Sothos outnumbered the Khumalos (or Ngunis) while the Abaka-Mambo far outnumbered both the Khumalos and the Sothos. The original Khumalo group referred to themselves as *AbeZansi* meaning 'those from the south', while the Sotho group was known as *AbeNhla* meaning 'those from the

north' and lastly the rest of the majority were *AbakaMambo*. In spite of their numerical superiority, the incorporated groups learnt the Nguni language that is today called IsiNdebele or simply Ndebele.

Whether the language spoken by the original Khumalo clan that rebelled against the Zulu kingdom was Zulu or not, cannot now be easily ascertained. What is possible to demonstrate today is the close affinity between the Zulu and the Ndebele languages. The languages are mutually intelligible and share probably over 85% of their lexicon. A number of views have been put forward to explain this similarity between the two languages. The predominant view is that Ndebele is a dialect of Zulu. For instance, Pelling (1971: 3) writes: "Ndebele, while still correctly regarded as a dialect of the Zulu cluster, has diverged even further from standard Zulu ..." One possibility, although not popular, is that Ndebele is not a dialect of Zulu but both are sister languages with a common ancestor, which is proto-Nguni.

3.   Language Change

In language contact studies, languages are presumed to come into contact with each other. In actual fact, it is always the speakers of the languages who are in contact. Their attitudes towards each other will influence the way they speak. It is, however, convenient for those who study language simply to speak about languages as though they had a life of their own.

Contact between people speaking different languages can have a wide variety of outcomes. In some cases only a few words are borrowed, in others whole new languages may be formed. The results of such contact differ according to several factors, including the length and intensity of contact between the groups, the types of social, economic and political relationship between them, the functions which communication between them must serve, and the degree of similarity between the languages they speak.

Most languages, including Ndebele, have been influenced at one time or another by contact, resulting in varying degrees of transfer of features from one to the other. Transfer of some features sometimes does not even require speakers of the different languages to have actual contact since it can be accomplished through book learning by teachers who then pass on the new vocabulary to other speakers via literature, religious texts, dictionaries, and so on.

As has been stated above, when the Khumalo clan left Zululand they had contact with other tribal groups. Between 1822 and 1837 the Ndebeles sojourned in the present Gauteng province. One phenomenon associated with their stay in the former Transvaal region was contact with missionaries. The London Missionary Society had established a mission station at Kuruman among the Tlapings. Although the Ndebeles did not generally embrace the Christian faith, some Tswana words from Christianity found their way into the Ndebele language. The following are some examples:

EXAMPLE A

(i)

Sotho Name	Ndebele Adoptive	English Equivalent
Modimo	*uMlimu*	God

It should be noted that this "new" Ndebele word does not refer to God's greatness which is captured by *uNkulunkulu* in Ndebele.

(ii)

Sotho Name	Adopted Ndebele Name
Motlokwa	*Mdlongwa*
Tau	*Dawu*
Mokgatla	*Mnkandla*
Nare	*Nyathi*

(iii)

English Word	Ndebele Word
two shillings	*isikhwitshimana*
two shillings and six pence	*ingogo*
one shilling	*iloso*
six pence	*ingcili*
one penny	*indibilitshi*
sugar	*umangungu*

Despite the rapidly expanding world of the Ndebeles following the introduction of numerous material items, their language matched these changes. New Ndebele words were coined to cope with these.

The Tswana people who were incoporated into the fast-growing Ndebele empire abandoned their language in preference of Ndebele. Even their tribal names were changed to sound more like Ndebele names as demonstrated in the second part of Example A. Change was inevitable, and it favoured the dominating tribe, the Ndebele. Trade is said to have intensified in the Transvaal and coins were introduced and appropriately named as illustrated in the third part of Example A.

One of the greatest changes to the Ndebele language was wrought after the group's arrival in Zimbabwe. As has been stated above, this was the time when they had come into contact with the Nyubi, Kalanga, Nanzwa, Nambya, Tonga and some Shona tribes. Some changes resulting from contact between the Shona and the Kalanga groups are as follows:

EXAMPLE B

(i)

Ndebele	Shona	English
Umuntu umile* *Umuntu umi*	*Munhu amire*	A person is standing

Samukele* *Sethule*	*Tigashire/Tigamuchire*	Welcome
Hambisa* *Tshitsha*	*Fambisa*	Hurry

Words marked by an asterisk represent constructions influenced by Shona (Hadebe 2002).

Another example of Shona influence is the use of the honorific plural or pronouns of power in the Ndebele language. Examples are:

(ii)

Ndebele	Shona
Salibonani* *Sakubona*	*Mhoroi*
Linjani* *Kunjani*	*Makadii*

Words marked by an asterisk represent constructions influenced by Shona (Hadebe 2002).

3.1   Vocabulary Change

Language change can also result from a people going through a certain crisis in its history. New vocabulary is coined to describe the existing phenomenon. Carter and McCarthy (1988: 19) state: "The vocabulary of a language is in constant flux; old items drop out, new terms come in, and as the new replace and augment the old, so the internal relations of the whole set alter." The lexicographer is caught in the dilemma where words are coming into the language while others are dropping out. In the Ndebele-speaking provinces of Zimbabwe the Ndebele people went through the war of liberation (1972–1979) and the subsequent civil war (1982–1987). The following were some of the words that underwent semantic shift and some that came into the language.

EXAMPLE C

(i) Words that underwent a semantic shift

Ndebele Word	Background
umthengisi	Before the war this word referred to a shopkeeper or any salesperson. After having undergone pejorative changes, it now refers to a traitor.
umjibha	This word referred to Zimbabwean migrant workers in Johannesburg. Staying in crime-ridden Johannesburg and surviving there, these workers were perceived to be brave and brilliant. During and after the war, the word *umjibha* switched meaning to refer to a male assistant to guerrillas, who helped in passing information between guerrillas and villagers.

abafana	During the war, villagers called guerrillas *abafana* (the boys) to avoid being detected by Rhodesian informers.

(ii) Words coined during the war

Ndebele Word	Source Word
itororo	The white Rhodesian government referred to the guerrillas as 'terrorists'.
amalwecatsha	This word referred to the nature of guerrilla fighters: 'those who fight while hiding'.

The guerrillas were trained and armed largely by the then socialist countries and independence was equated with a socialist revolution. Some terms related to the socialist ideology were brought into the Ndebele language.

(iii) Borrowed words

Ndebele Word	Source Word
ubhuzhwa	The French word *bourgeoisie*, transliterated to refer to one who grabs wealth.
ipovo	Portuguese word for 'civilian population'. Some Zimbabwean guerrillas were trained in the Portuguese-speaking countries of Angola and Mozambique. They brought with them Portuguese slogans and words that eventually found their way into Ndebele vocabulary.
ibhesi	From English *military base*
iphungwe	From Shona *pungwe*
ikatutsha	Derived from *katucha*
ibhazuka	Derived from bazooka
u-akha	Derived from *AK 47 rifle*
ikhiphu	War camps During the war, the Rhodesian government created security camps where villagers were detained.

3.2   Lexical Engineering

Lexical engineering is viewed by Singleton (2000: 152) as a discipline that "involves not only the coining of new expressions but also the modification or in some cases the suppression — or attempted suppression — of existing expressions". In Ndebele, there are a number of cases where some words have been coined while some have been suppressed, especially due to the influence of the Christian faith. Words that have come into the Ndebele vocabulary through the advent of Christianity include the following:

umfundisi (reverend and teacher)

Education was introduced to the Ndebeles by the missionaries who did preach-

ing as well as teaching in their newly established schools. The preacher was also seen as a schoolteacher. *Umfundisi* is derived from the verb *-fundisa* (to teach). When the colonial government in Rhodesia began to build schools for Africans in the 1960s, teachers were hired who were not necessarily preachers. In Zimbabwe today, there no more exists an obvious link between education and the Church and teachers are seen independently from religious duties. The word *umfundisi* has now narrowed to refer only to a preacher. A teacher is called *umbalisi*, derived from **-*balisa*** (to help to read).

umtshumayeli (preacher)

Umtshumayeli is the other word for 'preacher'. As Ndebele traditional religion did not have a preacher, this word should have been coined to denote the Christian preacher. However, the verb *-tshumayela* (to preach) seems to have been in the language but it referred to narrating. Now the verb *-tshumayela* has shifted its meaning to refer exclusively to preaching.

umvangeli (evangelist); *ivangeli* (evangelism); *-vangela* (to evangelize)

The influence of the Christian faith has led to the transliteration of the word 'evangelism' to *ivangeli* in Ndebele, as well as to the formation of *umvangeli* (evangelist) and *-vangela* (to evangelise). These very common words now appear as if they are indigenous to Ndebele. Another word seemingly indigenous to Ndebele is *isono* (sin). It is now difficult to tell whether the word *isono* is a transliteration from the English 'sin' or whether it is derived from the Ndebele verb *-ona* (to err). The noun derived from the verb *-ona* is *isono* (literally a wrong/crime). The word *isono* (sin) has a religious connotation of a very evil deed that is forbidden.

umbhabhathizo (baptism); *-bhabhathiza* (baptize); *umbhabhathizi* (a Baptist)

These three words which are now extensively used in Ndebele, are an indisputable part of the Ndebele vocabulary although it is known that they are transliterations from English.

Some words have shifted in meaning from what they originally meant and have taken a certain religious connotation.

-khonza (to worship)

In Ndebele, the word *-khonza* means 'to be loyal' or 'to like'. If a subject is loyal to his master, the Ndebele would say *uyamkhonza* (is very loyal). If one is very fond of someone or something, *uyamkhonza* is also used. For instance, the Ndebele might say *uyikhonzile intombi yakhe* (he is very fond of his girlfriend). This sense of the word *-khonza* is now rarely used and is even excluded from the

ISN, the sense of 'worship' being the only sense given. The Christians, however, do not confine *-khonza* to their worship but extend it to the traditional Ndebele religion as well. In Zimbabwe, Christianity is viewed as superior to and thus opposes and denigrates African religions. The traditional Ndebele belief system says there is an almighty creator of the earth *uNkulunkulu* (the Great, Great One). However, the Ndebele did not pray to or communicate with *uNkulunkulu* directly but instead they appeased their ancestral spirits in the event of misfortunes. The Christians, on the contrary, are very much opposed to any rites or ceremonies related to the traditional Ndebele religion. The Christians refer to these rites as ancestor worship — *ukukhonza amadlozi* (to worship spirits). Those who practice the traditional Ndebele religion do not consider their acts as *ukukhonza* (worship) but as the atonement of angry spirits. The Christians have stigmatised the traditional Ndebele religion by deliberately extending the term *-khonza* to include ceremonies and rituals not being worshipping at all.

-khuleka (to pray)

This word has also shifted in meaning to refer to praying. In Ndebele society when you visit a home, you have to announce your presence at the gateby shouting and saying *E Kuhle*! (All is well!). The head of the homestead will shout in response and grant you permission to enter his home (Hadebe 2000). This act is called *ukukhuleka* (announcing your presence). Christians believe that in praying they are asking God to listen to their requests and the Ndebele church-goers equate this act to announcing one's arrival at a homestead — *ukukhuleka.* Nowadays, the Ndebele people have adopted the Western way of knocking at somebody's door. As the practice of the traditional *ukukhuleka* has vanished in the urban centres, and the word now only refers to praying.

isiphambano (a cross)
isidlo (holy communion)

The Roman Catholic Church is the largest Christian denomination in Zimbabwe, and these two words have therefore become very important among the Ndebele Christians. The word *isiphambano* (a cross) is derived from the verb *-phambana* (to cross things), while *isidlo* (holy communion) is derived from the verb *-dla* (to eat). These two words are now part of the Ndebele vocabulary, being some of the many words most speakers take for granted as indigenous Ndebele words, so that their origins are hardly ever questioned.

3.3   Lexical Engineering with Reference to Loanwords

Hadebe (2002) observes that there was a period in the 1970s when loanwords from English were being avoided in the Ndebele language. This period of purism saw the substitution of loanwords from English by words believed to be indigenous as they were taken from Zulu. It should be noted that most Nde-

bele speakers look to Zulu as the source of indigenous vocabulary. It is unfortunate that this purism did not purge Ndebele of European loanwords as intended but substituted English by Afrikaans as the source language. That does not mean that Ndebele speakers preferred Afrikaans to English, but they adopted these Afrikaans words via Zulu believing them to be indigenous Zulu words. The cause for these Afrikaans adoptions was the contact of the Zulu people with the Afrikaners.

EXAMPLE D

(i) Adopted words from Afrikaans

Ndebele Word	Afrikaans Word	English Equivalent
ifasiteli	*venster*	window*
itafula	*tafel*	table*
ipulazi	*plaas*	farm*
isitulo	*stoel*	chair
ibhatshi	*baadjie*	jacket
ijazi	*jas*	coat
umfolo	*voor*	furrow
umjaho	*jaag*	chase

The English words marked by an asterisk have the Ndebele forms shown under (ii) as equivalents. These were discouraged, while the Ndebele forms originating from Afrikaans given under (i) were promoted.

(ii)

English Word	Ndebele Word — Avoided
window	*iwindi*
table	*ithebuli*
farm	*ifamu*

Ndebele has not all the time been permissive to foreign influence as is shown by this deliberate suppression of English vocabulary items. Unfortunately ordinary speakers are not aware of the false basis of this purism.

In an attempt to use indigenous words as opposed to loanwords, a number of words were coined in Ndebele to replace the loanwords. Some of these include *untshelelezane* (bicycle) as opposed to the loanword *ibhayisikili* (bicycle), *inkolo yomlilo* (motor car) as opposed to *imota* (motor car) and *umbalisi* (teacher) as opposed to *utitshala* (teacher). All these words are lemmatised in the ISN, although the Ndebele language corpus shows the prevalent use of the loanwords. This lexical engineering is a phenomenon not exclusive to Ndebele but common to languages in general. It is part of an attempt by society or sections of society to control vocabulary change.

There has also been influence from other African languages in contact with Ndebele, especially Shona, Kalanga, Sotho, Venda, Nambya and Tonga. A number of words currently taken for granted to be Ndebele originally came

from Shona, Kalanga or Sotho. One most debated word is the verb *-kwanisa* (be able). Language purists and teachers would insist that the word *-enelisa* (be able) should be used instead, being considered good and appropriate language. The ISN did not lemmatise *-kwanisa* even though it is more common than *-enelisa* for fear of it being labelled as 'bad language'. There are also words with sounds that are resisted by some sections of the Ndebele-speaking community, especially the elderly and those of Nguni origin. The affricate /dz/ is not original to Ndebele. Currently, however, there are a number of words with this sound. Instead of articulating the sound /dz/ some speakers replace it with /j/. Most words with these sounds have become variants as shown in the following examples:

> *dzikama* — *jikama* (be steady)
> *ubudzugwe* — *ubujugwe* (type of mushroom)
> *-dzimila* — *jimila* (to be lost and have memory lapses)

These examples illustrate the problems encountered by Ndebele lexicographers in headword selection and in defining.

4.   The Lexicographical Terminographical Interface

Seemingly lexicography and terminography have much in common. They are both concerned with describing lexical items in a user-friendly format within a dictionary. The main concern of both lexicography and terminography is thus the word.

Lexicography is often defined as the science of dictionary making. Svensen (1993: 1) defines lexicography as a branch of applied linguistics which consists in observing, collecting, selecting, and describing units from the stock of words and word combinations in one or more languages.

The purpose of terminography is to identify and analyze lexical items used in specialized domains of knowledge, such as commerce, medicine, law and computing. In principle, all domain-specific terms are of interest. In practice, terminographers are mostly preoccupied with 'new' terms: as domains change or grow, often at a startling pace, terminographers must document the associated lexical changes.

Whereas lexicography is essentially a semasiological enterprise, that is, starting with words and trying to explain their meanings, terminography has a significant onomasiological component, that is, starting with the analysis of a domain in order to establish its limits, relations to other domains and subdomains. Since understanding a domain means understanding the domain concepts, conceptual analysis is considered the cornerstone of terminography (Meyer and Eck 1996; Picht and Draskau 1985). Lexicography therefore deals with general language while terminography is domain specific.

As they should be complementary, these two disciplines are vital to each

other. The challenge therefore faced by lexicographers in Zimbabwe is the absence of terminography as a discipline. The possibility must be there for words created by terminographers to be standardized by way of including them in dictionaries. Although it does not exist as a recognized discipline in Zimbabwe, terminography is actively practised among the Ndebele people, as has been demonstrated in this article.

Through the production of a corpus-based Ndebele monolingual dictionary, the ISN, lexicographers have taken advantage of the closeness of these disciplines to standardize terminology that has originated with the people. Because of language contact and the resultant lexical change, Ndebele has rapidly changed and continues to change. The lack of terminography as a discipline has meant that words are not quickly identified and standardized in their various domains so that they can be put to use as part of the core vocabulary.

The question could be asked: Should the lexicographer exclude these unstandardized lexical items? The answer is no. The Ndebele lexicographer has had to double as both terminographer and lexicographer to bring these lexical items into the mainstream vocabulary, giving them a new tag as formal words.

The following is a sample of words which found a new status within the Ndebele lexicon. These words were introduced in the ISN for purposes of standardizing them because the corpus demonstrated that they overwhelmingly appear in the everyday use of the language. Of these terms it may be stated quite rightly, that they can no longer be excluded from the Ndebele vocabulary.

EXAMPLE E

ipasipoti bz 5. *Ipasipoti lugwadlana olunikwa umuntu nguHulumende, olutshengisa ukuthi umuntu ulemvumo yokuhamba kwamanye amazwe, kanye lokuthi ungubani lokuthi uyisizalwane saliphi ilizwe.*
(*passport* N *A passport is an official document issued by the government, which contains your personal details, which you need to show when you enter or leave a country.*)

imoningigawuni bz 5. *Imoningigawuni yigawuni umuntu ayigqokayo lapho evuka.*
(*morning gown* N *A morning gown is a long loose robe that is worn usually in the mornings by someone who has just woken up.*)

amavolontiya bz 6. *Amavolontiya ngamabutho ayelwela imali aqhatshwa nguRozi ukuba achithe umbuso kaLobhengula ngo 1893.*
(*volunteers* N *Volunteers are people who were conscripted and paid by Rhodes to destroy King Lobhengula's rule in 1893.*)

bhekha sz mwa. *Ukubhekha yikipheka ngewoveni into enjengesinkwa loba amakhekhe.*
(*bake* VB *To bake food like bread is to cook it in an oven without using extra liquid or fat.*)

itshampiyoni bz 5. **Itshampiyoni** *ngumuntu onguye othethe isicoco saphezulu kumbe onqobileyo emdlalweni othize. Fan ingqungquthela.*
(***champion*** N *A **champion** is someone who has won the first prize in a competition.)*

5.   Conclusion

It has been demonstrated that language contact naturally leads to language change through borrowing or lexical engineering. The result of this process is the addition of new words to the lexicon that need to be standardized and be recognized as formal. Ndebele lexicography has taken an important step towards formalizing new words in its lexicon.

Bibliography

Carter, R. and M. McCarthy. 1997. *Vocabulary and Language Teaching*. London/New York: Longman.

Hadebe, S. 2000. Developing Terminology in African Languages with Special Reference to Ndebele. Chiwome, E.M. et al. (Ed.). 2000. *Indigenous Knowledge and Technology in Africa and Diasporan Communities*: 225-231. Harare: Mond Publishers.

Hadebe, S. 2002. *The Standardisation of the Ndebele Language through Dictionary Making*. Unpublished D.Phil. Thesis. Harare and Oslo: University of Zimbabwe and University of Oslo.

Hadebe, S. (Ed.). 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.

Khumalo, L. 2004. *A General Introduction to Ndebele Grammar*. Cape Town: CASAS.

Meyer, I., K. Eck and D. Skuce. 1997. Systematic Concept Analysis within a Knowledge-based Approach to Terminology. Wright, Sue-Ellen and Gerhard Budin (Eds.). 1997. *Handbook of Terminology Management. Vol. 1*: 98-118. Amsterdam/Philadelphia: John Benjamins.

Pelling, J. 1971. *A Practical Ndebele Dictionary*. Harare: Longman Zimbabwe.

Picht, H. and J. Draskau. 1985. *Terminology: An Introduction*. Guildford: University of Surrey.

Singleton, D. 2000. *Language and the Lexicon: An Introduction*. London: Arnold.

Some Features of Monolingual LSP Dictionaries

G.-R. Mihindou, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa (guyr_mihindou@hotmail.com)*

Abstract: Dictionaries are classified according to a certain set of criteria. On the one hand, there are general purpose dictionaries, aiming to satisfy the needs of the majority of users, and focused on the general vocabulary, and, on the other hand, there are limited or specific dictionaries for specialists dealing with language for special purposes (LSP). The type of dictionary dealing with technical and sometimes semi-technical vocabulary is compiled to satisfy the needs of those who already have a knowledge of a specific field but want to improve their skills. General purpose dictionaries are well known and users are relatively familiar with them, but it is not necessarily the case with LSP dictionaries. This type of dictionaries needs to be described or presented with emphasis on some of its features. As a result LSP dictionaries and general purpose dictionaries will often be compared.

Keywords: LSP DICTIONARY, GENERAL PURPOSE DICTIONARY, MEANING, TECHNICAL VOCABULARY, GENERAL VOCABULARY, USER NEEDS, DATA DISTRIBUTION, MACROSTRUCTURE, MICROSTRUCTURE, ORDERING TYPOLOGY

Opsomming: 'n Aantal kenmerke van eentalige TSD-woordeboeke. Woordeboeke word geklassifiseer volgens 'n bepaalde stel kriteria. Aan die een kant is daar woordeboeke vir algemene doeleindes wat die behoeftes van die meerderheid gebruikers wil bevredig, en wat op die algemene woordeskat konsentreer, en, aan die ander kant, is daar beperkte of spesifieke woordeboeke vir vakkundiges wat taal vir spesiale doeleindes (TSD) behandel. Die soort woordeboek wat 'n tegniese en somtyds semitegniese woordeskat behandel, word saamgestel om die behoeftes te bevredig van diegene wat reeds kennis van 'n spesifieke gebied besit, maar wat hul vaardighede wil verbeter. Woordeboeke vir algemene doeleindes is goed bekend en gebruikers is betreklik vertroud met hulle, maar dit is nie noodwendig die geval met TSD-woordeboeke nie. Hierdie woordeboeke moet omskryf of bekend gestel word met die klem op sekere van hul kenmerke. TSD-woordeboeke en woordeboeke vir algemene doeleindes sal gevolglik dikwels vergelyk word.

Sleutelwoorde: TSD-WOORDEBOEK, WOORDEBOEK VIR ALGEMENE DOELEINDES, BETEKENIS, TEGNIESE WOORDESKAT, ALGEMENE WOORDESKAT, GEBRUIKERSBEHOEFTES, DATAVERSPREIDING, MAKROSTRUKTUUR, ORDENINGSTIPOLOGIE

Introduction

Lexicography is commonly understood as the practice of dictionary-making. This implies the compilation and the preparation of reference texts for publica-

tion. Therefore, an important product of lexicography is a dictionary or word-book (Hartmann and James 1998) which treats many branches of knowledge. The term "dictionary" encompasses many genres or types, which can be structurally classified according to a certain set of criteria. In fact, dictionaries have often been classified in categories based on formal features, like size for example: pocket dictionary, abridged dictionary, concise dictionary, etc.

The major dictionaries in the market are classified according to their function in which the data play a vital role, and according to their intended role in satisfying the potential user's needs. The general purpose dictionary (GPD) for example will present the semantic data in such a way that any user can understand the "common" vocabulary. The restricted dictionary is a special or specialized dictionary focusing on specific users who may be experts, semi-experts or non-experts. That is why Hartmann and James (1998) assert that a functional typology or categorization based on the contextual uses of dictionaries should focus on data types or categories. This article deals with the last category or type of dictionary which focuses on Language for Special Purposes (LSP). The LSP dictionary (LSPD) is a type of dictionary which aims to satisfy the needs of a specific group of users. Gouws (2000: 10) concedes that there is no single dictionary which can be everything to everybody; consequently a specialized dictionary has a very important role to play.

This implies that such a specialized dictionary will present the data in different ways, also treating and presenting the comment on semantics differently. The aim of this article is to analyse some aspects of the treatment and presentation of meaning in LSP dictionaries and their structures. The accent will mostly be on monolingual LSP dictionaries. But firstly a brief discussion of LSP dictionaries will be given.

The LSP Dictionary

In a strict sense, an LSP dictionary is a type of reference work intended to describe language variety used by experts in a particular subject field (Hartmann and James 1998). In a general sense, the LSP dictionary is part of the so-called restricted dictionaries (Burkhanov 1998), special dictionaries (Singh 1982) or technical dictionaries (Hartmann and James 1998). The term "restricted dictionary", according to Burkhanov (1998: 201), is used with reference to a large and heterogeneous group of dictionaries, starting from phraseological dictionaries via terminological and dialect dictionaries to hard word dictionaries. The term "restricted" does not necessarily imply a shorter or smaller size of work, because some restricted dictionaries are quite large. More so, the LSPD can rightly be considered a technical dictionary (Hartmann and James 1998).

By "technical" two kinds of dictionaries must be understood. On the one hand, there is the dictionary which deals with the very technical vocabulary of a specialized subject discipline. The objectives of such a type of dictionary are to satisfy and inform the experts. In fact, professionals in several specialized

fields defend their use of technical and complex language as being the most precise means of expressing technical and complex ideas. Therefore, the evidence is that scientists, doctors, bankers, and others need their jargon in order to communicate with each other succinctly and unambiguously (Crystal 1997: 383). As a result, this type of dictionary is conceived as an aid for users who are already specialists in their field, but want to acquire more skills. On the other hand, it can be a dictionary which considers experts, semi-experts and laypersons as users. In such a dictionary, the vocabulary will be semi-technical. So, as far as the semi-technical dictionary is concerned, there are no theoretical limits to the number of special purposes to which language can be put (Crystal 1997). Recently whole new areas of technical expression have emerged, concerning such domains as computing, broadcasting, commercial advertising and popular music. The lexicographer should therefore include the average user's knowledge needs in these semi-technical LSPDs.

According to Smit (1996: 142), Wiegand (1983: 116) says that the many statements in a general theory of lexicography can be formulated with relation to dictionary typology. As far as the LSPD is concerned, it is relevant to follow both the different typological criteria suggested by Wiegand (1983: 104), based on the purpose of and the language use in dictionaries.

The first suggestion determines the specific structure of a given lexicographic work. The second suggestion deals with the nature of the language currently treated in a dictionary. Here, according to his needs, the potential user becomes the key person in determining the type of language to be used in a dictionary. In this way, Wiegand (1983: 111) explains the theory of lexicographic language description, focusing on the set of eventual questions from the potential users and the different answers the lexicographer wants to develop as an input to a dictionary. Therefore the answers oriented at the potential users will play an important role in determining the type of dictionary.

In the same way, a special-field dictionary will be a dictionary dealing with a specific field and the potential user of this type of dictionary will ask particular questions or will express specific needs, which will guide the compiler. Once the potential users have been identified, their needs will influence the typology of the dictionary and the decisions of the lexicographer. The title of a dictionary indicates which kind of field and part of knowledge the text work covers, the type of lemmata, the structure of lexical items, etc.

The LSP Dictionary and the General Purpose Dictionary

An LSPD is very often opposed to a GPD. The GPD deals with the "commonest", the general vocabulary. But the LSPD treats a specific and precise vocabulary. Singh (1982: 18) observes that general dictionaries contain those words of the language which are in general use representing various spheres of life and presenting a complete picture of the general language. Special diction-

aries either cover a specific part of the vocabulary or are prepared for some definite purpose.

In the same light, Burkhanov (1998) asserts that GPDs are meant to represent the general vocabulary; they are mainly concerned with that part of the lexicon not restricted to a particular style or register. Lexical items, for instance, belonging to the regional and social varieties of a language, as well as the specialized terminology of particular academic disciplines and various subject fields, are not supposed to be included in the word list of a GPD. In some cases, however, they are included, but then provided with a label.

The expectations of the GPD user are very often aimed at some lexicographic data types such as grammatical information, pronunciation, peculiarities of spelling, etc. In this way, the GPD is closer to the expectations of the ordinary user. The LSPD, on the contrary, may account for only one lexicographic parameter, for instance dealing with only one aspect of lexical items or one type of semantic relation between those lexical items. Therefore, the dictionary of homonyms can serve as an example. LSPDs as technical dictionaries are governed by one common principle, namely to be non-ambiguous.

Van Schalkwyk (1987) maintains that it is important to keep in mind that a technical term should have only one logical and unambiguous meaning. This can be considered as the most precise aspect in the treatment of lexical items in LSPDs. Therefore, the consequence of making a comparison between an LSPD and a GPD leads the metalexicographer into another comparison between specialized lexicography, regarded as more prescriptive, and general lexicography, considered as more descriptive.

GENERAL LEXICOGRAPHY	LIMITED/RESTRICTED LEXICOGRAPHY
General vocabulary	Regional and social/specialized vocabulary
Deals with a nonrestricted lexicon	Deals with a restricted lexicon and particular register
Encompasses lexical items of linguistic usage	Concerned with specific terminology
Describes usages understood and accepted by all native speakers of a given language	Describes localized or specific usages obtained by convention for selected speakers.
Expected to provide lexicographic information	Expected to furnish lexicographic description
Various linguistic parameters of lexical items in the central list	Only one type of semantic relation between lexical items in the central list
Example of parameters: explanation of meaning, grammatical information, particularities of spelling	Example of parameters: inflectional or derivational proprieties of lexical items, synonymy, antonymy

Source: Burkhanov (1998)

It is also important to make a distinction between limited/restricted dictionaries and LSP/specialized dictionaries.

LIMITED/RESTRICTED DICTIONARIES	LSP/SPECIALIZED DICTIONARIES
Refer to a large and heterogeneous group of philological dictionaries	Refer to lexicographic works for linguistic and non-linguistic aspects of specialized terminologies
Regional and social varieties of language as well as specialized terminology of a specific subject field	Specialized terminology of a specific subject field only
Multi-field and single field terminology	Multi-field and single field terminology
Limited to a particular style or register	Focus on scientific and technical vocabulary
Restriction imposed on the selection of lexical items	Restriction imposed on the selection of lexical items
Examples of dictionaries: phraseological, pronunciation, terminological dictionaries, etc.	Examples of dictionaries: terminological, technical, translation dictionaries, etc.

Source: Hartmann (1983)

The distinction between these two categories of dictionaries can be illustrated as follows: The *Dictionary of 1000 French Proverbs* will fall under restricted dictionaries because of its limited register, *The Oxford Concise Dictionary of Linguistics* will be included under LSP dictionaries due to the specialization of the field it presents.

Svensén (1993: 2) who supports the assertion made by Geeraerts (1987) says, "The difference between general and technical dictionaries is not just a matter of selection. The methods of description and explanation are also different. A definition of concepts in a technical dictionary is often more detailed than that of the same concept in a general dictionary."

Hartmann and James (1998) attempt a classification based on general reference work vs. specialized reference work, as well as linguistic information vs. factual information.

GENERAL		SPECIALIZED	
Linguistic information	*Factual information*	*Linguistic information*	*Factual information*
General Dict.	Encyclopaedia	Dict. of Foreignisms	Dict. of Music
Learner's Dict.	Almanac	Dict. of Idioms	Dict. of Medicine
Historical Dict.	Companion	Dict. of Archaisms	Dict. of History
Bilingual Dict.	Biographical Dict.	Dict. of Proverbs	Dict. of Law
Dict. of Names	Directory	Dict. of Quotations	Dict. of Botany
Spelling Dict.	Catalogue	Dict. of Allusions	Dict. of Geology
Pronouncing Dict.	Bibliography	Dict. of Confusibles	Dict. of Chemistry
User guide	Compendium	Dict. of Slang	Dict. of Art
Dialect Dict.	Travel Guide	Dict. of Neologisms	Dict. of Physics
Thesaurus	Atlas	Dict. of Jargon	Term Bank

Source: Hartmann and James (1998)

This classification confirms the assertion by Gouws (2000: 11) on the risk of confusion between sets of reference works and therefore the pertinence of criteria used for any attempt at classification. The bilingual dictionary, for instance, considered as part of general reference works could also be classified among the specialized reference works. There are some bilingual dictionaries focusing on languages such as English and Afrikaans as well as other bilingual dictionaries dealing with special fields such as medicine, botany or scientific terms generally. The fact that *linguistic information* and *factual information* as criteria of classification are presented in both general lexicography and specialized lexicography indicates a possible hybridization for both lexicographic categories.

The viewpoint of Wiegand about special-field dictionary typology is presented in detail by Smit (2000: 174-177). Wiegand (1998: 742) refutes any classification on the dichotomy between language dictionaries and encyclopaedias based on the distinction semantic items vs. encyclopaedic items. He rather suggests a classification of reference works focusing on what he calls the genuine purposes of a reference work. Regarding the genuine purpose of a special-field dictionary, Wiegand (1994: 105) stipulates that it is an object used to obtain information that may be considered as special-field knowledge. Wiegand (1998) is also in favour of the idea that reference works should be grouped in two distinct classes: the first, K1, will be the class of linguistic objects whereas the second, K2 will therefore be the class of non-linguistic objects or "things"; K2 being the class of all reference works with a genuine subject-lexicographic purpose.

From the above discussion, it becomes clear that it is difficult to classify dictionaries very strictly. The chances are good that one set of criteria is considered and others are neglected. The languages themselves do not obey such rigid categorisation. Furthermore, the borders between categories are unclear. Gouws (2000: 11) rightly says that the differences between two neighbouring categories are often so few that a specific dictionary could be placed in either category. Svensén (1993: 21) confirms this, saying, "The distinction between general and technical dictionaries can often be difficult to define. In general terms, purely technical dictionaries are commoner than purely general dictionaries." Regarding differences between general and technical dictionaries, some classifications may be made. Therefore, Al-Kasimi (1977) proposes that dictionaries can be classified in various ways according to their range, perspective, purpose, and presentation. In this article, two types of LSPDs have been retained and adapted, based on the proposition of Singh (1982).

(a)   Dictionaries of technical terms: They deal with technical terms in a language, and, as has been said above, terminology plays the most dominant and vital role in this type of dictionaries. They are focused on special fields in life, and are often presented by special bodies and commissions formed especially for that purpose. They contain either terms peculiar to a particular subject field or general words with special mean-

ings for special fields. A dictionary of lexicography, a dictionary of medical terms, a dictionary of linguistics, etc., can serve as illustrative examples.

(b)   Dictionaries of special professions: These dictionaries present lexical items in a particular way according to the professional jargon. A dictionary of antiques and decorative arts, a reference book for glass collectibles, a dictionary of technical terms containing definitions of commonly used expressions in aeronautics, etc., can be considered as examples of these.

To summarize what has been said above concerning the LSPD, the following points should be kept in mind: Firstly, this type of dictionary depends totally on the aspects of the language covered. Secondly, it can be limited to one aspect of the language. And thirdly, this kind of dictionary deals with only one logical and unambiguous meaning.

Meaning in the LSP Dictionary

As has been said in the introduction, the aim of this article is to present some aspects of the treatment of meaning in the LSPD. Users consult a dictionary primarily to establish and understand the meaning of a lexical item. According to the quality of a text work, the meaning of one lexical item can have a more or less specific treatment in a dictionary. The treatment of meaning is determined on the one hand by the needs of the user, and on the other hand by the goals the lexicographer wants to achieve. Burkhanov (1998: 148) sees in meaning the major part of linguistic semantics and lexicological study, an important concept for lexicographic description of the lexicon and a possibility for making a considerable impact on the final lexicographic product.

Semantic data is presented in the dictionary in many ways. One of these is the description of meaning or the lexicographic definition. Here, the definition consists of a definiendum and a definiens. The definiendum is a lemma, the term or the word, which the lexicographer wants to explain in the dictionary article, whereas the definiens is the explanation of the meaning. Many lexicographers use the term "description of meaning" or "lexicographic definition of meaning". Wiegand (1992) prefers the term "paraphrase of meaning".

The lexicographer uses different methods to describe and to identify the meaning of a lexical item. According to Gouws (2000), not all lemmata can be defined in the same way and lexicographers rely on different types of definitions in the process of explaining meaning. Therefore, four different types of definitions are frequently used, namely:

(a)   The generally descriptive definition in which the meaning of the definiendum is explained by the definiens without any reference to the seman-

tic relationship between the definiendum and other lexical items. Monolingual LSPDs often use this category of the explanation of meaning.

(b)   The genus–differentia definition which does not only explain the meaning of the lemma, but also places the lemma in a class (genus) and operates by distinguishing through opposition between the lemma and other members of the class. In the context of explanations of meaning, this category is especially used in monolingual LSPDs.

(c)   The synonym definition which gives a synonym for the lemma. This category is used in LSPDs, especially in the context of translation equivalents in the target language.

(d)   The circular definition which implies a very strong resemblance between the definiendum and definiens.

All these different mechanisms employed by the lexicographer to describe, explain, and identify the meaning should be presented as clearly as possible if he wants to see his dictionary used successfully.

When the lexicographer speaks about meaning, he is dealing with more than a simple understanding of an idea. Words, references, symbols, and mental content are involved. It is almost impossible to speak about meaning in a dictionary without the word supporting this meaning. So meaning and word are intimately linked but not necessarily fused. The word the lexicographer includes as lexical item in a dictionary, like, for instance, a dictionary of econometrics, is a sum of symbols linked together in a way to be pronounced as one unit and interpreted in a certain way. This interpretation brings out the meaning. The process of this operation implies certain mechanisms in the mind of the user and the compiler.

As a result, one word can have one, or more than one, meaning. Therefore the lexicographer will be able to treat the lemma in an appropriate way according to the context or the co-text. He can, for instance, treat a lemma in such a way that the differentiations between denotation and connotation will be clear to the user and will indicate at the same time relations of meaning such as synonymy, antonymy, hyponymy, etc. This point will be illustrated in the examples below from an LSPD as a technical dictionary.

By a technical dictionary should be understood two types of dictionaries. On the one hand, there is the technical dictionary for experts, the terminological dictionary (Hartmann and James 1998) in which the emphasis is on expert knowledge, which implies the description of scientific and technical words in a certain field in a precise and specialized way. On the other hand, there is the LSP dictionary intended as a text work containing a semi-technical vocabulary, intended for relative neophytes or advanced learners (Hartmann 1998: 136).

As a technical dictionary the LSP dictionary uses a technical language when it refers to a specific register of the language dealing with a more or less

hermetic vocabulary and meanings as well. For example, a lexical item such as *absence* has different meanings, depending on the type of dictionary. The *Long-man Dictionary of the English Language* defines the word as follows:

> **absence** /ˈɑbsəns/ *n* **1** the state of being absent … **2** the period of time that one is absent **3** a want, lack ⟨an ~ of detail⟩
> **absence of mind** *n* inattention to present surroundings or occurrences

In *The New Shorter Oxford English Dictionary* the following definition is presented:

> **absence** /ˈɑbs(ə)ns/ … *n* … **1** The state of being away from a place or person; the time or duration of being away. ... **2** Non-existence or lack *of.* ... **3** Inattention due to thinking of other things. Esp. in *absence of mind.* **b** *Med.* Sudden tempo-rary loss of consciousness occurring without the awareness of the subject, esp. in petit mal; an episode of this (also *absence seizure*). ...

Both these dictionaries are considered as GPDs. The definiens, i.e. the explana-tion of the meaning, has been presented in such a way that "any" dictionary user can recognize and comprehend it as belonging to the general or common vocabulary. It therefore shows that in a GPD the emphasis is on the presenta-tion or the treatment of the "ordinary" meaning of lexical items encompassing a general view.

   Bell's South African Legal Dictionary, however, approaches *absence* differ-ently:

> **ABSENCE,** as a ground for relief under the Insolvency Act, No. 32 of 1916, means *absence* from the country; that is, from the Union and not merely from one Province thereof (*Estate Elaides* v. *Sackville Clothing Co. and another*, 1925 W.L.D. 144).

In *Black's Law Dictionary absence* is defined as follows:

> **absence,** *n.* **1.** The state of being away from one's usual place of residence. **2.** A failure to appear when expected.

These dictionaries may be considered technical dictionaries because of their very restrictive approach in terms of macrostructural selection and treatment. This restrictive approach may be balanced by bringing a semi-technical view of the lexical item *absence*, explained to both expert and neophyte dictionary users.

   Stroud's Judicial Dictionary of Words and Phrases describes five different contexts of use:

> **ABSENCE.** (1) Judicial separation obtained in the "absence" of the respondent (Matrimonial Causes Act 1857 (c. 85), s. 23) meant his or her non-appearance in the suit (*Phillips* v. *Phillips*. L.R. 1 P. & D. 169). ...
>
>   (2) Absence of seaman without leave: defined in Merchant Shipping Act 1894 (c. 60), s. 376 (1)(*b*), as "neglecting or refusing without reasonable excuse, to join or to proceed to sea in his fishing boat, or being absent without leave at

any time within 24 hours of his boat's sailing from any port, either at the commencement or during the progress of the engagement, or being absent at any time without leave and without sufficient reason from his boat": see *Edgill* v. *Alward* [1902] 2 K.B. 239, cited WILFUL DISOBEDIENCE. …

(3) "Absence" (Matrimonial Causes Act 1950 (c. 25), s. 14 (3)) means "physical absence" and not "non-appearance". The fact that a party has entered a limited appearance does not preclude the operation of the section (*Wilkinson* v. *Wilkinson* [1963] P. 1).

(4) "Absence at any school" (Family Allowances Act 1945 (c. 41), s. 21 (7)) does not include being taken into care by a local authority in one of its own schools, where there is no going home for the holidays (*Hill* v. *Minister of Pensions and National Insurance* [1955] 1 W.L.R. 899).

(5) "Absence without leave." Stat. Def., Army Reserve Act 1950 (c. 32), s. 28 (1); Air Force Reserve Act 1950 (c. 33), s. 27 (1).

Osborn's Concise Law Dictionary, however, explains the lexical item *absence* in a very short and simple definiens as follows:

> **Absence.** If a person has not been heard of for seven years, and the circumstances are such that, if alive, he would have been heard of, the presumption of death arises, but not as to the date of death (*Re Phene's Trusts* (1969) L.R. 5 Ch. App. 139). The court may, however, order that death be presumed at any time if sufficient evidence is shown. See the Matrimonial Causes Act 1973, s. 19; Domicile and Matrimonial Proceedings Act 1973, Sched. 6. …

The definitions of *absence* given in the two GPDs, namely the *Longman Dictionary of the English Language* and *The New Shorter Oxford English Dictionary* can now be reconsidered. In the first, the lemma has been treated as a polysemous word, differentiating three senses which occur in general usage. In the second, the treatment of meaning is presented in such a way that the user can quickly distinguish two different components. Firstly, the lemma is shown to be a polysemous word, and, secondly, the lemma has been labelled. An explanation is given according to common usage, employing the general vocabulary to explain the word in three different senses. An explanation of the word used in a specialized field, labelled *Med.*, i.e. in medicine, is also presented. As this specialized meaning differs from the use of *absence* in the general language, the context and field in which it is used must be indicated.

For LSPDs the approach is somewhat different, as can be seen from the definitions in the four technical dictionaries, namely *Bell's South African Legal Dictionary*, *Black's Law Dictionary*, *Stroud's Judicial Dictionary of Words and Phrases*, and *Osborn's Concise Law Dictionary*. These all give senses of *absence* as used in the legal parlance. *Stroud's Judicial Dictionary of Words and Phrases*, for instance, does not use labelling like that in *The New Shorter Oxford English Dictionary*. It develops the lexical item in five different senses which are all essentially focused on the field of law. The explanations are of use to specialists of law only. There are, for example, many quotations dealing with laws, like the

Matrimonial Causes Act, the Family Allowances Act, etc. These reference texts and the vocabulary within them are not for ordinary but for knowledgeable users. In this type of dictionary, the emphasis is not so much on the understanding of the item itself, but on the different contexts covered by the judicial concept of absence, because it is assumed that lawyers already have knowledge of the "ordinary meaning".

Structure of the LSP Dictionary

Data distribution and user needs

Most dictionaries depend on the nature and the extent of the data, especially from the viewpoint of the users' needs. Dictionaries are classified according to certain criteria first established by the metalexicographer. Before the compilation of a dictionary, the metalexicographer has to play an important role in planning the process of the dictionary. In this connection Lombard (2000) says:

> It is very important for the lexicographer to know who the target users of his labours are going to be before work on the dictionary can begin. Because the character of a dictionary and the nature of the material presented in it should be determined by who the target users are, one of the most important tasks of the compiler of any dictionary is to make the product user-friendly. User-friendliness thus becomes one of the criteria in terms of which a dictionary is assessed.

Lexicographers should consider the users' needs in order to establish the dictionary classification even if there is no entirely satisfactory method of classification.

Data plays an important role in the LSPD; it gives consistency to the dictionary and determines which lexical items will be included and which will be omitted. Raw material, data, once collected, will indicate the configuration of the future dictionary. In this case, the purpose of the LSPD is to serve users who already have knowledge of the field; the focus will be on the real needs of the users. In a dictionary of medical terms, for instance, the field is medicine, the potential users are doctors and specialists dealing with this discipline. The metalexicographer is supposed to focus the attention on the users and their needs. The compiler must also consider that in an LSPD focusing, for instance, on medical terms, certain types of users like students in medicine are semi-experts or semi-specialists, who must be enabled to retrieve the information they are looking for. One of the responsibilities of the compiler is to select data and treat it in such a way that the LSPD will satisfy both types of users: experts and non-experts.

When the definitions of *absence* in the four legal dictionaries are compared, it will be seen that those in *Black's Law Dictionary* are less specialized than those in the other three, namely *Bell's South African Legal Dictionary*, *Osborn's Concise Law Dictionary*, and *Stroud's Judicial Dictionary of Words and Phrases*. They al-

most equal those in GPDs. This indicates that these dictionaries are meant for different user groups.

Tarp (2000: 194) concedes that when a lexicographer conceives or produces a dictionary he has at least always a target group of users in mind, be they homogenous or heterogeneous, otherwise it will be difficult to establish a proper lexicographic work based on scientific principles. One of these principles is the obligation for the compiler to know the users' competencies in the subject field in which the LSPD is going to be compiled. The main question will then be: At what level of mastering the special subject field in question are the users?

From the knowledge of a specific discipline, other criteria based on the linguistic performances of the users should be added, such as:

—   At what level have they mastered the corresponding LSP in their mother tongue?

—   At what level have they mastered the corresponding LSP in a foreign language?

The first point is particularly important in the case of monolingual LSP dictionaries, the second point in the case of bilingual LSP dictionaries.

Macrostructure

The macrostructure of an LSPD is not too different from that of other dictionaries in terms of theory, because they all obey several criteria of metalexicography. But the term "macrostructure" covers two different aspects: one which is concrete, and the other which is abstract. So, on the one side, a dictionary presents several sections within which any user can identify a number of dictionary components such as the lemmas, to which other entries are directed. But this arrangement comes from another part of the structure, which is known as the data distribution structure or data distribution programme in which every detail present in the dictionary has been discussed and planned by a team of metalexicographers. This data distribution structure is the only one which can give the most concise and precise configuration of the future dictionary. In an LSPD, this data distribution structure will clearly offer the real intention of the compiler and the real user, and even anticipate and decide which kind of needs the user has. This structure guides the metalexicographer, lexicographer and all specialists interested in how a dictionary has been compiled and what procedure has been followed.

In macrostructural perspective, the structural skeleton of an LSPD basically has the same ordering as GPDs. There are, for instance, two major components in a given dictionary article: comment on form and comment on semantics. The first is considered part of the macrostructural procedure whereas the

second is classified as part of the microstructural treatment. However, LSPDs present fewer entries than GPDs, for example as far as labelling is concerned. The following example taken from *Stedman's Medical Dictionary* can be considered:

> **hermetic** (her-met'ik). Airtight; denoting a vessel closed or sealed in such a way that air can neither enter it nor issue from it.

In a normal GPD, this lexical item would have been treated differently. Apart from the phonetic markings, morphological features would have been included as well as grammatical information such as the part of speech as can be seen in the following definitions from *Chambers 21st Century Dictionary* and *The New Oxford Illustrated Dictionary*.

> **hermetic** /hɜːˈmɛtɪk/ and **hermetical** ▷ *adj.* **1** perfectly closed or sealed so as to be airtight. **2** belonging or relating to magic or alchemy. **3** obscure or abstruse. • **hermetically** *adverb*. • **hermeticity** |hɜːˈməˈtɪsɪtɪ| *noun*. ☉ 17c: named after Hermes Trismegistos, supposedly the Greek author of texts on magic and alchemy, and also the inventor of a magic seal.

> **hĕrmĕ'tic** *adj.* Of Hermes Trismegistus; of alchemy, magical, alchemical; ~ *seal*, *sealing*, airtight closure of vessel, esp. glass vessel, by fusion, soldering, or welding.
> **hĕrmĕ'tĭcallў** *adv*.

In the *Chambers 21st Century Dictionary*, the lemma **hermetic** is followed by a phonetic transcription, /hɜːˈmɛtɪk/, an additional headword **hermetical** as morphological variation, indicating an optional choice offered to the user, then the symbol ▷ used to mark the principal part of speech, in this case *adj.* written in italics. *The New Oxford Illustrated Dictionary* offers slightly less information: The phonetic information forms part of the lemma **hĕrmĕ'tic**, followed by the part of speech *adj.* in italics. In macrostructural perspective, this information may be considered important, and may be useful in GPDs but not necessarily in an LSPD like *Stedman's Medical Dictionary*, the focus here being mainly orientated towards the genuine interpretation of the meaning of the lexical item *hermetic* in the particular context of medicine.

The Chambers Dictionary with a more comprehensive perspective presents its article as follows:

> **hermetic** *hûr-met'ik* or **hermetical** *-met'ik-l*, *adj* belonging in any way to the beliefs current in the Middle Ages under the name of *Hermes*, the Thrice Great; belonging to the magic or alchemy, magical; perfectly closed, completely sealed; obscure, abstruse. – *adv* **hermet'ically**. – *n* **hermetic'ity** – *n sing* **hermet'ics** the philosophy contained in the **hermetic books**, esoteric science; alchemy. – **hermetically sealed** closed completely; made airtight by melting the glass. [Medieval L *hermēticus*, from *Hermēs Trismegistos* Hermes the thrice-greatest, the Greek name for the Egyptian Thoth, god of science, *esp* alchemy]

With regard to the dictionary articles of *hermetic* cited above, it becomes clear that the treatment of meanings, and therefore the purpose of GPDs and LSPDs differs: In the medical dictionary, the lexicographer has included only the essential data, just enough to capture the meaning and define its medical use. In the GPDs, more details are given. Moreover, this comparison shows that what is hierarchically positioned as denotation (of first importance) in an LSP, might be connotation (of second importance) in a GPD.

Ordering typology

Like GPDs, LSPDs often present their lemmata in alphabetical order. In the example of *hermetic* from *The Chambers Dictionary*, "hermetically", "hermecity", "hermectics", "hermetic books", "hermetically sealed" are all presented as sublemmata of the lexical item *hermetic*. This can be compared to the first part of the example of *hernia* from *Stedman's Medical Dictionary*:

> **hernia** (her′nē-ă) [L. rupture] Rupture (1); protrusion of a part or structure through the tissues normally containing it.
> **abdominal h.,** laparocele; a h. protruding through or into any part of the abdominal wall.
> **antevesical h.,** an interstitial h. projecting medially from the internal inguinal ring.
> **Barth's h.,** a loop of intestine between a persistent vitelline duct and the abdominal wall.
> **Béclard's h.,** a h. through the opening for the saphenous vein.
> **bilocular femoral h.,** Cooper's h.
> **h. en bissac,** properitoneal inguinal h.
> **Bochdalek's h.,** congenital diaphragmatic h.
> **h. of the broad ligament of the uterus,** a coil of intestine contained in a pouch projecting into the substance of the broad ligament.
> **cecal h.,** a h. containing cecum.
> **cerebral h.,** protrusion of brain substance through a defect in the skull.
> **Cloquet's h.,** a femoral h. perforating the aponeurosis of the pectineus and insinuating itself between this aponeurosis and the muscle, lying therefore behind the femoral vessels.
> **complete h.,** an indirect inguinal h. in which the contents extend into the tunica vaginalis.
> **concealed h.,** a h. not found on inspection or palpation.
> **congenital diaphragmatic h.,** Bochdaleck's h.; absence of the pleuroperitioneal membrane (usually on the left) or an enlarged foramen of Morgagni which allows protrusion of abdominal viscera into the chest.

The alphabetical ordering of *The Chambers Dictionary* as GPD is based on a word-by-word arrangement, whereas *Stedman's Medical Dictionary* as LSPD combines both alphabetical and thematic ordering in a single article. The lemma **hernia** is used here as an umbrella under which all types of hernia are treated

despite their alphabetical order. The thematic ordering is mostly used in LSPDs and rarely in GPDs.

The ordering typology can be summarized as follows: In GPDs the central list is established in alphabetical order, which implies that the lemmata, sub-lemmata, and even article stretches are arranged from A to Z. But LSPDs can present other arrangements according to the purposes decided on in the data distribution programme. It can be an alphabetic, thematic, or semantic ordering.

In a thematic ordering, lexical items are grouped in families or genera. Items such as *beauty*, *elegance* and *peace* can be grouped in a category *human being*, while words such as *cholera*, *AIDS*, and *malaria* can be incorporated in a disease category.

The semantic ordering is more often used in a synonym dictionary; terms are arranged according to the field and a concept covered by the lexical items. In these types of dictionaries, the treatment consists of a listing of a paradigm of synonyms. As a result, the dominant term is chosen as lemma and the other terms become synonyms, always arranged according to the semantic aspects of the lexical items.

Microstructure

In this section, it will be shown how the treatment of a given lexical item differs in different types of dictionaries. As a technical dictionary, the LSPD uses "technical" language when it refers to a specific register of the language dealing with a more or less hermetic vocabulary as well as its meanings. For example, a lexical item like *absence* has, as has already been shown above, different meanings, depending on the type of dictionary. Once more the different definitions of *absence* proposed by the *Longman Dictionary of the English Language* and *The New Shorter Oxford English Dictionary* can be considered.

The first dictionary treats the lemma by means of three definitions, the purpose being to give explanations of the lexical item *absence* in such a way that the user can easily retrieve "good" information. But if this treatment is observed carefully, it appears that the three definitions and senses combined in this article all refer to an adjective or a noun which should be explained somewhere else in the dictionary for a complete understanding. The definitions/senses **1** and **2**, for instance, define *absence* in terms of the adjective *absent*. In this particular case, the user cannot understand the meaning of *absence* if he/she does not know what the adjective *absent* means. To understand definition/sense **3** the user must know the meaning of the nouns *want* and *lack*. This type of definition results in a so-called "circular definition".

The second dictionary also gives three definitions/senses. Sense **1** is defined by means of a direct meaning description without reference to the adjective *absent*. In sense **2**, the user must also, as in the first dictionary, be acquainted with the meanings of two nouns, namely *nonexistence* and *lack*. Moreover, it combines these direct meanings with transferred meanings in sense **3**,

referring to a state of someone whose thought/mind is not present. This usage of *absence* is a transposition of the sense by analogy.

As has been said above, three of the four legal dictionaries which can be considered technical (or LSP) dictionaries, namely *Bell's South African Legal Dictionary*, *Stroud's Judicial Dictionary of Words and Phrases* and *Osborn's Concise Law Dictionary*, explore the meaning of the lexical item *absence* in a very strict and specific way. To understand the meaning of *absence* in *Bell's South African Legal Dictionary* for instance, the user needs first to know to what "the Insolvency Act, No. 32 of 1916" refers. This necessitates a certain legal background. It implies that for the compilers of these dictionaries, the user is expected to have a certain legal knowledge, which can help him/her to appreciate the relevancy of these meaning descriptions. As a result, only users who are legal specialists will, in this case, obtain enough "good" information from *Bell's South African Legal Dictionary*. This restrictive approach may be broadened by bringing a somewhat semi-technical view of the lexical item *absence*, providing explanations for users who are both experts and neophytes.

Of all the technical dictionaries mentioned above, *Osborn's Concise Law Dictionary* especially offers a more accessible meaning description of *absence* than those of the other two: *Bell's South African Legal Dictionary* and *Stroud's Judicial Dictionary of Words and Phrases*. Its explanation summarized in one dictionary article, is both accurate and simple. It encompasses the legal time and circumstance which precede the declaration of absence as well as the legal procedure attached to it. This implies that in terms of user perspective, qualified lawyers, law students and neophytes can easily and quickly retrieve the information for which they are looking. In this sense, the given description of meaning is more specific. It can be assumed that the meaning of *absence* as presented in the GPDs, the *Longman Dictionary of the English Language* and *The New Shorter Oxford English Dictionary*, is already known to the users consulting the legal dictionaries.

Access structures

In many GPDs, data is presented to users according to certain priorities set up in advance by the lexicographer in order to satisfy their needs. This is also the case with LSPDs. Some aspects such as the degree of expertise in a given discipline, the level of the language used in the data, as well as the scale of accuracy in the treatment of meaning in a given dictionary are factors generally influencing the process of compilation, shaping the form of the lexicographic work and opening a way to satisfying the dictionary user.

Because technical terms designate a distinct concept, they have a nominating and a defining function. As a result, they are expected, as is suggested by Gläser (2000: 88), to be precise, unequivocal, unemotional, stylistically neutral, without evaluative connotations, and suitable for patterns of word-formation.

According to Bergenholtz, Tarp and Wiegand (1999), a distinction can be made between a simple data distribution structure focusing only on articles in the central list of a dictionary and an expanded data distribution focusing on the outer texts of a dictionary or some of these. In this article, with regard to examples of dictionary articles given below, only the first one (simple structure) is taken into account. Despite the user-driven approach, LSPDs are also dictionary-function orientated. Two functions can be distinguished (cf. Bergenholtz and Tarp 1995: 20-28):

(a)   communication-directed functions, which aim to assist the user in solving problems in the sphere of communication (cf. Gouws (2002: 56), as is the case with the lemma *heart* from *The Oxford Medical Companion*:

> **HEART.** The heart is the muscular pump which maintains circulation of the blood, situated in the thorax to the left of the midline. The right and left sides are separate functional units, each comprising a filling chamber, or atrium, and a much thicker and more muscular ventricle. The side receives deoxygenated venous blood returning from the tissues and pumps it into the lungs via the *pulmonary arteries. From here, after oxygenation by pulmonary ventilation, it returns to the left side and is pumped by the ventricle back into the tissues via the *aorta and the *arteries.

This treatment of the lexical item *heart* includes in the definition references to definitions of concepts related to the lemma **heart**, e.g. *aorta* and *arteries*.

(b)   knowledge-directed functions which intend to increase the user's knowledge with regard to a specific topic, as illustrated by the lemma *heart* from *Taber's Cyclopedic Medical Dictionary*:

> **heart** (hart) [AS. *heorte*]. A hollow, muscular, contractile organ, the center of the circulatory system. Its wall possesses three layers: the outer epicardium, a serous layer; the middle myocardium, composed of cardiac muscle; and the inner endocardium, a layer which lines the four chambers of the heart and covers the valves. The heart is enclosed in a fibroserous sac, the pericardium, the space between the pericardium and the epicardium forming the pericardial cavity. SEE: *circulation* for illus.

This treatment of the lexical item *heart* includes in the definition comments and descriptions of various concepts related to the lemma **heart**, notably *epicardium*, *myocardium*, *endocardium* and *pericardium*.

In LSPDs, articles with definitions conveying cultural and historical information are also included, as can be seen from the following examples from *The Oxford Medical Companion*:

> **COS** (now Kos), is an island in the Dodecanese, near the shores of Turkey, north of Rhodes. An early site of 'temple-healing' (see TEMPLE MEDICINE) is found in Asklepieia, dedicated to *Aesculapius. Cos was also the birthplace of *Hippocrates *c*. 560 BC.

> **BROCA,** PAUL (1824–80). French surgeon and anthropologist. He was a pioneer of *orthopaedics as well as *neurosurgery. In 1861 he showed that the left inferior frontal cerebral convolution (Broca's gyrus) was related to speech. He was one of the first to localize and *trephine for *brain abscess.

Lemmata such as these are only found in GPDs with an encyclopedic bias.

Conclusion

GPDs deal with the general signification of lexical items, while LSPDs are more technical, more specific in their treatment of meaning. The object of definition in both GPDs and LSPDs is not only a concept taken from the field of a special science but also a situation, the analysis of which does not demand expert knowledge and may be limited to separating the various moments or contexts succeeding one another in a specific situation (Doroszewski 1973). Definition is the ability to imagine vividly such an object or a situation and to transpose clearly this image into words. LSP or specialized dictionaries are different from and must be distinguished from limited dictionaries or restricted dictionaries even if they share some features. Where GPDs deal with more than one meaning, LSPDs are usually focused on only one unambiguous meaning. As far as dictionary functions are concerned, LSPDs contain both knowledge- and communication-directed functions within several dictionary articles.

Bibliography

Al-Kasimi, A.M. 1977. *Linguistics and Bilingual Dictionaries*. Leiden: Brill.

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

Bergenholtz, H., S. Tarp and H.E. Wiegand. 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Hoffman, L., H. Kalverkämper and H.E Wiegand (Eds.). 1999. *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research*: 1762-1832. Berlin: De Gruyter.

Bird, Roger (Ed.). 1983[7]. *Osborn's Concise Law Dictionary*. London: Sweet and Maxwell.

Brown, Lesley. 1993. *The New Shorter Oxford English Dictionary on Historical Principles*. 2 volumes. Oxford: Clarendon Press.

Burkhanov, I. 1998. *Lexicography: A Dictionary of Basic Terminology.* Rzeszów: Wydawnictwo Wyższej Szkoły Pedagogicznej.

Chambers. 1998. *The Chambers Dictionary*. London: Chambers.

Coulson, J. et al. (Eds.). 1978. *The New Oxford Illustrated Dictionary*. Sydney: Bay Books.

Crystal, D. 1997[2]. *The Cambridge Encyclopedia of Language.* New York: Cambridge University Press.

Doroszewski, Witold. 1973. *Elements of Lexicology and Semiotics*. The Hague/Paris: Mouton; Warszawa: PWN-Polish Scientific Publishers.

Garner, Bryan A. 1999[7]. *Black's Law Dictionary.* St. Paul, Minn.: West Group.

Geeraerts, D. 1987. Types of Semantic Information in Dictionaries. Ilson, R. (Ed.). 1987. *A Spectrum of Lexicography: Papers from AILA, Brussels 1984*: 1-10. Amsterdam: John Benjamins.

Gläser, Rosemarie. 2000. Should LSP Dictionaries also Include Professional Jargon and Slang? *Lexikos* 10: 86-98.

Gouws, R.H. 2000. *General Theory of Lexicography: An Introduction*. Unpublished M.Phil Class Notes. Stellenbosch: University of Stellenbosch, Department of Afrikaans and Dutch and Department of African Languages.

Gouws, R.H. 2003. Using a Frame Structure to Accommodate Cultural Data. Emejelu, J.D. (Ed.). *Éléments de lexicographie gabonaise. Tome II*: 54-69. New York: Jimacs-Hillman.

Hartmann, R.R.K. 1983. *Lexicography: Principles and Practice*. London: Academic Press.

Hartmann, R.R.K. and G. James. 1998. *Dictionary of Lexicography*. London/New York: Routledge.

Hensyl, William R. (Ed.). 1990[25]. *Stedman's Medical Dictionary*. Baltimore/Hong Kong/London/Sydney: Williams and Wilkins.

James, John S. 1971. *Stroud's Judicial Dictionary of Words and Phrases.* London: Sweet and Maxwell.

Lombard, F.J. 2000. *Course in General and Computer Lexicography*. Unpublished Training Course Notes. Stellenbosch: Bureau of the WAT.

Longman. 1984. *Longman Dictionary of the English Language*. Harlow, Essex: Longman.

Matthews, P.H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford/New York: Oxford University Press.

Mertvago, P. 1996. *Dictionary of 1000 French Proverbs*. New York: Hippocerne Books.

Milne, Alexander, Clifford Cooper and Brian D. Burne. 1951[3]. *Bell's South African Legal Dictionary*. Durban: Butterworth.

Robinson, Mairi (Ed.-in-chief). 1999. *Chambers 21st Century Dictionary*. Edinburgh: Chambers.

Singh, R.A. 1982. *An Introduction to Lexicography.* Mysore: Central Institute of India Languages (C.I.I.L).

Smit, M. 1996. *Wiegand's Metalexicography as a Framework for a Multilingual, Multicultural, Explanatory Music Education Dictionary for South Africa*. Unpublished D.Litt. Thesis. Stellenbosch: University of Stellenbosch.

Smit, M. 2000. Using Frames in Special-Field Lexicography: An Ethnomusicological Case Study. *Lexikos* 10: 173-188.

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

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

Thomas, Clayton L. 1977[13]. *Taber's Cyclopedic Medical Dictionary*. Philadelphia: F.A. Davis.

Van Schalkwyk, D.J. 1987. Professional Terminology in African Languages: Theory and Practice. *Logos* 7(2): 9-12.

Walton, John, Jeremiah A. Barondess and Stephen Lock (Eds.). 1994. *The Oxford Medical Companion*. Oxford/New York/Tokyo: Oxford University Press.

Wiegand, H.E. 1983. Ansätze zu einer allgemeinen Theorie der Lexikographie. Schildt, J. and D. Viehweger. 1983. *Die Lexikographie von heute und das Wörterbuch von morgen: Analysen — Probleme — Vorschläge*: 92-127. Berlin: Akademie der Wissenschaften der DDR, Zentralinstitut für Sprachwissenschaft.

Wiegand, H.E. 1992. Elements of a Theory towards a So-called Lexicographic Definition. *Lexicographica* 8: 175-289.

Wiegand, H.E. 1994. Zur Unterscheidung von semantischen und enzyklopädischen Daten in Fachwörterbüchern. Schaeder, B. and H. Bergenholtz (Eds.). 1994. *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 103-132. Tübingen: Narr.

Wiegand, H.E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband.* Berlin: De Gruyter.

Karakterisering van het *Etimologie-woordeboek van Afrikaans* (EWA)

Fons Moerdijk, *Instituut voor Nederlandse Lexicologie, Leiden en Universiteit van Amsterdam, Amsterdam, Nederland (moerdijk@inl.nl)*

Samenvatting: Zesendertig jaar na *Afrikaanse etimologieë* van Boshoff en Nienaber verscheen er in 2003 een nieuw etymologisch woordenboek van het Afrikaans: het *Etimologiewoordeboek van Afrikaans* (EWA). Tussen die twee woordenboeken bestaat een wezenlijk verschil. Het werk van Boshoff en Nienaber is vrijwel alleen toegankelijk voor specialisten en niet specifiek bedoeld voor een breed publiek. EWA is juist allereerst bestemd voor de leek die zich interesseert voor de oorsprong en de geschiedenis van de eigen taal. Daarnaast biedt het ook de etymoloog en historisch taalkundige een hoop gegevens en inspiratie voor nader onderzoek. Als etymologisch woordenboek bezit EWA een aantal bijzondere, eigen kenmerken. De selectie van de woorden is bepaald door het beoogde gebruikerspubliek. Als gevolg daarvan bevat EWA verscheidene typen woorden die in traditionele etymologische woordenboeken doorgaans veronachtzaamd worden, zoals afleidingen, samenstellingen, neologismen en gemeenzame spreektaalwoorden. De artikelstructuur wordt gekenmerkt door een hoge graad van uniformiteit en consistentie. Het accent ligt op de zogenaamde etymologia proxima, de etymologie waarbij men één stap teruggaat en de meest directe herkomst van het woord biedt. Voorts is speciale aandacht besteed aan polysemie: niet alleen de herkomst van de dominante betekenis van een woord wordt gegeven, maar ook die van zijn eventuele andere betekenis(sen). Benoemingsmotieven komen ook veelvuldig aan bod. In een eventuele tweede druk zouden opvallende verschillen in vorm of betekenis altijd verklaard of in ieder geval besproken moeten worden. Dat blijft nu nogal eens achterwege. Ook zou voorafgaand aan zo'n volgende editie meer onderzoek naar woordhistorisch materiaal uit primaire bronnen voor met name de 17de–19de eeuw verricht moeten worden. Het zou mooi zijn als deze eerste uitgave van EWA de aanleiding tot een dergelijk onderzoek zou vormen.

Sleutelwoorden: BENOEMINGSMOTIEF, ETYMOLOGIA PROXIMA, ETYMOLOGIA REMOTA, ETYMOLOGIE, ETYMOLOGISCH WOORDENBOEK, LEXICOGRAFIE, LEENVERTALING, MACROSTRUCTUUR, MICROSTRUCTUUR, POLYSEMIE, WOORDGESCHIEDENIS

Abstract: Characterisation of the *Etimologiewoordeboek van Afrikaans* (EWA). In 2003, thirty six years after *Afrikaanse etimologieë* by Boshoff and Nienaber, a new etymological dictionary of Afrikaans has been published: the *Etimologiewoordeboek van Afrikaans* (EWA). There is a substantial difference between these two dictionaries. The work of Boshoff and Nienaber is almost entirely accessible only to specialists and not specifically meant for the general public. EWA is first and foremost meant for the layperson who is interested in the origin and the history of his/her own language. In addition it also offers a lot of data and inspiration for further research to etymologists and historical linguists. EWA is an etymological dictionary with a number of special characteristics. The selection of the entry-words has been determined by the intended users. Consequently, EWA contains several kinds of words which are generally neglected in traditional etymological dictionaries, such as derivations, compounds, neologisms and colloquialisms. The entry structure is characterized by a high degree of uniformity and consistency. The focus is on the so-called etymologia proxima, the nearest etymology that goes one step back and provides the most

direct origin of a word. Furthermore, special attention has been paid to polysemy: not only the origin of a word in its dominant sense has been given but also that of its other eventual sense(s). Explanations of naming are also frequently indicated. In a future second edition significant differences in form or meaning should always be explained or certainly discussed. At present, these have rather often been neglected. Prior to such a subsequent edition, more research on word historical materials from primary sources especially for the 17th–19th century should also be done. It would be excellent if this first edition of EWA could lead to such an investigation.

Keywords: ETYMOLOGIA PROXIMA, ETYMOLOGIA REMOTA, ETYMOLOGICAL DICTIONARY, ETYMOLOGY, LEXICOGRAPHY, LOAN TRANSLATION, MACROSTRUCTURE, MICROSTRUCTURE, POLYSEMY, WORD HISTORY, WORD ORIGIN

1.   Inleiding

Zesendertig jaar na *Afrikaanse etimologieë* van Boshoff en Nienaber, gepubliceerd in 1967, heeft het Afrikaans er een nieuw etymologisch woordenboek bijgekregen: het *Etimologiewoordeboek van Afrikaans* (EWA). Op 5 september 2003 werd het op het bureau van het *Woordeboek van die Afrikaanse taal* (WAT) officieel gepresenteerd. Bij die gelegenheid hield ik een feestrede met de titel 'De bijzondere kenmerken van EWA'. Ik dankte dat voorrecht en genoegen aan mijn nauwe bemoeienis met het woordenboek. Tot driemaal toe was ik in Stellenbosch om er een opleiding etymologie te geven aan de beoogde EWA-redacteuren. Ik zorgde voorts voor het ontwerp van de microstructuur en vaste tekstmodellen, adviseerde bij de selectie van de bronnen en de op te nemen woorden, las en becommentarieerde de kopij. Ik voel mij dan ook zo'n beetje 'de Adam van EWA'.

    Dit artikel is een brede uitwerking van die rede die ik toen gaf. Het is geen recensie. Mijn nauwe betrokkenheid bij EWA maakt mij ongeschikt als recensent. Ik heb me ten doel gesteld om vanuit mijn perspectief de belangrijkste kenmerken aan te geven en te bespreken. Daarbij is mijn verbondenheid met EWA juist een voordeel. Ik heb kennis van adviezen, afspraken en beslissingen die aan bepaalde aspecten en werkwijzes ten grondslag lagen. Informatie daarover werkt verhelderend.

    Mijn karakterisering van EWA zal bestaan uit een algemene typering, een beschouwing van de macrostructuur, de selectie en de aard van de trefwoorden, met daarbinnen speciale aandacht voor de leenvertalingen, een behandeling van de microstructuur, waarbij ik de belangrijkste aspecten daarvan belicht, te weten de verhouding tussen nabije en verre etymologie, polysemie, benoemingsmotieven en dateringen. Vervolgens ga ik in op de rol van EWA voor het Nederlands en bespreek ik enige tekortkomingen, op te vatten als desiderata voor een eventuele tweede editie.

    Bij de genoemde presentatie van EWA kreeg een beperkt aantal mensen een cd-rom-versie ten geschenke. Ik behoorde tot die gelukkigen en maak op verscheidene plaatsen in dit artikel dankbaar gebruik ervan.

2.   Algemene typering

Etymologische woordenboeken zijn geen naslagwerken ter bevordering of ter controle van het eigen taalgebruik. Men slaat ze op voor enig woordhistorisch onderzoek, of men leest ze uit interesse in de herkomst en geschiedenis van de eigen woordenschat en de eigen cultuurhistorie die daarin weerspiegeld worden. Afhankelijk van deze functies kunnen we ze als volgt typeren:

(a)   uitsluitend naslagwerk voor wetenschappelijk onderzoek, dus voor etymologen en anderszins taalkundig geschoolden;

(b)   overwegend naslagwerk voor wetenschappelijk onderzoek, maar tot op zekere hoogte ook leeswoordenboek voor een breed niet taalkundig-geschoold lekenpubliek;

(c)   overwegend leeswoordenboek voor een breed lekenpubliek, maar tot op zekere hoogte ook geschikt als basis voor wetenschappelijk onderzoek;

(d)   uitsluitend leeswoordenboek.

Volgens deze indeling zou het genoemde werk van Boshoff en Nienaber een etymologisch woordenboek van het type (b) of misschien zelfs (a) zijn, terwijl EWA een vertegenwoordiger is van type (c). *Afrikaanse etimologieë* is vooral een naslagwerk voor etymologen en taalkundigen. Het is niet opgezet als leeswoordenboek voor een niet taalkundig geschoold publiek. EWA vormt in dit opzicht het spiegelbeeld van Boshoff–Nienaber. In de Inleiding bij het werk (blz. i) verwoordt Dirk van Schalkwyk de doelgroep van EWA als volgt: 'Die EWA is nie in die eerste plek vir die etimologies kundige persoon daargestel nie, maar wel vir die breë intellektuele lekepubliek. Omdat etimologiese woordeboeke interessante leesstof bied, is dit egter 'n inligtingsbron vir elke taallief-hebber.' Bedoeld in eerste instantie voor een breed lekenpubliek, biedt het daarnaast zeker ook etymologen en historisch-taalkundigen veel van hun gading. Het is opgezet en bewerkt met inachtneming van bestaande wetenschappelijke principes en methodes en gebaseerd op controleerbare bronnen. Mikkend op een breed publiek, is er zeker sprake van popularisering, maar dan wel van een popularisering waarbij aan wetenschappelijke eisen is voldaan (vgl. Seebold 1981: 264-265).

3.   Macrostructuur

Bij het ontwerpen van de macrostructuur, de lijst van geselecteerde trefwoorden en hun ordening, heeft de redactie zich haar beoogde brede lezerspubliek over het algemeen goed voor ogen gehouden. De trefwoordenlijst is niet, zoals in tal van traditionele etymologische woordenboeken, het resultaat van eenzij-

dige aandacht voor een keur van ongelede erfwoorden, substraatwoorden en oude ontleningen, het traditionele lievelingseten van de etymologische smulpaap. Leidende gedachte voor de selectie van die trefwoorden was niet: is het woord interessant voor de etymologische vakman, d.w.z. ongeleed en/of oud genoeg, maar: is het woord interessant voor de ontwikkelde leek, die in de herkomst van de woorden van zijn taal belang stelt? Is het een woord dat door zo iemand zal worden opgezocht? Door dat andere perspectief hebben ook tal van jonge en gelede woorden en zelfs behoorlijk wat woorden uit de informele spreektaal in EWA onderdak gevonden.

Samenstellingen en afleidingen zijn ten onrechte de stiefkinderen van de etymologie. Toch zijn ze van belang voor dateringen, voor een getrouwe weerspiegeling van de verhoudingen binnen de woordenschat, maar ook om de omvang van het leengoed van een taal vast te kunnen stellen. In dat perspectief is het zinvol vooral bij samenstellingen na te gaan of er sprake is van een volkomen inheems woord dan wel van een leenvertaling (= een woord dat wel is gevormd met inheemse elementen, maar die elementen zijn vertalingen van een uitheems woord). Het EWA bevat vele, vaak even vernuftige als amusante voorbeelden van leenvertalingen, met name uit het Engels. Ik geef ter illustratie de volgende greep uit het rijke aanbod: *agterbanker* naar Eng. *backbencher* (1910); *aptytwekkertjie* naar Eng. *appetiser* (1862); *bergfiets* naar Eng. *mountain bike* (2000); *blitsverkoper* naar Eng. *bestseller* (1889); *broekiekous* naar Eng. *pantyhose* (1963); *ontkleedans* naar Eng. *striptease* (1936); *padvark* naar Eng. *road hog* (1891); *ruk-en-rol* naar Eng. *rock-and-roll* (met een veelzeggende schertsende vervorming van *rock* in *ruk*!); *riller* naar Eng. *thriller* (1889); *sagtebal* naar Eng. *softball* (1926); *sjoebroekie* naar Eng. *hotpants* (1970); *skootrekenaar* naar Eng. *lap-top* (1984); *skopboks* naar Eng. *kickboxing* (1971); *terugflits* naar Eng. *flashback* (1916); *wensdenkery* naar Eng. *wishful thinking* (1932); *wetenskapfiksie* naar Eng. *science fiction* (1851). Een voorbeeld van een afleiding is *kreuner* naar Eng. *crooner* (een even vermakelijke als doeltreffende klankvervorming!).

Wat de behandeling van het verschijnsel leenvertaling betreft, zouden Nederlandse (en andere) etymologische woordenboeken een voorbeeld kunnen nemen aan het EWA. Wanneer we in EWA lezen dat zelfs Afrikaans *vlootblauw* een leenvertaling is van Engels *navy blue*, dan roept dat de vraag op: Hoe zit dat eigenlijk met het Nederlandse *marineblauw*? Het zou mij niet verbazen als bij nader en grootschalig onderzoek zou blijken dat veel voor inheems-Nederlands gehouden samenstellingen in werkelijkheid ook leenvertalingen van Engelse woorden zouden blijken te zijn.

Veel van deze leenvertalingen getuigen van de geestigheid en vindingrijkheid van de Afrikaanse taalgebruiker. Die eigenschappen zijn ook af te meten aan inheemse relatief jonge vormingen zonder buitenlands model, zoals bijvoorbeeld *moffiepropeller* voor 'strikdas', *moltrein* voor 'metro' en *soldoedie* 'soldate', een blend van de eerste lettergreep van *soldaat* en *doedie* 'jong meisie'.

Naast geschiktheid voor het beoogde lezerspubliek was er een tweede voorwaarde verbonden aan de selectie van de op te nemen woorden: zoveel mogelijk een weerspiegeling vormen van de diversiteit in herkomst en van de

interne gelaagdheid van de woordenschat van het Afrikaans als zodanig. Het aandeel van het Nederlands en het Engels daarin is groot en het ligt dan voor de hand dat beide talen ook in het woordenboek de boventoon voeren.

EWA bevat, zo leert de digitale versie, 8 054 alfabetisch geordende trefwoorden. Op die 8 054 trefwoorden moeten we enerzijds de verwijslemmata in mindering brengen, anderzijds moeten we rekening houden met een toename, omdat per woord voor meer betekenissen verschillende herkomsttalen vermeld kunnen zijn. In 6 006 gevallen is de herkomsttaal voor woorden én betekenissen aangegeven. Juist omdat voor nogal wat betekenissen meer herkomstmogelijkheden worden geboden, is het niet zinvol het aantal keren dat talen als herkomsttalen worden genoemd, in percentages om te zetten. Bovendien valt uit de data op zich nog niet af te leiden hoe zeker of onwaarschijnlijk de etymologische voorstellen zijn. Wel geven de aantallen die je op de cd-rom krijgt voor de verschillende herkomsttalen een indicatie van hun spreiding en relatieve gewicht. Het zal geen verbazing wekken dat het Nederlands met 4 523 treffers koploper bij uitstek is. Het Engels is met 1 727 (daarbij zijn alle variëteiten samengenomen, waaronder Zuid-Afrikaans Engels met 27) een sterke tweede. Ze worden op ruime afstand gevolgd door Maleis 50 (inclusief 3x 'Kaapse Maleis'), Duits 50, Khoi 43, Frans 40, Zulu 39, Nama 36, Xhosa 31, Portugees 21 (inclusief 6x Maleis Portugees), Nguni 18, Sotho 15, Tswana 12, Indonesisch (inclusief Indonesisch Nederlands, Oost-Indisch Nederlands) 7, San 6 en Vlaams 3. Arabisch, Korana, Nederduits, Pedi, Spaans, Swahili, Tamil, worden twee keer genoemd en Chinees, Hebreeuws, Hindi, Italiaans, Oerdoe, Russisch, Shona en Venda één keer.

Bij de woorden en betekenissen waarvoor geen herkomsttaal is aangeduid, hebben we te maken met Afrikaanse woorden en betekenissen. Hun aantal moet dan in de duizenden lopen. Voor het aantal betekenissen bij polyseme leenwoorden kunnen we een redelijke indicatie krijgen, omdat ze binnen het sjabloon meestal de vaste formule 'het in Afr. self ontwikkel' meekregen. Het zijn er 525. Onder de inheemse vormingen nemen samenstellingen (1 018) en afleidingen (300), ook geheel volgens de verwachting, een prominente plaats in.

Ofschoon men zich duidelijk op het Standaard-Afrikaans heeft geconcentreerd, zijn ook woorden uit de 'geselstaal' opgenomen (391) en zijn in zeer bescheiden mate ook enige streektaalwoorden (36) opnamewaardig bevonden. De aanzienlijke invloed van het Engels op met name de informele omgangstaal (Ponelis 1998: 122-123 en 126) heeft zijn reflex in woorden als *grênd*, *grill*, *groan*, *guarantee*, *neklis*, *newwer*, *nip* (< informeel Z.-A. Eng. *nip*(*s*)), *nok* (< Eng. *knock*), *quit*, *ref* (< Eng. *ref*(*eree*)), *useless*.

Veel woorden die direct en herkenbaar uit het Nederlands komen en waarover etymologisch niets meer dan dat te zeggen viel, zijn niet opgenomen. Ze konden echter niet allemaal weggelaten worden. Dat zou immers tot gevolg gehad hebben dat het Nederlandse aandeel in de opbouw van de Afrikaanse woordenschat te kort gedaan zou worden. Uit oogpunt van representativiteit heeft wel een zeker aantal ervan een plaats in het woordenboek gekregen. Een

gevolg van die praktijk is, dat samenstellingen waarover wel een etymologisch interessante of aardige opmerking te maken viel, wél opgenomen kunnen zijn, maar dat de ongelede woorden zelf die er een bestanddeel van vormen, er níét in staan. Zo bevat EWA een samenstelling *kothuis* met de informatie dat dit woord door volksetymologie is ontstaan uit Engels *cottage*, maar naar artikelen voor de woorden *kot* en *huis* zal men tevergeefs zoeken. We vinden er wel *melk-baard*, *botterkop*, *koubeitel*, *kouekoors*, maar niet *melk*, *kop*, *beitel* en *koors*.

Er zijn ook nogal wat woorden waarvan de etymologie zo summier is, dat ze overbodig aandoen. Ponelis (2002: 33) somt er een aantal op. Men moet echter niet te snel oordelen tot overbodigheid. Allereerst niet, omdat er, zoals we zagen, bij nogal wat inheems lijkende samenstellingen toch sprake kan zijn van leenvertalingen. Voorts worden deze 'voor zich sprekende' woorden interessanter als ze vanuit het Afrikaans in een andere taal zijn ontleend. Zo valt er weliswaar niets spectaculairs te beleven aan de vermelding dat *huilboom* een samenstelling is van *huil* en *boom*, maar opname is toch te rechtvaardigen vanwege het feit dat het woord door het Zuid-Afrikaans Engels is overgenomen. Overname van Afrikaanse woorden door het Zuid-Afrikaans Engels is relevant (EWA bevat 661 gevallen daarvan). Het wijst op zekere 'populariteit' van het Afrikaanse woord. Ook de aanwezigheid van dateringen kan al een rechtvaardiging voor opname zijn. Dat *gewas* uit Ndl. *gewas* stamt, is etymologisch allerminst opzienbarend, maar via de dateringen voor betekenissen van het Nederlandse woord als 'ongeveer 1600' en '1608' krijgen we toch de aanwijzing dat het woord tot het vroege Hollands-Afrikaans behoord zal hebben.

Als er echter buiten de vorming op zich geen relevante feiten voor de woordgeschiedenis te vermelden zijn, zouden transparante samenstellingen en afleidingen buiten de deur gehouden worden, zo was met de redactie afgesproken. Inbrekers hou je blijkbaar altijd. Dat is bijvoorbeeld aan de orde bij een artikel als dat voor *versdrama*, waarvoor alleen 'drama wat in versvorm geskryf is' als betekenis, en 'samestelling van *vers* en *drama*' als etymologie wordt gegeven. En bij *erdewerk* wordt naast de opgave van twee betekenissen slechts vermeld dat het woord een samenstelling is van *erd* en *werk*. Er is geen datering, geen doorgave naar Zuid-Afrikaans Engels noch andere informatie die aan het woord enig etymologisch belang had kunnen geven. Dergelijke woorden hadden inderdaad beter weggelaten kunnen worden ten gunste van woorden die er etymologisch meer toe doen.

Spreiding over de woordsoorten geeft het volgende beeld: zelfstandig naamwoord 5 685, werkwoord 1 238, bijvoeglijk naamwoord 832, bijwoord 271, tussenwerpsel 219, telwoord 19, voorzetsel 18, voegwoord 17, voornaamwoord 15, lidwoord 2.

4.   Microstructuur

De microstructuur, de informatie die binnen het artikel over de trefwoorden wordt gegeven en de ordening daarvan, valt op door uniformiteit en gestaalde consequentie. Die zijn allereerst het gevolg van een vooraf ontworpen bewer-

kingssjabloon, waarin de velden voor de diverse informatiecategorieën zijn vastgelegd: lemma, woordsoort, wisselvorm, label, betekenisnummer, definitie en etymologie. Bij 'etymologie' zijn er verschillende velden voor 'nabije etymologie' en 'verre etymologie' (zie daarvoor de volgende paragraaf), met daarbinnen velden als benoemingsmotief, bepaalde klank- en vormingsprocessen en bij ontlening: brontaal, bronwoord met zijn betekenis (als die afwijkt van die van het leenwoord). Bij woorden en betekenissen worden binnen het artikel ook hun vroegste dateringen vermeld, mits beschikbaar.

Maar er is meer. De woordenschat van het Afrikaans is in een aantal hoofdcategorieën te verdelen. Voor die categorieën en hun onderverdelingen zijn van tevoren vaste tekstmodellen ontworpen. Ook die werkwijze stond garant voor een hoge uniformiteitsgraad.

Van veel belang voor de opvallende interne consistentie van het EWA was ten slotte de aanstelling van een z.g. 'tegniese redakteur', Gerhard J. van Wyk, die verantwoordelijk was voor de definitieve redigering, de controle op de nakoming van de gemaakte afspraken, en de eindcontrole van de zetproeven.

De informatie in de artikelen is duidelijk gepresenteerd in goed leesbare, volle tekst, die niet, zoals te vaak bij het werk van Boshoff en Nienaber, is samengeperst als een sloopauto.

4.1   Het onderscheid tussen etymologia proxima en etymologia remota

In EWA wordt een duidelijk onderscheid gemaakt tussen de z.g. etymologia proxima en de etymologia remota. De etymologia proxima is de etymologie waarbij het directe etymon voor een woord wordt gegeven. Je gaat slechts één stap terug in de herkomstontwikkeling. De etymologia remota, de verwijderde etymologie, betreft de verdere stappen terug tot eventueel de uiteindelijke oorsprong, het einde van de etymologische evolutieketen. Die twee vormen van etymologie moeten goed uit elkaar gehouden worden. EWA geeft bij het eerste artikel *pen* 'veer van 'n vogel' en 'schrijfwerktuig' de volgende etymologie:

Uit Ndl. *pen* (Mnl. *penne*).
Mnl. *penne* uit Latyn *penna* 'veer'.
Eng. *pen*.

Het bestanddeel 'Uit Ndl. *pen* (Mnl. *penne*)' is hier de etymologia proxima, het gedeelte 'Mnl. *penne* uit Latyn *penna* 'veer'. Eng. *pen*' is de etymologia remota, met vermelding van een vorm in een verwante Germaanse taal.

De nadruk ligt in EWA op de directe etymologie. De verwijderde etymologie wordt normaal slechts in hoofdlijnen geschetst en ontbreekt vaker waar die voor de lekengebruiker niet interessant of te gecompliceerd werd geacht. Een bijzonder markante karaktertrek van EWA is, dat beide soorten etymologieën ook typografisch zijn onderscheiden. De verwijderde etymologie komt achteraan in een nieuwe alinea en in een kleiner lettertype, een vernieuwend

procédé, waarbij beide vormen van etymologie terecht en helder uiteengehou-
den worden.

Een consequentie van de nadruk op de etymologia proxima is, dat Ger-
maanse achtergronden niet zo aan bod komen als bijvoorbeeld in Nederlandse
etymologische woordenboeken voor oude erfwoorden het geval is. EWA bevat
geen stelselmatige opsomming van externe verwanten in andere Germaanse
talen in de diverse stadia van hun ontwikkeling. Hier ligt een belangrijk onder-
scheid met *Afrikaanse etimologieë*. Wel treft men met zekere regelmaat naast de
standaardopgave van de Nederlandse familieleden een vermelding aan van
huidige verwanten in belangrijke Germaanse talen als het Engels en het Duits.
De ultieme gang naar Indo-europese wortels wordt niet gemaakt, laryngalen
en substraten vindt men er niet in terug. De vakman blijft hiervoor aangewezen
op de etymologische woordenboeken van het Nederlands en andere Germaan-
se talen. Invoering in EWA zou overigens slechts het zoveelste geval betekend
hebben van klakkeloze overname van dergelijke gegevens uit eerdere etymolo-
gische woordenboeken. Bovendien, voor kritische verwerking daarvan is nog
altijd een etymologische kennis van het Germaans en Indo-europees vereist,
die bij de EWA-redactie nu eenmaal (nog) niet aanwezig was.

Afwijkingen van de algemene lijn komen wel voor. Zo wordt *mielie* 'graan-
soort, of die kop, pit of saad van hierdie plant' teruggevoerd op Ndl. *milie* (al
Mnl.) 'Turkse koring'. Dit is de directe etymologie, waarmee volstaan had kun-
nen worden. Van dit Ndl. *milie* wordt echter vervolgens aangegeven dat het is
ontleend uit Portugees *milho* 'manna, mielie', dit weer uit Latijns *milium*, dat
waarschijnlijk verband houdt met Latijns *mille*, mv. *mil(l)ia* 'duizenden'. Erg ver
terug gaat ook de etymologie van *frikkadel*. Na de directe etymologie 'uit Ndl.
frikadel, *fricadel*' wordt nog de volgende verwijderde etymologie gepresenteerd:
'Ndl. *frikadel* uit Fr. *fricadelle* uit Provensaals *fricadel*, wat teruggaan op Gallo-
Romaans **frigicare*, intensief van Latyn *frigere* 'rooster''.

De neiging om toch meer informatie bij de verwijderde etymologie te ge-
ven is zeer begrijpelijk. Vaak zitten daar zeer interessante, soms zelfs spectacu-
laire, elementen in, waarvan men terecht aanneemt dat ze ook boeiend zijn
voor de lekenlezers. Men wil hun die dan ook niet onthouden. Immers, zoals
Malkiel (1976: 20-21) stelt, vanuit zuiver theoretisch-wetenschappelijk stand-
punt mag juist één stap terug de perfecte temporele reikwijdte zijn, commer-
cieel kan dat wel eens niet zo verstandig zijn, omdat de gebruiker nu eenmaal
graag mee teruggevoerd wordt naar een punt waar zijn intellectuele honger
gestild kan worden. De nabije etymologie is nu eenmaal vaak minder presti-
gieus dan de verre etymologie. Vandaar dat veel etymologische woordenboe-
ken zoveel zeer gesofisticeerde spatio-temporele bouwsels met veel overtollig-
heden (Malkiel 1976: 20-21) bevatten. Dat laatste heeft EWA in ieder geval
zoveel mogelijk willen vermijden. De verhouding tussen het Nederlands en het
Afrikaans is in dit opzicht wel te vergelijken met de verhouding tussen het
Latijn en de Romaanse talen. Etymologische woordenboeken van de Romaanse
talen moeten de geschiedenis van de woorden tussen het Latijn en het huidige

tijdvak vastleggen, maar niet de geschiedenis van de Latijnse woorden vóór de Romeinse tijd. Voor Frans *chien* volstaat opgave van herkomst uit Latijns *canis*, maar de herkomst dáárvan hoeft feitelijk niet meer gegeven te worden (Bloch–Von Wartburg 1994: xx).

4.2   Betekenissen

Etymologische woordenboeken moeten inzicht geven in de lexicale opbouw van een taal. Daartoe behoort niet alleen de verklaring van de herkomst van de woorden, maar ook die van de betekenissen. Op betekenisgebied laten de traditionele etymologische woordenboeken het doorgaans echter schromelijk afweten. Men doet alsof woorden slechts één betekenis hebben. En indien men die goed bekend achtte of als die niet nodig was om niet-verwante woorden met dezelfde vorm uit elkaar te houden, dan werd zelfs die ene betekenis niet eens gegeven. Dat is wat het Nederlandse taalgebied betreft bijvoorbeeld het geval in het vermaarde *Etymologisch Woordenboek der Nederlandsche Taal* van Franck, Van Wijk en Van Haeringen en het *Nederlands Etymologisch Woordenboek* van Jan de Vries. Het *Etymologisch Woordenboek* van Van Dale geeft bij ieder trefwoord nog wel één betekenis, maar plaatst die tussen rechte haken. Het zal wel niet zo bedoeld zijn, maar dit geeft betekenissen toch ook het aanzien van bijkomstigheid. Een gunstige uitzondering vormt het nieuwe, grote *Etymologisch Woordenboek van het Nederlands* (EWN), waarvan onlangs onder hoofdredacteurschap van dr. Marlies Philippa, met prof. dr. Frans Debrabandere en dr. Arend Quak, het eerste deel verscheen. Daarin wordt voor het eerst bij verscheidene woorden systematischer aan de belangrijkste betekenisontwikkelingen aandacht besteed.

Binnen de moderne etymologie overheerst thans wel de opvatting, dat etymologische woordenboeken niet alleen de herkomst maar ook de belangrijkste betekenisontwikkelingen moeten vastleggen (vgl. Pfeiffer 1993: VIII-IX, Pfister 1980: 1, Hiersche 1993: 65 en 68-69, Seebold 1981: 264-265). Het EWA sluit daarbij aan en geeft ook de semantiek waar ze etymologisch recht op heeft. Heeft een woord meer betekenissen, dan behandelt het die consequent allemaal. Het geeft het ontstaan van verschillende betekenissen op verschillende tijdstippen aan, vermeldt met nadruk welke betekenissen het Afrikaans zelf ontwikkeld heeft en welke het aan andere talen, het Nederlands en het Engels voorop, heeft ontleend. EWA levert bijvoorbeeld voor het woord *mol* naast de bekende knaagdierbetekenis ook de betekenissen 'persoon wat ondergronds op 'n donker plek werk', 'persoon wat bysiende of kortsigtig is' en 'persoon wat as dubbelagent optree'. Van ieder van die betekenissen wordt aangegeven waar ze vandaan komt. In de primaire betekenis 'ondergronds knaagdier' gaat Afrikaans *mol* terug op Nederlands *mol*. In de betekenis 'persoon wat as dubbelagent optree' is Afrikaans *mol* ontleend aan Engels *mole* (1925). De twee andere betekenissen 'het in Afr. self ontwikkel'. Nederlandse etymologische woordenboeken geven slechts de etymologie voor *mol* 'ondergronds knaagdier'.

Door zijn etymologische aandacht ook voor andere betekenissen dan alleen de basisbetekenis laat EWA ons aan de ene kant goed zien welke betekenisontwikkelingen het Afrikaans zelf heeft doorgemaakt, aan de andere kant presenteert het betekenisgegevens die tot uitbreiding of bijstelling van onze kennis van het oudere Nederlands zullen leiden. Het Afrikaans *afval* dankt zijn eerste betekenis 'overtollig voedsel of vlees dat weggegooid wordt' aan het Nederlands, maar kent als tweede, inheems ontwikkelde betekenis: 'kop, pootjies en pens van 'n geslagte dier as gereg', een vorm van culinaire recycling die ons Nederlanders (gelukkig) vreemd is (het moet toch een smakelijk gerecht zijn, want in die tweede betekenis is *afval* zelfs tot in het Zuid-Afrikaans Engels doorgedrongen). Afrikaans *asem* heeft de betekenissen 'adem, ademhaling, uithoudingsvermogen' van gewestelijk Ndl. *asem,* maar heeft daar zelf de betekenis 'iemand om 'n taak te verrig' bij ontwikkeld. *Koerasie* gaat terug op Nederlands *courage* (1548; zelf een ontlening aan het Frans) in de betekenis 'moed'. Het Afrikaans heeft daar twee eigen betekenissen aan toe weten te voegen: 'manhaftigheid wat deur die gebruik van sterk drank aangewakker word' en 'seksuele viriliteit of oormoed'. In de kindertaal heeft *piepie* er naast een primaire betekenis 'urine' of 'urineren', betekenissen die berusten op Nederlands *piepie* (vooral in de verbinding *piepie doen*), een tweede, inheems ontstane betekenis 'penis' bijgekregen.

Soms worden door EWA betekenissen ten onrechte als inheems beschouwd. Afrikaans *kuier* vinden we terug met vier betekenissen: 'rustig wandel', 'besoek aflê, sonder om te oornag', 'as gas by iemand tuisgaan', en in de informele spreektaal 'gesellig verkeer, fuif'. Volgens EWA zou alleen de betekenis 'kuieren, rustig wandelen' op het Nederlands teruggaan. Het WNT heeft twee artikelen *kuieren*: 'op zijn gemak wandelen' en 'praten, kouten'. Ze verschenen in een aflevering uit 1909. Gegevens uit dialectwoordenboeken die later verschenen, konden daar dus nog niet in verwerkt zijn. Vanzelfsprekend was het voor de redactie van EWA onmogelijk om alle Nederlandse dialectbronnen na te pluizen. Zij moest zich vrijwel uitsluitend op het WNT en Van Dale baseren. Daardoor moest men wel de informatie missen die het *Woordenboek der Zeeuwse Dialecten* over *kuieren* geeft: naast de wandelbetekenis vinden we daar ook: 'op bezoek gaan, een uitgangsdag hebben'. De Afrikaanse betekenis 'op bezoek gaan' is niet inheems, maar moet ook gerekend worden tot het Hollands-Zeeuwse erfgoed in het Afrikaans.

4.3   Benoemingsmotieven

Binnen het bewerkingssjabloon is een apart veld ingeruimd voor het benoemingsmotief, de op vorm en betekenis gebaseerde verklaring van de naamgeving. EWA besteedt daar volop en consequent aandacht aan. In het tekstmodel is hiervoor een standaardformule ontworpen, die systematisch aanvangt met 'so genoem omdat …'. Liefelijk is *trompoppie* 'majorette': 'Die meisie word so genoem omdat die musiek waarop opgetree word, deur slaginstrumente ver-

skaf word'. Van een bijna dichterlijk gehalte is de verklaring bij *lugkasteel*: 'so genoem omdat die ongegronde verwagting gevisualiseer word as 'n kasteel wat as 't ware in die lug dryf en nie op iets konkreets rus nie'. Aansprekend is het benoemingsmotief van *lawaaiwater* 'sterk drank, veral brandewyn': 'so genoem omdat 'n persoon wat hom aan sterk drank vergryp, geneig is om lawaaierig te raak'. Niet bepaald vleiend voor bejaarden is de verklaring bij *ouvolk* 'groot klipsalmander': 'so genoem omdat die vel uit harde skubbe bestaan en omdat die reptiel vir lang tye bewegingloos in die son sit. Sowel sy voorkoms as sy gedrag kom dus ooreen met dié van 'n ou mens'. Niet te versmaden zijn voorts: *advokaat* 'drank uit brandewyn, eiers, suiker en gerasperde neutmuskaat berei': 'so genoem omdat die drank volgens 'n ou bygeloof bevorderlik sou wees vir die advokaat se stem en om sy brein te verskerp', *balbyter* 'soort mier met 'n groot kop': 'so genoem omdat sulke miere by veral seuns se kortbroeke inkruip wanneer hulle plat op die grond sit en hulle dan rondom die testikels byt', *fransmadam* 'kleinerige, lekker seevis': 'so genoem omdat die vis se blink en groot bruin oë aan dié van 'n Franse dame herinner'.

Soms wordt bij ontleningen het benoemingsmotief van het woord in de brontaal besproken. Dat is bijvoorbeeld het geval bij *malmok* 'seevoël wat in die suidelike oseane voorkom'. Dit woord zou zijn oorsprong hebben in Nederlands *malmok*, een samenstelling van *mal* 'dwaas' en *mok* 'meeuw' (1685) met de betekenis 'noordelike stormvoël'. Het benoemingsmotief dat wordt gegeven, geldt niet het Afrikaanse, maar het Nederlandse woord en luidt: 'Die Hollanders wat hierdie seevoëls in Groenland teëgekom het, het hulle so genoem omdat die voëls hulle weens hul 'domme malligheid' so maklik laat doodslaan het'. Afrikaans *kugel* 'jong, bedorwe, welgestelde, veral Joodse meisie wat materialisties en ligsinnig is en geaffekteerd deur die neus praat' is afkomstig uit Zuid-Afrikaans Engels slang *kugel* (1970), dat op zijn beurt te vereenzelvigen valt met Engels *kugel* 'niervetpoeding gemaak van noedels, aartappels of brood, soms met die byvoeging van rosyne'. Het benoemingsmotief dat wordt gegeven heeft dan betrekking op het Zuid-Afrikaans Engelse woord en wordt heel plastisch als volgt verwoord: 'Die meisie word wsk. so genoem omdat sy daarmee gereduseer word tot 'n styselrige poeding, 'n vergelyking wat haar oppervlakkigheid en aangeplaktheid weerspieël'. *Fliek* 'bioskoop' is ontleend aan Australisch Engels *flick* (1926). Eng. *flick* 'bioscoop' is ontstaan uit *flickers* 'rolprente', zo genoemd 'vanweë die flikkerende ligte tydens 'n vertoning'. Hier geldt het benoemingsmotief het Engelse woord.

Het benoemingsmotief is regelmatig vermengd met interessante en/of anekdotische cultuurhistorische bijzonderheden. Zo luidt de verklaring bij *boereraat* 'medisyne, meestal kruie-aftreksels, soos in die ou dae gebruik toe dokters skaars was': 'so genoem omdat boere geskiedkundig ver van dokters af was en dus dikw. hulle eie rate teen siektes moes vind'. En bij *gentoe* (*jento*, *jentoe*, *jintoe*) 'prostituut' lezen we: 'So genoem na 'n skip, die Gentoo, wat in die middel van die 19de eeu aan die Kaap aangedoen het met 'n groep vroulike passasiers wat prostitute geword het (de redacteur voegt daar heel diplomatiek

aan toe: 'Daar is baie teenstrydighede oor waar die skip vandaan gekom het en wat die herkoms van die vroue was'). In dat perspectief ontmoeten we ook historische personen. Bij *jakopewer* 'rooi of oranjekleurige seevis', lezen we dat het door volksetymologie is ontstaan uit Ndl. *jacob-evertsen* (1648). Het benoemingsmotief luidt: 'so genoem omdat die vis se huid aan die rooi, gevlekte puisiegesig van 'n Ndl. admiraal uit die 17de eeu, Jacob Evertsen, herinner'. *Klaaslouwbos* is een bepaald soort gewas. Als benoemingsmotief vernemen we: 'vermoedelik so genoem omdat ene Klaas Louw toegelaat het dat die plant op sy lande vervuil, en sodoende daartoe bygedra het dat dit 'n lastige onkruid geword het'.

Het benoemingsmotief kan ook een bepaalde betekenis betreffen. *Fiskaal* komt voor in de betekenissen 'gerechtsambtenaar' en 'vogel met 'n V-vormige wit rugmerk'. Bij die tweede betekenis wordt meegedeeld dat de vogel zo wordt genoemd 'omdat die voël, soos die amptenaar, insamelingswerk doen deur insekte en klein diertjies as 'n kosvoorraad te vergaar, of omdat die voël net so deur kleiner voëltjies en diertjies gevrees word as wat die beampte tydens die VOC-bewind deur die koloniste gevrees is'. De betekeniskloof tussen *kriek* 'krekel' en *kriek* 'horloge' wordt gedicht met de toelichting: 'wsk. so genoem omdat die eentonige tikgeluid daarvan herinner aan die geluid wat 'n kriek maak'.

Soms is het gegeven benoemingsmotief dubieus. Voor *krukker* 'iemand wat nie meer fiks of vaardig genoeg is om op sportgebied te kompeteer nie' stelt EWA: 'so genoem omdat die persoon weens beserings, dikw. dié wat die gebruik van 'n kruk noodsaak, nie meer aan 'n bepaalde sportsoort kan deelneem nie'. Hier lijken we toch eerder uit te moeten gaan van een afleiding met *-er* van *kruk* en van verbinding met de betekenis 19) die het WNT daarvoor geeft: 'Als minachtende benaming voor iemand die "niet meekan", die in zekere opzichten ten achteren of in een bepaald vak (en derg.) onbedreven is, enz.; een sukkel. Soms ook medelijdend of minachtend gebezigd met betrekking tot gezondheid of lichamelijken toestand, zelden en thans nooit met betrekking tot het uiterlijk.'

4.4   Dateringen

Dateringen worden gegeven bij woorden en bij afzonderlijke betekenissen. Men treft ze vooral aan bij Nederlandse en Engelse woorden. Bij Afrikaanse woorden wordt in tal van artikelen met het bronjaartal aangegeven in welke woordenlijst of welk woordenboek van het Afrikaans ze voor het eerst zijn opgetekend, bij sommige worden vroegste dateringen vermeld die men uit secundaire bronnen, boeken en tijdschriftartikels heeft overgenomen.

Dateringen voor Nederlandse woorden zijn voornamelijk gehaald uit het WNT. Dat is een zeer gecompliceerd woordenboek. Bovendien wordt het vinden van dateringen bemoeilijkt, doordat in de oude delen van dit woordenboek (tot 1942) de citaten niet gedateerd zijn. Ook Nederlandse taalkundigen

zelf slagen er niet altijd in bij de juiste gegevens uit te komen. Voor Nederland-se woorden zullen bovendien, nu de cd-rom van het WNT daarvoor geraad-pleegd kan worden, nog behoorlijk wat vroegere dateringen gevonden gaan worden. Het eerste deel van het EWN vertoont die tendens al. Nogal wat dateringen zullen dan ook met enig voorbehoud bezien moeten worden.

Van de ruim 4 000 keren dat een herkomstdatum voor een Nederlands woord of Nederlandse betekenis is opgegeven, komt ongeveer de helft uit de 16de en de 17de eeuw; 1 524 ervan zijn al in het Middelnederlands bekend.

Problematisch was voor het Nederlands de datering van dialectwoorden. Pas redelijk laat in de negentiende eeuw verschijnen de eerste dialectwoorden-boeken waarin ze worden opgetekend. Die eerste optekening is natuurlijk niet hun ontstaansjaar. In werkelijkheid zijn ze (veel) ouder. Dergelijke woorden toch voorzien van het publicatiejaar van het dialectwoordenboek, zou een ver-tekening van die werkelijkheid hebben gegeven. Daarom heb ik de redactie geadviseerd in die gevallen geen datering te geven. Dat advies heeft men zich in het algemeen goed ter harte genomen, maar soms is het veronachtzaamd. Bij *swernoot* wordt gewezen op een Nederlands, gewestelijk (oostelijk) *swernood*, nevenvorm van *zweernood*. Dit laatste woord krijgt een datering 1870 mee. Het is afkomstig uit een dialectwoordenboek van Schuermans. Heel in het alge-meen kan als waarschuwing gelden, dat men jonge dateringen bij gewestelijke Nederlandse woorden (ruwweg: vanaf de paar laatste decennia van de 19de eeuw) moet wantrouwen; ze zijn normaliter (veel) ouder.

Aarzeling was er bij mij ook met betrekking tot het geven van zeer vroege dateringen bij tal van Engelse woorden (bijvoorbeeld: *groan* 'ongeveer 700's'; *quit* 'ongeveer 1440'; *necklace* 1590; *grill* 1668). Grofweg kan men stellen dat het Engels vóór de negentiende eeuw nauwelijks of geen invloed op de woord-voorraad van het Afrikaans gehad zal hebben. Daarom leken dergelijke vroege dateringen, die betrekking hebben op het eerste voorkomen in het Engels zelf, minder zinvol. Er is toch besloten ze te geven. Het zijn vaak woorden uit een dagelijkse levenssfeer en het is wel tekenend voor de algemene, zich uitbrei-dende invloed van het Engels dat ze daarbinnen worden overgenomen. Juist aan hun vroege jaartal voor het Engels zijn ze in die hoedanigheid goed te herkennen.

Voor veel dateringen van Afrikaanse woorden geldt, wat we boven te kennen gaven voor de dateringen van Nederlandse dialectwoorden: ze hebben betrekking op de eerste optekening in de lexicografische traditie, maar het gaat daarbij om woorden die in de gesproken volkstaal al (veel) langer bestaan en waarvoor dus een (veel) hogere ouderdom moet worden aangenomen. Om die vast te kunnen stellen en om de ontwikkelingen naar vorm en betekenis door de eeuwen heen na te gaan, is onderzoek van primaire bronnen (reisverslagen, journalen, rechtsdocumenten enz.) over de periode van de tweede helft van de 17de eeuw tot en met de 19de eeuw noodzakelijk. Zo'n onderzoek moet het ook mogelijk maken de historische fases in de verhouding tussen het Nederlands en het Afrikaans beter te onderscheiden. Voor het huidige woordenboek ont-

braken daartoe de middelen en bezat de redactie nog niet de vereiste woord-
historische en etymologische kennis, maar voor een mogelijke tweede druk zou
er alles aan gedaan moeten worden om het in enige substantiële vorm van de
grond te krijgen.

5.   EWA en het Nederlands

In het voorgaande hebben we al op verscheidene plaatsen op overeenkomsten
en verschillen van het Afrikaans met het Nederlands gewezen. Die focussing
op het Nederlands ligt natuurlijk voor de hand. Het aandeel van het Neder-
lands in de woordenschat van het Afrikaans is bijzonder groot. Niet alleen
vormt de 17de-eeuwse informele Hollandse omgangstaal een belangrijke basis
voor het Afrikaans, maar ook later, vooral vanaf de 19de eeuw, heeft het Af-
rikaans de invloed van de geschreven Nederlandse cultuurtaal ondergaan. De
geschiedenis van de Nederlandse woordenschat is daarom niet alleen van
belang voor de geschiedenis van de Afrikaanse woordenschat, maar de ge-
schiedenis van de Afrikaanse woordenschat is ook van groot belang voor de
geschiedenis van de Nederlandse woordenschat.

In dat licht is belangrijk en interessant dat EWA de aandacht vestigt op
heel wat oude of gewestelijke Nederlandse woorden en betekenissen, zoals bij-
voorbeeld: *agtelosig*, verouderd Ndl. *achtelozig*; *astrant* 'vrijpostig, brutaal' nog in
gewestelijk Ndl.; *frot* 'tikkertje', verband met gewestelijk Ndl., in versch. zuide-
lijke dialecten *frot* 'slag' (op zijn beurt in relatie met fr. *frotter* 'aanraken'); *kwan-
suis* 'schijnbaar', verouderd Ndl.; *nakend* uit Ndl. spreektaal (in 17de-eeuwse
Hollandse spreektaal nog dominante variant van *naakt*), nu nog gewestelijk;
nuut, uit gewestelijk Hollands *nuut*, naast *nieuw* in de schrijftaal; *oorhoops*, ver-
ouderd Ndl. *overhoops*, nevenvorm van *overhoop*; *polvy* 'hak van 'n skoen', ver-
ouderd Ndl. *pollevij*; *runnik* 'die geluid van 'n perd voortbring', gewestelijk Ndl.
runniken; *skraps* 'schaars, onvoldoende, zuinig, krap, schriel' , verouderd Ndl.
schrap, nog gewestelijk in Vlaanderen; *tet* 'vrouwenborst, tepel', nog in
gewestelijk Ndl.; *uitlander* 'buitenlander', verouderd Ndl.; *verbrands*, verouderd
Ndl. *verbrand* (16de/17de eeuw), uitroep van verbazing, teleurstelling of ont-
steltenis.

EWA bevat op diverse plaatsen grondwoorden van afleidingen die in het
Nederlands verloren gegaan zijn of alleen gewestelijk nog bekend zijn. *Parman-
tig* 'onbeskof, uitdagend, vermetel' heeft zich ontwikkeld uit Ndl. *parmantig*
(1599) 'deftig, fier'. Daarnaast vermeldt EWA een substantief *parmant* 'onbe-
skofte, uitdagende, vermetele persoon'. Dat laatste is te verbinden met een ver-
ouderd Ndl. bijvoeglijk naamwoord *parmant* (1599), dat volgens het WNT
gewestelijk o.a. in Friesland en Holland nog als substantief bekend is. Afri-
kaans *toet* gaat terug op Ndl. *toeten*, de verouderde grondvorm van het beken-
de frequentatief *toeteren*. *Troetel* 'liefkozen', 'baie liefde of sorg gee' is te verbin-
den met verouderd Ndl. *troetelen* (1532), dat verdrongen is door zijn thans zeer
gebruikelijke afleiding met *ver-*: *vertroetelen*.

Een voorbeeld van een oude Nederlandse betekenis die in het Afrikaans bewaard is gebleven, levert *fluks* 'groot van gestalte, frisgebouw'. Het WNT vermeldt een verouderde betekenis 'flink, stevig, fors (van voorkomen of lichaamsbouw)' en Van Riebeeck in hoogsteigen persoon levert als citaat: 'Zij waren zeer fraai en zeer fluksche mannen van bijzondere goede statuur'. Opmerkelijk is het woord *kont*. In de Afrikaanse betekenis 'vroulike skaamdeel' leeft de oorspronkelijke betekenis van dit oude, Germaanse woord voort. Het is ook de oudste betekenis van het Ndl. *kont*, nu alleen nog gewestelijk bewaard gebleven in o.a. Vlaanderen, Groningen en Friesland. In het Standaard-Nederlands betekent *kont* nu 'zitvlak, billen', door het WNT gekarakteriseerd als 'blijkbaar jonge' betekenis. Het bevat geen bewijsplaatsen van voor de 18de eeuw, wat de afwezigheid ervan in het Afrikaans verklaart. Het woord had die betekenis eenvoudigweg nog niet, toen Van Riebeeck en de zijnen aan de Kaap neerstreken.

Gegevens in EWA zullen ook beter inzicht verschaffen in vroegere gewestelijke verbreidingen. Het znw. *raas* 'bestraffing, vermaning' krijgt de volgende etymologie: 'Uit gewestelike Ndl. *raas* 'beknorring''. Die etymologie berust duidelijk op gegevens uit het WNT. Dat vermeldt als betekenis 5): 'Beknorring'. De toepasssing blijft volgens de bronnen (de Groningse dialectwoordenboeken van Molema en Ter Laan) beperkt tot Groningen. Dat het Afrikaanse woord inderdaad uit die Groningse toepassing komt, is niet aannemelijk. Het woord is of was in andere (steeds regionale) betekenissen bekend: 'onzin, gekheid' bij 17de-eeuwse Vlaamse en Zeeuwse schrijvers, 'dom gepraat' en 'geschreeuw, getier' in Vlaams-België; 'behoefte om dol te doen' in Limburg (Valkenburgs). Een betekenis als 'bestraffing, berisping' kan teruggaan op 'geschreeuw, getier' en is oorspronkelijk waarschijnlijk over een groter gebied verspreid geweest.

Lexicale divergentie van het Afrikaans met het Nederlands is niet alleen veroorzaakt doordat in het Afrikaans woorden bewaard bleven die in het Nederlands zelf verdwenen zijn of er alleen gewestelijk nog voortleven. Beide zijn natuurlijk ook geheel eigen nieuwe wegen gegaan. Voor divergentie zorgt daarbij vooral ook het verschil in reactie op de invloed van het Engels. Bij de voorbeelden van leenvertalingen die we gaven bij de behandeling van de macrostructuur, zagen we al dat het Afrikaans veel eigen vormingen heeft waar het Nederlands simpelweg de Engelse woorden onvertaald ontleend heeft. EWA levert ook tal van voorbeelden waarbij het verschil komt, doordat het Nederlands een leenvertaling met andere elementen heeft of het woord niet kent, zoals onder meer het geval is bij: *aanddrag* naar Eng. *evening dress* (1896), Ndl. *avondkleding*; *baaibroek* naar Eng. *bathing-trunks* (1895), Ndl. *zwembroek*; *bloedoortapping* naar Eng. *blood transfusion* (1879), Ndl. *bloedtransfusie*; *bloufilm* naar Eng. *blue movie* (1965), Ndl. *pornofilm*; *eetstaking* naar Eng. *hunger strike* (1889), Ndl. *hongerstaking*; *dagbreek* naar Eng. *day-break* (1530); *gidshond* naar Eng. *guide dog* (1932), Ndl. *blindegeleidehond*; *grondboontjiebotter* naar Eng. *peanut butter* (1903), Ndl. *pindakaas*; *kweekhuiseffek* naar Eng. *greenhouse effect* (1937), Ndl. *broeikaseffect*; *proefbuisbaba* naar Eng. *test-tube baby* (1978), Ndl. *reageerbuisbaby*;

selfhelpwinkel naar Eng. *self-service shop*, Ndl. *zelfbedieningswinkel, zelfbedienings-zaak*; *skouhuis* naar Eng. *show house* (1962), Ndl. *modelwoning*; *surrogaatmoeder* naar Eng. *surrogate mother* (1972), Ndl. *draagmoeder*; *trompoppie* naar Eng. *drum-majorette* (1938), Ndl. *majorette*.

Ook buiten de leenvertalingen heeft het Afrikaans inheemse vormingen en het Nederlands Engelse leenwoorden, vgl. *pendeltuig* tegenover *shuttle*, *rekenaar* versus *computer*. Natuurlijk doet zich ook het omgekeerde voor: het Afrikaans heeft Engelse woorden waar het Nederlands inheemse vormingen heeft, vgl. *sokker* tegenover *voetbal*. De steeds meer toenemende invloed van het Engels op het Afrikaans, sterk vooral, zoals we boven al zagen, in de informele omgangs-taal, zal voor een verdere lexicale verwijdering zorgen.

EWA bevat voorts de nodige inheemse woorden of uit de 'swart tale' over-genomen woorden die binnen de eigen cultuur gelden voor tal van zaken, pro-cessen, activiteiten, stereotiepen en opvattingen, levensomstandigheden, flora en fauna, enz., waarvoor het Nederlands geen equivalenten heeft. En het ligt voor de hand dat men voor nieuwe zaken, nieuwe internationale begrippen van na de Nederlandse invloed eigen woorden vormt, zoals *paddavoet* < *padda* 'pad' en *voet*, zo genoemd 'omdat die voetbedekking aan die gewebde poot van 'n padda herinner'. Het Nederlands heeft hier het veel minder beeldende *zwemvliezen*.

Niet alleen voor het lexicon maar ook voor de historische klank- en vorm-leer van zowel het Afrikaans als het Nederlands kan EWA een nuttige leveran-cier van gegevens zijn. Er staan in EWA heel wat woorden die in hun klank-vorm direct aansluiten bij de klankvorm waarin die woorden gewestelijk in het Nederlands taalgebied nog bekend zijn of vroeger bekend geweest moeten zijn: woorden met *eu* in plaats van *oo* zoals *meule*, woorden met *sk* in plaats van *sch*, woorden met *e/i-* wisseling, met velarisering, enz. Ik wijs ter adstructie op een interessant woord als *skaduwee*, ontstaan uit de oude verbogen vorm *schaduwe* met verzwaring van de onbeklemtoonde eindvocaal tot *-ee*, en voor de morfo-logie op *oubaas, foeitog, kleinhuisie, swartskaap, slimjan*, die blijk geven van de neiging van het Afrikaans om één woord te vormen waar het Nederlands een verbinding of woordgroep heeft.

6.   Voor verbetering vatbaar

Vóór hun opleiding hadden de EWA-redacteuren geen enkele bemoeienis met etymologie of historische taalkunde gehad. Modern-taalkundigen inwijden in de etymologie is zoiets als een woestijnvolk leren schaatsen.Tegen die achter-grond is het verbluffend wat zij binnen relatief korte tijd hebben gepresteerd. Niettemin blijven er nog enige wensen tot verbetering over. Op een aantal daarvan wil ik in deze paragraaf ingaan. Ze betreffen de problematiek van de vormverschillen, de 'betekeniskloven', ontlening uit het Nederlands of uit het Engels en normatieve redactionele opmerkingen.

6.1   Vormverschillen en betekeniskloven

Afspraak in het algemeen was dat aan vormverschillen die veroorzaakt zijn door systematische klankontwikkelingen, stilzwijgend voorbijgegaan kon worden. Wel was bespreking noodzakelijk van aparte, afwijkende, onregelmatige, opvallende vormverschillen. Ondanks die afspraak bevat het woordenboek nogal wat gevallen waarbij op dergelijke vormverschijnselen onvoldoende of in het geheel niet gelet is. Alertheid voor (zelfs minimale!) vormverschillen is niettemin altijd geboden. Ze leiden vaker tot andere etymologische hypotheses.

EWA geeft *blikners* 'deurgeskawe op die sitvlak en tussen die bene, gewoonlik wanneer 'n perd bloots gery word' en beschouwt dit als een inheemse samenstelling van *blik* en *ners* = *naars,* nevenvorm van *aars* (na metanalyse van het onbepaald lidwoord). Ter vergelijking wordt gewezen op Ndl. *blikaars.* Dat brengt ons bij een verouderd Ndl. adjectief *blik* 'van vel ontbloot', horend bij een algemene betekenis 'open, bloot'. De specifiekere betekenis 'ontveld' komt al voor bij Mnl. *blec.* Voor *blikaars* geeft het WNT als betekenis: 'eene ontvelling aan de schaamdelen, inzonderheid veroorzaakt door paardrijden', precies die gespecialiseerde betekenis die EWA voor het Afrikaanse woord geeft. Het WNT bevat bovendien enige citaten, waaronder een met de vorm *Blick-naers*: 'Dat de Huyt ghestadigh traent, ghelijck in eene Blick-naers', uit *Heel-konste*, 3de druk 1660, van J. van Beverwijck. Daaruit blijkt dat het 17de-eeuwse Nederlands ook al de vorm met *-n-* had. Afrikaans *blikners* kan nu heel goed direct op die oude gewestelijke Hollandse nevenvorm teruggaan.

Een zelfstandig naamwoord *verkoue* wordt teruggevoerd op Ndl. *verkoudenheid.* Gewezen wordt voorts op *verkouenis* (1733) en *verkouentheid* (1803). Formeel is het substantief *verkouden* echter niet op een lijn te stellen met die afleidingen. Een znw. *verkouden* is in het Nederlands niet bekend. Daarom moet het zijn ontstaan aan een ontwikkeling in Afrikaans zelf danken. Waarschijnlijk gaat het om een geval van conversie van het bijvoeglijk naamwoord *verkoue* (dit wel < Ndl. bnw. *verkouden*).

Doedoe, een kindertaalwoord voor 'slapen', zou zijn oorsprong hebben in Ndl. *dodijnen* (1617): ''n kindjie in die arms wieg en aan die slaap sing'. Hier wordt het hele bestanddeel *-ijnen* veronachtzaamd. *Narsing*: 'narcis' wordt herleid uit Ndl. *narcissen*, meervoud van *narcis.* Daarvoor neemt men een tussenvorm *narsen(s)* aan, die velarisering heeft ondergaan. Maar hoe moeten we het ontstaan van die tussenvorm begrijpen? Waarom zou die *is*-klank zomaar wegvallen? Zijn er parallellen in Afrikaans?

Onduidelijk en onjuist is in het artikel *glips* 'fout, vergissing, flater e.d.' de link met Engels *slip* 'flater, vergissing', waarmee volgens de redacteur *glips* naar vorm en betekenis overeenstemt. Wat de betekenis betreft klopt dit, maar wat de vorm betreft natuurlijk niet: een *g* is geen *s.* Voor *straaljagter* wordt Ndl. *straaljager* als mogelijke oorsprong geopperd. Als dit zo zou zijn, zou in ieder geval het vormelijke verschil tussen *-gt-* van het Afrikaanse en *-g-* van het Nederlandse woord, ofwel tussen de twee verschillende woorden *jager* en *jagter*, besproken moeten worden.

Dekodeer wordt gekarakteriseerd als een verafrikaanste vorm van Eng. *decode*. Vanwaar dan die uitgang *-eer*? Ligt hier overigens herkomst uit Ndl. *decoderen* niet meer voor de hand? *Skandeer* 'in versvoet verdeel' zou ook een afleiding met *-eer* zijn van de verafrikaanste vorm van Eng. *scan*. Hoe dan de *d* te verklaren? Overigens geldt ook hier dat herkomst uit Ndl. *scanderen* wel een betere oplossing is.

Bij *glimp* is een nevenvorm *glimps* vermeld, maar op die *-s* wordt in de etymologie zelf verder ten onrechte niet ingegaan. Ik kan niet bepalen in hoeverre het hier om een voor het Afrikaans bekend verschijnsel gaat. Wat ik wel weet is, dat EWA diverse voorbeelden bevat van woorden met *-ps* in plaats van *-s*: *blaps* 'fout, flater, gemors' < Eng. *blab* 'onbeduidend gepraat'; *deps* < Ndl. *deppen*; *gaps* 'iets vinnig gryp of steel' naast *gap* < Ndl. *gappen*; *glips* 'fout, flater', mogelijk verband met Ndl. *glippen*; *gops* 'eenvoudige, onopgevoede agteraf persoon' < Eng. *gob* 'Amerikaans matroos of gewone zeeman'; *keps* 'pet' < *kep* 'pet'; *klops* als afleiding van *klop* 'slaan'; *raps* 'ligte hou of liggies slaan' < Eng. *rap*; *sieps* 'weergave van een slaangeluid'; *skraps* < Ndl. *schrap*; *tops* 'baie goed' naast *top*.

Niet alleen opvallende vormverschillen of vormafwijkingen behoeven verklaring. Het is ook een vereiste voor een goede etymologie dat betekeniskloven worden gedicht, of in ieder geval worden besproken, als overbrugging niet mogelijk blijkt. Blote vermelding van betekenissen waartussen het verband volstrekt ondoorzichtig is, roept vragen op bij de lezers waarmee je ze niet mag laten zitten. Toch gebeurt dat in EWA wel eens. Bij *siebie* 'klein hondjie' wordt als een van de mogelijke etymologieën gegeven 'uit 'n S.A. swart taal *Sibi*, die naam van 'n stasie naby Pienaarsrivier in die Limpopo-provinsie'. Wat de naam van een 'stasie' te maken heeft met een klein hondje, is mij een raadsel. Hier was nadere toelichting geboden geweest.

Dalk wordt beschouwd als een samentrekking van *dadelik*, dat teruggaat op Ndl. *dadelijk*. Maar *dalk* heeft een betekenis 'miskien, moontlik', *dadelik* betekent 'onmiddellik, sonder versuim, op die oomblik'. Het opvallende betekenisverschil tussen beide woorden wordt ten onrechte niet verklaard. Overigens is ook in de woordgeschiedenis van Nederlands *dadelijk* geen aanknopingspunt te vinden voor een betekenis 'misschien, mogelijk'.

Deps 'iets wegvat of vir die grap steel' (omgangstaal) zou een afleiding zijn met *-s* van Ndl. *deppen*, klanknabootsend uit *dabben* 'stamp met voorpote' (t.o.v. perde)'. Wat het verband zou moeten zijn tussen 'wegnemen, stelen' en 'stampen met de voorpoten, van paarden', blijft onverklaard.

6.2   Uit het Engels of uit het Nederlands?

Bij nogal wat woorden is gekozen voor een herkomst uit het Engels, waar herkomst uit het Nederlands evengoed en vaak zelfs meer in aanmerking zou zijn gekomen. Ik kan mij niet herinneren dat hierover afspraken zijn gemaakt, of dat de redactie daar zelf een criterium voor heeft ontwikkeld. Maar het komt zo regelmatig voor dat we in ieder geval van een tendens kunnen spreken. Zo

zouden bijvoorbeeld *casino*, *frokkie*, *lasagna*, *trevira* uit het Engels en niet uit het Nederlands stammen, dit in weerwil van het bestaan van Ndl. *casino*, Ndl. *frok*, Ndl. *lasagna* en Ndl. *trevira*.

Voorkeur voor het Engels is ook herhaaldelijk te constateren bij samenstellingen die als leenvertaling uit het Engels worden gekenmerkt, terwijl diezelfde samenstellingen ook in het Nederlands aanwezig zijn. *Jeugherberg* zou een leenvertaling zijn van Eng. *youth hostel,* over de aanwezigheid van een mogelijke herkomst uit Ndl*. jeugdherberg* wordt niet gerept. *Skokdemper* wordt opgevat als leenvertaling van Engels *shock-absorber* en niet als ontlening aan Ndl. *schokdemper*. *Dagdroom* wordt beschouwd als leenvertaling van Eng. *daydream*, terwijl Ndl. *dagdroom* als bronwoord rechtstreekser in aanmerking komt. Ook *swartskaap* wordt beschouwd als leenvertaling van Engels *black sheep* of eventueel zelfs van Frans *la bête noire*, wat gelukkig wel minder waarschijnlijk wordt geacht. Niet duidelijk wordt waarom Ndl. *zwart(e) schaap* hier niet mee mag doen en dus het zwarte schaap is. Ook bij *heupfles* wordt eenzijdig voor de moeilijke oplossing van een leenvertaling van Eng. *hip flask*, gekozen, terwijl Ndl. *heupfles* toch voor het grijpen lag.

Wat voor woorden geldt, geldt ook voor betekenissen. *Speen* wordt gegeven met twee betekenissen: 'tepel van een dier' en 'rubbermondstuk op 'n bababottel'. Voor de eerste betekenis wordt Nederlandse herkomst aangenomen, maar voor de tweede betekenis wordt een leenbetekenis van Engels *nipple* waarschijnlijker geacht. Waarom, is mij niet duidelijk. Het Nederlandse *speen* is ook in die betekenis zeer gebruikelijk.

Bij heel wat gevallen heb ik op grond van historische achtergrond en overeenkomende vorm juist omgekeerd de neiging eerder aan Nederlandse dan aan Engelse herkomst te denken. Los van die subjectieve ervaring, moet objectief gelden: als er geen criterium is noch een concrete aanwijzing dat aan herkomst uit een van beide talen de voorkeur gegeven moet worden, moet men beide mogelijkheden open laten, zoals bijvoorbeeld keurig het geval is bij *eskapade*: 'Uit Ndl. *escapade* (1824) of Eng. *escapade* (1653)'.

6.3   Normatieve redactionele opmerkingen

Etymologische woordenboeken zijn krachtens hun aard descriptief. Men verwacht er geen kritiek op woorden, hun vorming of hun gebruik. Af en toe steken ze in EWA toch de kop op, zoals bijvoorbeeld bij *loopgraaf*: 'Ndl. *loopgraaf* is 'n lomp vertaling van D. *laufgraben* en behoort eintlik *loopgracht* te wees', of bij *hartomleiding*: 'Dit is wel 'n hartoperasie wat uitgevoer word, maar dit is bloedvate en nie die hart nie wat omgelei word. Dit is dus in werklikheid foutief om van 'n 'hartomleiding' te praat'.

6.4   Kennis van het Nederlands

Soms vormt een niet geheel toereikende kennis van het Nederlands een pro-

bleem voor de redactie. Voor de bepaling van het (nog) al dan niet voorkomen van woorden of betekenissen is men dan geheel aangewezen op het WNT of de Grote Van Dale. Het gaat niet altijd goed. Bij het lidwoord '*n* wordt als etymologie gegeven: 'Verkorting as gevolg van swak beklemtoning van *een*, met lg. uit Ndl. *een*.' Daarna volgt de mededeling: 'Die lw. '*n* is ook gewestelik in Gronings opgeteken.' Ik weet niet waar die mededeling op gebaseerd is. Zij is in ieder geval onjuist. De uitspraak met zwak beklemtoonde vocaal is in het Nederlands algemeen, althans niet aan enig gewest gebonden. De spelling '*n* is een informele schriftelijke weergave van die uitspraak.

Ook onbekendheid met de Nederlandse spellingsgewoonten kan soms leiden tot merkwaardige uitspraken. Bij *aamborstig* wordt Ndl. *aamborstig* als juiste oorsprong aangewezen, met een datering 1710. Voor *aamborstich*, louter een spellingsvariant met *ch* voor *g*, wordt apart 1612 opgegeven. De kwalificatie 'vroeër ook' is daarbij nogal vreemd.

7.   Besluit

EWA is een bijzonder product van Nederlands-Afrikaanse samenwerking. Van Nederlandse zijde was er allereerst de lancering van het idee door prof. P.G.J. van Sterkenburg, directeur van het Instituut voor Nederlandse Lexicologie (INL). De benodigde fondsen kwamen er dankzij dr. J. Kist, penningmeester van het INL, en mr. drs. L. Jongsma, voorzitter van de Raad van Bestuur van de uitgeversmaatschappij Sdu. Ik tekende, zoals ik in de openingsalinea al aangaf, voor de etymologisch-lexicografische opleiding en begeleiding. Van Afrikaanse zijde was er de omarming van het idee door prof. A.H. van Wyk, voorzitter van de toenmalige Beheerraad van de WAT. Het WAT leverde in de persoon van dr. D.J. van Schalkwyk, toen hoofdredacteur van het WAT, de bestuurder van het project en zorgde voor administratieve ondersteuning door M.H. van der Westhuizen en T. Harteveld, die zich ook met het zetwerk belastte.

Maar EWA is en blijft bovenal het werk van 'tegniese redakteur' G.J. van Wyk en de auteurs A.E. Cloete, A. Jordaan, dr. H.C. Liebenberg en prof. H.J. Lubbe. In het werk klinkt ook nog op verscheidene plaatsen de naam door van prof. J. Combrink, met dr. Liebenberg redacteur van het eerste uur, maar ons helaas te vroeg ontvallen. Zij hebben een boeiend en onderhoudend werk afgeleverd en verdienen daarvoor respect, hulde en: Baie dankie.

Bronnen

Bloch, Oscar en Walther von Wartburg. 1994. *Dictionnaire étymologique de la langue française*[10]. Paris: Presses Universitaires de France.

Boshoff, S.P.E. en G.S. Nienaber. 1967. *Afrikaanse etimologieë*. Pretoria: Suid-Afrikaanse Akademie vir Wetenskap en Kuns.

EWA. — Van Wyk, G.J. (Tegn. Red.), A.E. Cloete, A. Jordaan, H.C. Liebenberg en H.J. Lubbe. 2003. *Etimologiewoordeboek van Afrikaans*. Stellenbosch: Buro van die WAT.

Hiersche, Rolf. 1993. Zum Deutschen Etymologischen Wörterbuch. Bremmer, Rolf H. en Jan van den Berg (Red.). 1993. *Current Trends in West Germanic Etymological Lexicography*: 65-76. Leiden: E.J. Brill.

Malkiel, Yakov. 1976. *Etymological Dictionaries. A Tentative Typology*. Chicago/London: The University of Chicago Press.

Pfeiffer, Wolfgang. 1993. *Etymologisches Wörterbuch des Deutschen*. Berlin: Akademie Verlag.

Pfister, Max. 1980. *Einführung in die Romanische Etymologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Ponelis, Fritz. 1998. Convergentie van het Afrikaans met het Engels. Klooster, W.G., H.J.W.M. Broekhuis, E.H.C. Elffers-van Ketel en J.P.A. Stroop (Red.). 1998. *Eerste Amsterdams Colloquium Nederlandse Taalkunde*. *Lezingen gehouden op 10 en 11 december 1997*: 121-131. Amsterdam: Leerstoelgroep Nederlandse Taalkunde Universiteit van Amsterdam.

Ponelis, Fritz. 2002. Etimologiewoordeboek van Afrikaans (recensie). *Lexikos* 12: 329-334.

Seebold, Elmar. 1981. *Etymologie. Eine Einführung am Beispiel der deutschen Sprache*. München: Verlag C.H. Beck.

WNT. — 1864–1998. *Woordenboek der Nederlandsche taal*. Leiden.

Revising Matumo's *Setswana–English–Setswana Dictionary*

D.J. Prinsloo, *Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (danie.prinsloo@up.ac.za)*

Abstract: The aim of this article is to design a revision strategy for the Setswana to English side of the *Setswana–English–Setswana Dictionary* compiled by Z.I. Matumo in 1993. An existing general organic Setswana corpus as well as a dedicated corpus compiled for the purposes of the revision will be used as a basis for macro- and microstructural aspects of the proposed revision. Lemma candidate lists for inclusion in and omission from the existing dictionary will be generated from these corpora, existing articles will be critically analysed and models for revised/updated articles will be presented. Key components of the revision strategy include the design and use of a multi-dimensional Ruler and Block System for the measurement and balancing of alphabetical stretches for the revised dictionary in terms of time, average length of articles and number of pages per alphabetical category. It is not possible to present all aspects of the revision within the scope of a journal article but the most prominent ones as well as a selection of typical issues will be dealt with.

Keywords: LEXICOGRAPHY, LEMMATISATION, REVISION, INFORMATION RETRIEVAL, MACROSTRUCTURE, MICROSTRUCTURE, RULER, BLOCK SYSTEM, DICTIONARY, AFRICAN LANGUAGES, SETSWANA (TSWANA)

Opsomming: Hersiening van Matumo se *Setswana–English–Setswana Dictionary*. Die doel van hierdie artikel is om 'n hersieningstrategie te ontwerp vir die Setswana na Engelse kant van die *Setswana–English–Setswana Dictionary* wat in 1993 deur Z.I. Matumo saamgestel is. 'n Bestaande algemene Setswanakorpus asook 'n spesifieke korpus wat saamgestel is vir die doel van die hersiening sal as basis vir mikro- en makrostrukturele aspekte van die voorgestelde hersiening gebruik word. Lemmakandidaatlyste vir insluiting in en weglating uit die bestaande woordeboek sal vanuit hierdie korpusse gegenereer word, bestaande artikels sal krities ontleed word en modelle vir die hersiene bygewerkte artikels sal aangebied word. Sleutelkomponente van die hersieningstrategie sluit die ontwerp en gebruik van 'n multi-dimensionele Liniaal en Bloksisteem in vir die meting en balansering van alfabetiese reekse vir die hersiene woordeboek in terme van tyd, gemiddelde lengte van artikels en aantal bladsye per alfabetiese kategorie. Dit is nie moontlik om alle aspekte van die hersiening binne die bestek van 'n tydskrifartikel aan te bied nie maar die vernaamstes, asook 'n aantal tipiese kwessies, sal behandel word.

Sleutelwoorde: LEKSIKOGRAFIE, LEMMATISERING, HERSIENING, INLIGTINGSONTSLUITING, MAKROSTRUKTUUR, MIKROSTRUKTUUR, LINIAAL, BLOKSISTEEM, WOORDEBOEK, AFRIKATALE, SETSWANA (TSWANA)

Introduction

Substantial revision and updating of a dictionary require detailed and meticulous planning on microstructural and macrostructural levels and is not less laborious than the planning and design of a new dictionary. Lexicographers often err in tackling such revisions in a haphazard way; eager to simply add new words to the dictionary rather than to take an holistic approach towards delivering a well-balanced and improved product.

> Many people think that the bulk of the work done by lexicographers, or dictionary makers, is that of collecting new words and defining them. Inclusion of the latest words is indeed a major part of our work, but no less important is the revising and updating of the entries for words that are already in our dictionaries. . ... During revision every aspect of a dictionary entry is examined and if necessary changed. (Stevenson 2004)

> The most obvious way the dictionary will develop is by the addition of more words. We already have a small list of words for inclusion in the next edition, and we look forward to obtaining more from our readers as well as from our own researchers. (Matumo 1993: ix)

Landau (2001) distinguishes between *updating* and *revision* of a dictionary. He regards updating as an exercise which should ideally be performed annually or biennially while substantial revision or in his terms *a complete re-examination of the previous edition* should be performed about every ten years.

> Dictionaries may be updated by the substitution of some new entries for old entries, and for the first few years after publication, such a procedure may work very well. But when a dictionary passes the ten- or fifteen-year-old mark, updating takes on a desperate character. (Landau 2001: 397)

The envisaged revision of Matumo's *Setswana–English–Setswana Dictionary* (henceforth referred to as MSD), published by Macmillan in 1993, thus qualifies in terms of Landau for such substantial revision.

On macrostructural level, the most prominent issue in the revision of a dictionary remains the decisions on lemmas to be included or excluded as echoed by Busane (1990: 30):

> One of the basic problems of lexicography is to decide what to put in the dictionary and what to exclude.

On microstructural level, the proposed revision of MSD will focus on a critical analysis of the data types and microstructural architecture with a view to creating a more user-friendly design with enhanced quality based on corpus data.

Background and original dictionary

MSD (1993) is the fourth edition of what is titled since 1993 the *Setswana–Eng-lish–Setswana Dictionary*. The first edition dates back to approximately 1875, the second to 1895, and the third to 1925, entitled *Secwana–English Dictionary*. The latter was compiled by J. Tom Brown and formed the basis for MSD.

The features of the 'new' (1993) edition are summarized as follows:

— Completely reset in the most up-to-date orthography.
— Greatly increased number of headwords.
— Grammatical details in contemporary dictionary style.
— Tables of noun classes, concords and prefixes.
— References to many Setswana traditions.
— Proverbs quoted to illustrate delicate shades of meaning.
— Descriptive, not prescriptive, particularly with regard to borrowed or coined words. (Matumo 1993: Back cover)

In the Introduction Matumo says:

> I am as conscious as anyone else that there are shortcomings in this dictionary. Language is a fluid and developing organism, and a dictionary freezes it momentarily so that its vocabulary can be studied. This means that in an important sense a dictionary is already out of date on its day of publication. (Matumo 1993: ix)

Electronic Setswana corpora

The proposed revision of MSD is based on two Setswana electronic corpora. Firstly, the general Setswana Pretoria Corpus, compiled at the University of Pretoria, consisting of a variety of printed matter totalling 4.5 million running words (tokens) and 131 000 different words (types). Secondly, a dedicated Setswana corpus consisting of publications most likely to be studied by the target users of the revised dictionary, of approximately 1 million running words and 50 000 types.

Macrostructural revision strategies

As far as the choice of lemmata is concerned, the challenge to the lexicographer is the question as to whether, on the one hand, lemmas most likely to be looked for by the target users are included, and, on the other hand, whether all lemmas currently included in MSD can be justified in terms of such a likelihood. If frequency of use is an important criterion as is the case in the revision of MSD, the question is whether frequently used words were not accidentally left out or whether all the lemmas included in MSD deserve a place in the dictionary. Further the question could be raised if the space they occupy should rather be

more fruitfully used for other words that either have a high frequency in the general corpus or a high frequency in the dedicated one. (See De Schryver and Prinsloo (2003) for a detailed discussion of the issue of balancing out general corpora and dedicated corpora in an effort to compile a lemma list for a restricted dictionary.) Even if the lexicographer ignores frequency counts and decides on the basis of his/her intuition that current entries should be retained, the question is whether they should be lemmas in their own right or treated in the articles of other lemmas.

Consider the following examples of words that occur more than a thousand times in the general corpus, frequently in the dedicated corpus and which were entered as translation equivalents in the English–Setswana side of MSD but that were not lemmatised in the Setswana–English side.

Table 1:   Frequently used words not included as lemmas in MSD

Lemma	Freq. Gen. Corpus	Freq. Ded. Corpus	Meaning
le	12 5616	8 851	and
fa	56 463	2 987	here
bona	16 697	1 431	they; see
batho	8 424	899	people
botlhe	1 662	167	all
bosigo	1 478	95	night
sekolo	1 218	123	school
bonala	1 105	22	visible
tseo	1 064	58	those
otlhe	1 055	93	all

The occurrence of such instances underline the view of De Schryver and Prinsloo (2000) that utilization of a corpus is indispensable in assuring that words most likely to be looked for by target users are not omitted simply because they did not cross the compilers' way.

Different types of omissions/inconsistencies are apparent in Table 1. Firstly, a common failure is to complete a typical paradigm of which only a limited number of elements exist, e.g. quantitatives (cf. Gouws and Prinsloo (1997: 47) for a perspective on limited versus unlimited elements). The forms for classes 8 or 10 *tsotlhe* (2 336), class 15 *gotlhe* (397), 1st pers. plural *rotlhe* (217) and class 14 *jotlhe* (183) are given, but not classes 2 *botlhe* (1 662), class 6 *otlhe* (1 055), class 5 *lotlhe* (409), class 7 *sotlhe* (67), etc.

A second example in this regard is the demonstrative second position, class 7: *seo* 'that one' (1 301) is given, but not classes 8 or 10 *tseo* 'those' (1 064), class 5, *leo* 'that one' (949), etc. All demonstratives given in the guidelines to the dictionary should be treated in the central text. (See Prinsloo (1996) for a discussion on dead references pertaining to words given in the guidelines to a dictionary.)

In order to combat what Gouws and Prinsloo (1998: 21) call the decontextualisation of lexical items, brought about by the alphabetical sorting of lemmas in a dictionary, tables such as those given for the quantitatives and demonstratives in the front matter fulfil a valuable function in restoring such lexical and grammatical relations. It is however imperative that the members of such a paradigm be lemmatised in the central text and that appropriate and correct reference be made from each individual lemma to the tables as reference addresses. Compare also in this regard the inclusion of numerous colour plates of different trees and cattle in the back matter of Kgasa and Tsonope (1995) without cross-referencing from the articles of these trees and cattle in the central text.

When candidates for deletion from the lemma list of MSD must be decided on, consider the following extract from a list of multiword lemmas in MSD.

Table 2:   A selection of multiword lemmas in MSD

ka go dira	ke gone	ka thelelo	ka mmanene	tolwana ya leitlho
mokgatha-thete	ke a	kgaphasetse	ka kgaga	thini ya tsebe
ke ne ke	kaololwa	kgeleisitse	ka ke	tladi mothwana
ke mong	kago e e godileng	ka moso	ka jeno	tladi ya tlapana
ke mang	kabayanya	ka mmanete	ka gope	nkgiwa

Singled out for attention here are the numerous clusters presented as multiword lemmas. The lemmatisation of multiword items such as *ke ne ke* 'I was', *ka go dira* 'by acting', *kago e e godileng* 'a building that is high or tall', etc. cannot be critisized in principle. Gouws (1991) and Zgusta (1971) emphasize that there are numerous multiwords that should be regarded as single lexical items and therefore be presented as multiword lemmas in the central text of the dictionary. However, in MSD multiword lexical units are often confused with frequently used free combinations.

The potential for the successful retrieval of information by target users is also low for most lemmas in Table 2. Of the 330 occurrences of *kago* 'the process of building, a building' in the corpus, *kago e e godileng* occurred only once and clusters such as *kago ya phemelo* 'protection building/structure' and *kago ya bokgoni* 'successful structure' occur more frequently with counts of 17 and 10 respectively but were not lemmatised as multiword lemmas. Since *kago* was lemmatised, no real harm is done in lemmatising *kago e e godileng* as well because alphabetically it directly follows the article of the lemma *kago* and may therefore catch the eye of the user. In the case of *ka go dira*, and many other similar ones, the value of the entry is however questionable since it is unlikely that the user will know how to look it up in the alphabetical stretch for K especially since no cross-referencing is provided from the article for *dira* to *ka go dira*.

Even if users do consult lemmas starting with or consisting of *ka*, they are confronted with another problematic aspect of lemmatisation in MSD, i.e. ex-

tensive stacking of a large number of lemmas, in this case 38, consisting of 12 lemmas for *ka* and 26 lemmas for *ka* plus a noun, verb, etc. Even a cross-reference to these 38 possible lemmas that are not marked as homonyms, e.g. by superscript homonym markers, would be user-unfriendly. A much better solution would be to treat frequent clusters such as *ka go dira* (167) *ke go dira* (87) and *kgona go dira* (51) in the article of *dira*.

Building and applying a multi-dimensional Ruler

Apart from the macrostructural aspect relating to inclusion versus omission of individual lemmata, such control should be exercised in terms of balancing out entire alphabetical categories in the dictionary as a whole.

> Nothing is more difficult to predict or control than a dictionary begun from scratch. (Landau 2001: 398)

This remark is equally applicable to dictionaries that were compiled without the availability of a corpus. (See De Schryver and Prinsloo (2000) and Prinsloo and De Schryver (2003) for numerous examples of inconsistencies regarding over- and undertreatment in terms of alphabetical categories.)

Consider the following example where substantial inconsistency between the length of articles in the first few alphabetical categories compared to the last few in Kriel (1983) is apparent even to the naked eye, without any help from measuring instruments.

(1)

aka	**2**	**ala**

deel vir, vonnis vir (of) tot; uitspraak gee vir, – *lehu*, ter dood veroordeel. *ahlolêlwa*, gevonnis word, veroordeel word, geoordeel word. *ahloleng*, julle moet oordeel; *se* -e, moenie oordeel nie. *ahlohlê*, mag/kan oordeel. *ahlotšwe*, gevonnis wees, uitspraak is gegee. *moahlodi*, regter, beoordelaar. *baahlodi*, regters, beoordelaars.

aka, *a.ka. (-ile, -etše)*, lieg, leuens vertel, jok,. onwaarheid spreek (dial. kyk: *aketša*).

aka, *a.ka*, inhaak, vashaak, haak, aanhaak, soen, omarm, lieg, liefkoos; *akwa*, gehaak/ingehaak word; *akêla*, haak vir; *akelana*, mekaar liefkoos, vriendskaplik verkeer; *akelwa*, ingehaak word vir; *akiwa*, ingehaak.

akere, *'a kê.'rê*, akker.

aketša, *a ke.tša*, leuen vertel, lieg, jok; *akeditše*, het (gelieg) 'n leuen vertel. *sa aketše*, nie lieg nie.

akga, *a.kga*, werp, gooi, slinger, swaai, beweeg. *akgaakga*, heen en weer beweeg (soos branders), slinger, skommel; *akgaakgwa*, heen en weer geslinger word; – *diatla*, arms swaai, met leë hande loop. – *dinao*, voet in die wind slaan; *akgwa*, beweeg/geslinger word; -*akgêga*, skommel, swaai; -*akgêla*, slinger, swaai, werp. *akgêla*, slinger na/vir, tou om die horings gooi, met 'n vangtou vang, uitkrap, soos kole uit 'n vuur. *akgelwa*, geslinger word, gevang word met 'n tou. – *dikobo*, klere uitpluk.

(2)

ribega, *ri bê.ga*, onderstebo keer, toemaak; bedek.
ribegetša, *'ri be ge.tša*, onderstebo draai.
ribesela, *ri bê sê.la*, omslaan, omkeer.
ribete, *'re bê.tê*, klinknael.
ribetela, *ri bê tê.la*, klink, vasnael, vasklink.
riboga, *'ri bo ga*, ontvang, swanger word.
ribogolla, *'ri bo go l.la*, regop draai, *ribolla*, openbaar, oopmaak, blootlê.
ribolla, *ri bo l.la*, omgekeerde oprig/regstel.
rifa, *ri.fa*, bedek, toemaak.
rifi, *'ri.fi*, rif, rotsbank; – *ya gauta*, goudrif.

robetše, *ro be.tše*, slaap, het ontslaap, *yo a robetšego*, ontslapene.
robja, *rô.bja*, gebreek word.
roboka, *rô'bô.ka*, smok, aanrand.
robong, *ro'bo.ng*, nege.
roborobo, *rô bô rô bô*, spoorwegbus
roboto, *ro bô tô*, robot, verkeerslig.
roga, *'ro.ga*, vloek, skel, skeltaal gebruik; swets, beledig.
rogaka, *ro ga.ka*, vloek, skel; *-wa*, gevloek, vervloek word; *-ile*, het vervloek; – iša, laat vloek.
rogana, *ro ga.na*, vloek, mekaar uitskel.
roganela, *ro ga nê.la*, vloek terwille van.
rogo, *'rô.gô*, rog, *mo-*, bredie.

In order to address such inconsistencies on the macrostructural level, Prinsloo and De Schryver (2002, 2003) and De Schryver (2003), studied the balance between alphabetical categories for English, Afrikaans and a number of African languages.

The question was whether a specific distribution, preferably one that could accurately be measured, exists between the different categories in a given language. They found that this is indeed possible. A remarkable consistency in respect of the balance between alphabetical stretches has been detected by comparing dictionaries and corpora. This consistency is observed with regard to, on the one hand, the number of lemmas treated for or the number of pages dedicated to each alphabetical category, and, on the other hand, the lemmatised as well as unlemmatised alphabetical word lists culled from corpora. For purposes of the revision of MSD, Rulers were compiled from the general corpus as well as from the dedicated corpus.

The concept *Ruler* is defined as a practical instrument of measurement for the relative length of alphabetical stretches in alphabetically ordered dictionaries. They are designed according to the generally accepted principle that alphabetical categories in any given language do not contain an equal number of words. For example, a single glance at a few popular English dictionaries reveals that the alphabetical categories or alphabetical stretches for A, B, D, M, R and especially C and S, contain large numbers of lemmas, occupying almost 50% of the dictionary, while categories such as J, K, Q, U, V, X, Y and Z are relatively small, and consequently fill only a few pages. For a dictionary such as the *Macmillan English Dictionary* (Rundell 2002), where the alphabetical categories are marked with coloured thumb tags, one does not even have to open the dictionary in order to appreciate this breakdown which can also literally be measured by putting an ordinary ruler against the dictionary to roughly measure the 'thickness' of each alphabetical stretch in millimetres. Likewise, an alphabetical list of types generated from the Sesotho sa Leboa corpus shows

that roughly 17% of all words in this language fall under the single category M while categories such as C, J, Q, U, V, W, X, Y and Z are virtually empty.

Consider the Ruler for Setswana in Figure 1, based on the average of the percentage breakdown of types in (a) the general Setswana corpus and (b) the dedicated Setswana corpus.
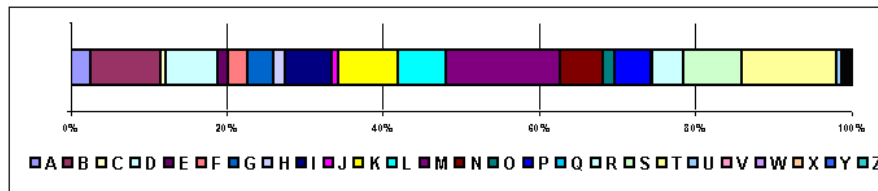
Figure 1:   A Ruler for Setswana

For the revision of MSD, the focus is shifted from an alphabetical breakdown in the sense of the balance between the 26 letters of the alphabet (A to Z) by reorganising the data given in Figure 1 into a *percentage* breakdown in the form referred to as a Block System in Table 3.

Table 3:   A Block System for Setswana

1	ALAF	21	FELE	41	KOUS	61	MOTL	81	SELE
2	AROG	22	FOLO	42	LAEL	62	MPHE	82	SERA
3	BADI	23	GAGW	43	LEBO	63	NATE	83	SETO
4	BANN	24	GATS	44	LEKI	64	NGWA	84	SIMO
5	BATW	25	GOLO	45	LERI	65	NKUK	85	SUAS
6	BIRO	26	GWET	46	LETS	66	NTEM	86	TALE
7	BOGA	27	HUBE	47	LOKO	67	NTSH	87	THAA
8	BOLA	28	IJES	48	MAAD	68	NYOR	88	THIB
9	BONK	29	IKGO	49	MAHA	69	OOMA	89	THWE
10	BORU	30	INOL	50	MALE	70	PANT	90	TLAM
11	BOUT	31	IPUS	51	MARA	71	PHAK	91	TLHA
12	DAAM	32	ITIS	52	MATL	72	PHIM	92	TLHO
13	DIFA	33	ITSH	53	MEFA	73	PITL	93	TLWA
14	DIKG	34	JOKO	54	MESU	74	PUDU	94	TSAP
15	DINK	35	KANY	55	MMAL	75	RAMO	95	TSHE
16	DIRA	36	KERO	56	MMOL	76	RENG	96	TSHW
17	DITH	37	KGAR	57	MOFI	77	ROKG	97	TSUN
18	DITU	38	KGOM	58	MOKG	78	RURU	98	UBAU
19	EGEP	39	KHAN	59	MONG	79	SEBA	99	WABO
20	ETLH	40	KODU	60	MORW	80	SEHI	100	ZIMB

While based on the same statistics, the Block System opens the door to a number of very practical applications and a multi-dimensional utilization in the revision process of MSD. For lexicographers and editors it gives clear guidance in terms of page allocation, average length of articles, progress in terms of time and even remuneration intervals for part-time compilers.

With the prescribed number of pages set at roughly 300 for each side of the dictionary, it means that 3 pages should roughly correlate with each block/ percentage point; the average article length should be 3 lines, and the average compilation time per article 10 minutes. Even remuneration scheduled at the markers 25% **GOLO**, 50% **MALE**, 75% **RAMO**, and 100% **ZIMB**, is being negotiated.

An actual compilation test was performed by treating a selection of 100 typical lemmas logging the average length and time used for the compilation of each article, with and without consultation of the corpora.

It is important that a sound perspective be maintained on the value of the multidimensional Ruler and Block System as dictionary compilation tools. They should not be regarded as absolute or precision instruments of measurement. The real value of the Ruler lies in the fact that it focuses the attention of the compiler on potential ill-balanced areas. This will now be illustrated for MSD.

Table 4:   MSD lemma and page breakdown versus the Setswana Ruler

	MSD: Lemmas %	MSD: Pages %	Setswana Ruler	MSD lemmas vs the Ruler
A	1.2	1.3	2.6	-1.4
B	**4.7**	**4.6**	**9.0**	**-4.3**
C	0.0	0.0	0.6	-0.6
D	6.0	6.4	6.6	-0.6
E	1.2	1.3	1.4	-0.2
F	3.7	3.3	2.4	1.3
G	5.2	5.3	3.4	1.8
H	0.9	0.9	1.5	-0.6
I	5.3	4.9	5.9	-0.6
J	0.7	0.7	0.8	-0.1
K	**12.2**	**11.9**	**7.7**	**4.5**
L	6.7	6.8	6.1	0.6
M	12.5	13.7	14.6	-2.1
N	4.0	4.0	5.5	-1.5
O	1.3	1.3	1.6	-0.3
P	5.9	6.0	4.6	1.3
Q	0.0	0.2	0.2	-0.2
R	3.9	3.5	3.9	0.0
S	8.5	8.6	7.5	1.0
T	**15.4**	**14.1**	**12.2**	**3.2**
U	0.5	0.4	0.6	-0.1
V	0.0	0.0	0.3	-0.3

W	0.1	0.2	0.4	-0.3
X	0.0	0.2	0.1	-0.1
Y	0.1	0.0	0.3	-0.2
Z	0.0	0.0	0.2	-0.2
		99.8	100	

In the revision of MSD, the Ruler suggested under-treatment of the alphabetical stretch B and over-treatment of the stretches K and T in terms of the number of lemmas treated and the number of pages allocated to these categories. It is now the lexicographer's task to analyse these categories in order to ascertain why these alphabetical categories deviate from the Ruler and if corrective action is required. The corpora supply further assistance in the form of candidate lists for inclusion and for omission discussed above.

In the case of the presumed under-treatment of B in MSD, the lexicographer should particularly study the list of candidates for *inclusion* to see if frequently used words were not left out. In the case of K and T the focus should primarily be on the candidate lists for omission to determine whether inclusion of words that do not occur even once in the corpora are justified or not. A detailed analysis of these stretches cannot be given here but a brief analysis will be attempted. By analysing B on suspected under-treatment, gross inconsistencies and omissions were indeed and immediately detected.

Table 5:   Frequently used words in the alphabetical stretch B not included as lemmas in MSD

Lemma	Translation	Tot. freq. in both corpora	In MSD: Yes/No	In Brown: Yes/No	Transl. Eq. in MSD Eng.▸Sets.: Yes/No
banna	men	2 341	No	Yes	Yes
batho	people	9 323	No	Yes	Yes
bona	they; see	18 128	No	Yes	Yes
botlhe	all	1 829	No	Yes	Yes
batsadi	parents	1 516	Yes	Yes	No

The policy of MSD is to include plural forms as lemmas, e.g. *batsadi*. However, lemmas such as *banna*, *batho*, *bona*, *botlhe* and *bosigo* were excluded even though they

(a)   occur more than a thousand times in the corpora,

(b)   were included in the 1925 *Secwana–English Dictionary* of Brown of which MSD is a revision, and

(c)   are given as translation equivalents in the reverse side of MSD.

For the alphabetical stretches K and T, the lexicographer should critically evaluate the huge number of hapaxes (words occurring once only in a corpus)

and zero frequencies given in the candidate lists for deletion in MSD, i.e. 1 664 lemmas (56.7% of all lemmas) for K and 1 812 (49.2%) for T.

The use of Rulers and Block Systems in the compilation or revision of dictionaries, does not mean, however, that the status of *hapax* or *zero-occurrence* in corpora is per definition a directive for omission. In the compilation of a lemma list for a restricted dictionary for very specific target users, De Schryver and Prinsloo (2003: 42-44) justified an extreme case of lemma selection/omission by including words that have a zero frequency in the dedicated corpus as lemmas but excluding words occurring up to nine times in the dedicated corpus.

Microstructural revision strategies

On the microstructural level, comment on semantics is the most important component or data type that, for a bilingual dictionary, should be presented mainly in the form of translation equivalent paradigms. Gouws (1989: 113) states that it is the information type most generally consulted by target users, most substantial and considered as the central component of the article.

> Vir die deursneewoordeboekgebruiker is betekenis die inligtingstipe wat die algemeenste in woordeboeke nageslaan word. As 'n mens na die struktuur van 'n woordeboekartikel kyk, is dit ook duidellik dat betekenisbeskrywing nie net die omvangrykste komponent van die artikel is nie maar dat dit ook as die sentrale deel van 'n woordeboekartikel beskou moet word.

In MSD, this is clearly not the case. Translation equivalents are to a large extent overshadowed by morphological and grammatical information, by the piling up of source language synonyms, etc. Compare the first few articles taken from a single, randomly selected page in MSD.

(3)

matlhagatlhaga ABS. N. CL. 6 *ma-*, NO SING., industriousness; activity.
matlhajana N. CL. 6 *ma-*, PL. OF *letlhajana*, shelves.
matlhaje N. CL. 6 *ma-*, PL. OF *letlhaje*, same as *matlhajwa*, a species of berry-yielding bush; Diospyros lycioides.
matlhakang N. CL. 6 *ma-*, NO SING., DER. F. *tlhakana*, a mixed, or motley lot.
matlhakola N. CL. 6 *ma-*, COLL. PL. OF *letlhakola*, Euclea spp.
matlhakola N. CL. 6 *ma-*, NO SING., DER. F. *tlhakodisa*, a remnant. ID. EXPR., *matlhakola a a dipêpa*, a bare remnant.
matlhaku N. CL. 6 ma-, PL. OF *letlhaku*, cut branches.

It is clear from (3) that comment on semantics takes a secondary place to detailed comment on form made even more prominent by the use of capital letters and to the piling up of source language synonyms sometimes even resulting in the total omission of any comment on semantics:

(4)

todi N. CL. 9N-, SING. OF *ditodi*, same as *lelodi* and *kgobati*.

Another aspect that should be corrected in the revision of MSD is inconsistent labelling and grammatical descriptions:

(5)

gotlhe ENUM. QUAL. CL. 15 < go-, see tab. p. xviii, all; altogether; entirely.
jotlhe CL. 14 QUANT. S., all; the whole, see tab. p. xix.
rotlhe QUAT. USED WITH SUBST. AND IN PARTICULAR THE ABST. PROV., rona, all of us.
tsotlhe QUANT. QUAL., USED WITH CONSTRUCTIONS OF CLS. 8 AND 10, all.
yotlhe QUANT. QUAL. USED TO QUAL. NOUNS AND PRON. CL. 9, all; nama yotlhe, all the meat.

In (5), a variety of grammatical labels, abbreviations and treatment styles are used to refer to quantitatives including punctuation errors and incorrect cross-references. As for punctuation, errors that need to be corrected include double commas, double full stops, grammar labels not followed by a full stop, etc.

    For articles such as (6) that contain a translation equivalent paradigm of unrelated meanings, a homonymic approach should be considered as in (7).

(6)

ntlha N. CL. 9N-, SING. OF *dintlha*, a point; an item; a side; the first. INTERJ. EXPR., surprise; wonder.

It could be argued that translation equivalents such as 'a point', 'a side', 'the first' and 'idea' are not merely different senses but unrelated meanings that should accordingly be treated as homonyms:

(7)*

ntlha[1] *num.* **1** first: **ke motho wa ntlha go nwa tee**, *he is the first person to drink tea;* **2** beginning: **lwa ntlha o ne a itumetse**, *in the beginning he was happy*
ntlha[2] *conj.* but, by the way: **ntlha e ka re re a latlhega**, *but it seems we are getting lost*
ntlha[3] *n.* side: **o ntse ntlha ya lokotswana**, *he is sitting on the side of the wall*
ntlha[4] *n.* end, point: **o bone ntlha ya teng e bogale**, *beware of its end, it is sharp*
ntlha[5] *n.* point, idea: **o tsile ka ntlha e e botlhokwa**, *he came with a good point*

For the lemmatisation of verbs, the treatment of a randomly selected verb, *ga-gaba* in MSD as well as in a few other Setswana dictionaries can be considered.

(8)

(a)   *Dikišinare ya Setswana–English–Afrikaans* (Snyman 1990)
       gágábā, slither (eg a snake) // seil (eg 'n slang)

(b)   *Thanodi ya Setswana* (Kgasa and Tsonope 1995)
       gagaba GGG tpt. **–ile.** Tsamaya ka diatla le mangole

(c)   *Secwana–English Dictionary* (Brown 1925)
       Gagaba, v.i., pft. gagabile, creep or crawl, on hands and knees; crawl, as
       a cat hunting.

(d)   MSD
       gagaba V. S. SIMP., same as gogoba, creep or crawl, on hands and knees;
       crawl, as a cat hunting.

In comparison, consider the following extract from the concordance lines generated for *gagaba* from the corpora:

Table 6:   Concordance lines for *gagaba*

..aetsega a tsena mo lobaleng **a**	**gagaba**	**ka mangole le diatla** a reedi
ba a sale a tshwana le **noga, a**	**gagaba**	**ka mpa** mo loroleng
aana mmoki a boka a ba a sala a	**gagaba**	**ka dimpa** fa fatshe. Moji a ts..
a tlhoka: **Maru a bo a tlhaga a**	**gagaba**	go tswa borwa. Botsho ba matl
go ikatametsa fa go tsona ka go	**gagaba**	**ka matsogo le mangole.** O ne a
se bonela mo lefifing se tla se	**gagaba**	**jaaka katse e ratela legotlo,**
tsaya motlhala wa mo mosong. Ba	**gagaba**	**ka iketlo, dikoloi di tletse**

A single glance at these concordance lines reveals that *creep* or *crawl* are indeed core senses of *gagaba* in relation to humans, animals and reptiles but also senses such as *slow movement* of e.g. clouds or traffic.

    In (8)(a) the translation equivalent *slither* with reference to 'snake' is given but not in any of (8)(b)–(8)(d). In (8)(b) the definition is limited to 'move with hands and knees' which defines one of the core senses of *gagaba* but excludes this kind of movement for all animals and reptiles. In (8)(c) and (8)(d) movement of humans and animals are well captured but not that of reptiles nor the sense of slow movement. In an attempt to improve on MSD's article for *gagaba*, and in fact on all of (8)(a)–(8)(d), the following treatment is suggested for the lemma *gagaba*.

(9)

gagaba *v.* **1** crawl, creep: *~ ka diatla le mangole*, crawl on hands and knees;
    ~ jaaka katse e ratela legotlo crawl like a cat stalking a mouse **2** slither: *noga
    e ~ ka mpa mo loroleng* the snake slithers on its belly in the dust; **3** move slowly;
    maru a ~ go tswa borwa clouds move in from the South

Articles (7) and (9) represent an attempt to improve on typical articles for nouns and verbs in MSD such as (6) and (8) by putting much more emphasis

on the comment on semantic, less on the comment on form, and to maximally use corpus data for sense distinction, frequent collocations, authentic examples, etc. in the treatment of such lemmas.

Conclusion

In this article an attempt has been made to formulate a typical revision strategy for substantial revision of a Setswana dictionary representing a case where in Landau's terms, revising should take on *a desperate character*. In all the official African languages of South Africa, many dictionaries exist that are outdated and in need of such a fundamental revision. Since electronic corpora exist for these languages, the strategies presented here could be considered for such revisions. Much emphasis has been placed on revision on the macrostructural level because it is believed that the dilemma of what to include in or exclude from the lemma list of especially a single-volume paper dictionary in terms of Busane (1990), is likely to remain 'forever'. It is therefore imperative for the lexicographer to be able to motivate inclusion/omission of lemmas in terms of sound lexicographic and statistical principles and only then to proceed to maximally utilise concordance lines to enhance microstructural treatment of these lemmas.

Endnote

*   The original draft of this article for the lemma *ntlha* is credited to Mr Thapelo Otlogetswe.

References

Brown, J.T. 1925. *Secwana–English Dictionary*. Lobatsi: London Missionary Society.

Busane, M. 1990. Lexicography in Central Africa: The User Perspective with Special Reference to Zaïre. Hartmann, R.R.K. (Ed.). 1990. *Lexicography in Africa*: 19-35. Exeter Linguistic Studies 15. Exeter: Exeter University Press.

De Schryver, G.-M. 2003. Drawing up the Macrostructure of a Nguni Dictionary, with Special Reference to isiNdebele. *South African Journal of African Languages* 23.

De Schryver, G.-M. and D.J. Prinsloo. 2000. Electronic Corpora as a Basis for the Compilation of African-language Dictionaries, Part 1: The *Macrostructure. South African Journal of African Languages* 20(4): 291-309.

De Schryver, G.-M. and D.J. Prinsloo. 2003 Compiling a Lemma-sign List for a Specific Target User Group: The Junior Dictionary as a Case in Point. *Dictionaries* 24: 28-58.

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

Gouws, R.H. 1991. Toward a Lexicon-based Lexicography. *Dictionaries* 13: 75-90.

Gouws, R.H. and D.J. Prinsloo. 1997. Lemmatisation of Adjectives in Sepedi. *Lexikos* 7: 45-57.

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

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

Kriel, T.J. 1983. *Pukuntšu Dictionary*. Pretoria: J.L. van Schaik.

Landau, S.I. 2001. *Dictionaries: The Art and Craft of Lexicography.* Cambridge: Cambridge University Press.

Matumo, Z.I. 1993. *Setswana–English–Setswana Dictionary*. Gaborone: Macmillan.

Prinsloo, D.J. 1996. Review: Robert Botne and Andrew Tilimbe Kulemeka: *A Learner's Chichewa and English Dictionary* (Afrikawissenschaftliche Lehrbücher 9). *Journal of African Languages and Linguistics* 17(2): 199-202.

Prinsloo, D.J. and G.-M. de Schryver. 2002. Designing a Measurement Instrument for the Relative Length of Alphabetical Stretches in Dictionaries, with Special Reference to Afrikaans and English. Braasch, A. and A. and C. Povlsen (Eds.). 2002. *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 13–17, 2002*: 483-494. Copenhagen: Center for Sprogteknologi, University of Copenhagen.

Prinsloo, D.J. and G.-M. de Schryver. 2003. Effektiewe vordering met die *Woordeboek van die Afrikaanse Taal* soos gemeet in terme van 'n multidimensionele Liniaal [Effective Progress with the *Woordeboek van die Afrikaanse Taal* as Measured in Terms of a Multidimensional Ruler]. Botha, W. (Ed.). 2003. *'n Man wat beur. Huldigingsbundel vir Dirk van Schalkwyk*: 106-126. Stellenbosch: Buro van die WAT.

Rundell, M. (Ed.). 2002. *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan.

Snyman, J.W. (Ed.). 1990. *Dikišinare ya Setswana–English–Afrikaans Dictionary/Woordeboek*. Pretoria: Via Afrika.

Stevenson A. 2004. *Revising a Dictionary* [online]. Available at <http://www.askoxford.com/worldofwords/worddetectives/revising/?view=uk>. [Accessed 28 May 2004.]

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

Semi-Automatic Retrieval of Definitional Information: A Northern Sotho Case Study [*]

Elsabé Taljard, *Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (etaljard@postino.up.ac.za)*

Abstract: Corpus-based terminology is currently gaining ground on the international front. It is therefore important that terminologists working on the South African Bantu languages not only take note of this development, but that they should also follow this trend, even if they do not have the same measure of access to highly sophisticated software. The aim of this article is therefore to establish whether it is possible to retrieve definitional information on key concepts from untagged, running text by making use of affordable and easily accessible software such as *WordSmith Tools*. In order to answer this question, a case study is done in Northern Sotho, using textual material on linguistics as basis for a special field corpus. Syntactic and lexical patterns serving as textual markers of definitional information are identified and the success rate of the computational retrieval of definitional information is analysed and evaluated. Attention is also paid to the retrieval of specifically conceptual information, which turned out to be a fortunate by-product of semi-automatic retrieval of definitional information. Finally, it is illustrated how definitional information retrieved can be utilised in the writing of a formal terminological definition.

Keywords: TERMINOLOGY, SOUTH AFRICAN BANTU LANGUAGES, DEFINITIONAL INFORMATION, SEMI-AUTOMATIC INFORMATION RETRIEVAL, TERMINOLOGICAL DEFINITIONS, CONCEPTUAL RELATIONSHIPS, LEXICAL PATTERNS, SYNTACTIC PATTERNS, TEXTUAL MARKERS, KEYWORD-IN-CONTEXT (KWIC), WORDSMITH TOOLS

Opsomming: Semi-outomatiese herwinning van definisie-inligting: 'n Noord-Sothogevallestudie. Korpus-gebaseerde terminologie is tans besig om veld te wen op die internasionale front. Dit is daarom belangrik dat terminoloë wat binne die Suid-Afrikaanse Bantoetale werk, nie net sal kennis neem van hierdie ontwikkeling nie, maar dat hulle ook hierdie neiging sal volg, selfs al het hulle nie dieselfde mate van toegang tot gesofistikeerde rekenaarprogrammatuur nie. Die doel van hierdie artikel is daarom om vas te stel of dit moontlik is om definisie-inligting oor sleutelkonsepte uit ongemerkte, lopende teks te herwin deur bekostigbare en toeganklike sagteware soos *WordSmith Tools* te gebruik. Ten einde hierdie vraag te beantwoord, is 'n gevallestudie in Noord-Sotho gedoen, met gebruikmaking van teksmateriaal oor die linguistiek as basis vir 'n gespesialiseerde korpus. Sintaktiese en leksikale patrone wat as tekstuele merkers van defini-

---

[*] An earlier version of this article was presented as a paper at the Eighth International Conference of the African Association for Lexicography, organised by the Department of Germanic and Romance Languages, University of Namibia, Windhoek, Namibia, 7–9 July 2003.

sie-inligting dien, word geïdentifiseer en die suksesratio van rekenaarmatige herwinning van definisie-inligting word ontleed en beoordeel. Aandag word ook gegee aan die herwinning van spesifiek konseptuele inligting, wat 'n onverwagse byproduk van die semi-outomatiese herwinning van definisie-inligting is. Ten slotte word geïllustreer hoe definisie-inligting aangewend kan word by die skryf van 'n formele terminologiese definisie.

Sleutelwoorde: TERMINOLOGIE, SUID-AFRIKAANSE BANTOETALE, DEFINISIE-INLIGTING, SEMI-OUTOMATIESE INLIGTINGSHERWINNING, TERMINOLOGIESE DEFINISIES, KONSEPTUELE VERHOUDINGE, LEKSIKALE PATRONE, SINTAKTIESE PATRONE, TEKSTUELE MERKERS, KEYWORD-IN-CONTEXT (KWIC), WORDSMITH TOOLS

1.   Electronic corpora and terminology — an overview of the current international and national scenario

The use of electronic or machine-readable corpora in general lexicography is a well-established practice, not only on the international front, but also within the South African set-up, where the nine National Lexicography Units for the South African Bantu languages[1] are all to a greater or lesser extent using a corpus-based approach for the compilation of their various dictionaries. According to Ahmad and Rogers (2001: 728), however, the use of corpora for terminological or LSP purposes has been accepted much more slowly, with the use of corpora largely being restricted to general lexicography, one reason being the more prescriptive orientation of terminology, and the other its onomasiological orientation, i.e. concept-based approach. With reference to the international scene, Pearson (1998: 1) identifies three possible reasons for the seeming reluctance of terminologists and terminographers to recognize the vital role that electronic corpora can and should play in terminology work.

In the first instance she points out that in the past it was indeed difficult to get hold of specialized corpora needed for terminological purposes. However, the increasing availability of especially electronic texts has now made it possible for terminologists to build their own specialized corpora with relatively little effort. The unavailability of suitable textual material that can be used in the compilation of electronic corpora is indeed also relevant for the South African context, and specifically for the South African Bantu languages. Due to the political and educational dispensation of the past two decades, very little has been produced in the South African Bantu languages in the line of special field texts. Subject material is written largely in English and/or Afrikaans, thus denying terminologists the opportunity to base their terminological work on authentic special field texts. Fortunately, during recent years, more and more technical material written in the South African Bantu languages has become available, mainly on the Internet, often as translations of mostly English source texts. Furthermore, although the system of 'Bantu education' as implemented by the previous regime surely has no redeeming qualities, one of the by-prod-

ucts of this system was indeed the publication of school textbooks on special subject fields in at least some of the Bantu languages. Due to the fact that the medium of instruction in former black schools is mostly English, these textbooks are no longer used, but can with a little effort be sourced from school storerooms and archives of libraries. Despite their political baggage, these textbooks can be of great value for the building of special field corpora. South African terminologists, especially those working in the Bantu languages, should therefore be encouraged to make use of this material to compile their own specialized corpora.

A second reason why corpora have not been utilised for specialized lexicography and/or terminology is the perception that terms are context independent — a notion that has dominated terminological work for quite some time. This issue is raised by both Pearson (1998: 1) and Sager (2001: 761), who point out that it is only recently that the emphasis has shifted to studying terms in their communicative, i.e. linguistic context. Within the 'traditional' approach, terms are regarded as separate items, "forming part of a semi-artificial language, deliberately devoid of any of the functions of other lexical items" (Sager 2001: 761). Within the modern, corpus-based approach, terms are viewed as lexical items accorded term status based on the communicative setting in which they appear. Real texts are therefore now regarded as primary sources of terminological data, providing information on the meaning, usage and appropriateness of a term, as conditioned by the textual environment in which it appears. Moreover, as Shreve (2001: 773) indicates, when terms co-occur in a text, they establish conceptual relationships. Thus, by analysing the text, important information on the larger conceptual structure of a specific knowledge domain can be retrieved. In South Africa, there still seems to be some reluctance on the part of terminologists to rely on authentic text for terminological data. Currently, terminological work done by terminologists working in the Bantu languages is mainly of a translational nature, consisting of the finding of translation equivalents for English/Afrikaans terms, based mainly on mother-tongue intuition and consultation with subject-field specialists. Even so, this should not exclude the possibility of utilising corpora — as Bowker and Pearson (2002: 14, 20) point out, corpora are useful complements to all other types of resources and should not be viewed as replacements for these. Making use of especially comparable and parallel corpora within this particular terminological environment could contribute much to the quality of the terminological effort.

In the third instance, it is generally accepted that the input of special field experts is indispensable in the identification and definition of terms. The basic premise here is the "conviction that terms are different from words and can only be defined by suitably qualified subject specialists" (Pearson 1998: 1). 'Traditional' terminologists therefore rely to a large extent on subject experts for the identification and definition of terms. Again, having access to electronic corpora does not imply disregarding the input of experts — the information gleaned from corpora can be used to supplement, support and validate the

judgement made by experts, which in practice often represents one person's opinion. Within the South African situation, consultation with experts is still the preferred methodology for terminology work. Unfortunately, there seems to be a lack of commitment on the part of special field experts who are mother-tongue speakers of the South African Bantu languages to develop terminology in these languages, the basic assumption being that English is the language of special subject fields. Furthermore, ongoing consultation with subject-field specialists throughout the terminological process is a time-consuming and labour-intensive exercise, and thus makes the seeming reluctance of subject-field specialists to co-operate understandable, although not defendable.

If South African terminologists, especially those specializing in the Bantu languages wish to keep abreast of terminological developments on the international scene, corpus-based terminology is no longer an option, but an imperative. As was pointed out above, the increasing availability of subject-field texts written in these languages — many of them in electronic format — now enables terminologists to build their own corpora for special purposes. Furthermore, access to user-friendly and affordable software such as *WordSmith Tools* opens the door for terminologists to query and analyse these corpora automatically or at least semi-automatically[2]. It has already been illustrated by Taljard and De Schryver (2002) that it is indeed possible to extract terms semi-automatically from corpora based on subject-field texts, thus reducing (but of course not eliminating) the dependence of the terminologist on the co-operation of the subject-field specialist.

2.   Rationale

In this study, the feasibility of retrieving definitional information semi-automatically from special field corpora is investigated. For the purpose of this study, the term 'definitional information' is used to refer to any information to be found in an electronic special field corpus regarding the meaning and usage of a term, as well as the conceptual relationships it has with other terms. In this regard, two issues will be addressed. In the first instance, Pearson (1998: 5) states that authors writing within certain specified communicative settings are likely to provide explanations of at least some of the terms they use. This hypothesis is tested with regard to Northern Sotho, using a special purpose corpus consisting of a collection of texts on linguistics as authentic data. Secondly, the possibility of retrieving definitional information in a semi-automatic way from the corpus is investigated. By way of conclusion, an example will be given as to how the definitional information retrieved from the corpus can be used for the writing of a terminological definition. Before these issues are addressed, however, current methodological options open to South African Bantu language terminologists with regard to the generating of definitional information are investigated.

3.   Generating definitional information: the current South African scenario

It has already been pointed out that it is indeed possible for South African ter-
minologists to compile their own special field corpora, and, by following the
methodology suggested by Taljard and De Schryver (2002), to semi-automati-
cally extract terms from electronic texts. Should the terminologist now want to
add definitions to the extracted terms, he/she is currently left with two op-
tions: (a) to formulate definitions with the help of a subject-field expert, or (b)
to provide translational equivalents for the terms in English/Afrikaans, then
search for definitions in either an LSP dictionary or existing term lists, and as a
last step, translate the definitions from English/Afrikaans into the appropriate
Bantu language. Both these options have inherent pitfalls that often impact
negatively on the quality of the terminological activity, and therefore warrant
some discussion.

A number of potential problems regarding the process of consultation
with experts have already been touched upon: it is time-consuming, it is la-
bour-intensive and the results often represent the view of only one person. The
ideal is of course that multiple experts be involved, but this is in practice not
always a realistic aim. Furthermore, terminologists need a very specific skill in
order to elicit the correct and appropriate information from subject-field spe-
cialists. As Bowker and Pearson (2002: 17) point out, terminologists must be
careful not to ask leading questions, which would result in obtaining distorted
information. Furthermore, not being a subject-field expert, the terminologist is
often not equipped to distinguish between relevant and irrelevant information,
or to judge whether the proffered information is merely a personal opinion
expressed by the expert. It is therefore clear that terminologists should receive
at least some measure of training in the conducting of interviews with subject-
field specialists. With terminology training being in its infancy in South Africa,
it cannot be taken for granted that terminologists do indeed get this kind of
training.

The second option is also not without its disadvantages. According to
Bowker and Pearson (2002: 15), one of the inherent problems with LSP diction-
aries is their incompleteness. Printed dictionaries tend to become out-dated
rather quickly, especially in the fields of science and technology, which are
characterised by rapid development. Consulting an LSP dictionary therefore
does not guarantee retrieval of the current state of knowledge in a particular
subject field. Furthermore, LSP dictionaries often do not provide adequate
contextual or usage information — another point raised by Bowker and Pear-
son (2002: 16). Even if the terminologist succeeds in finding definitions in
existing LSP dictionaries, there is always the danger of incompatibility of the
target group served by the dictionary and the target group which the termi-
nologist has in mind. A definition sourced from an LSP dictionary with post-
graduate chemistry students and chemistry experts as target users will for
example not be suitable if the target audience for whom the terminologist is

compiling a terminological reference source is senior secondary school learners. A mechanistic translation of terminological definitions would in this instance serve no purpose. Translation of definitions is another area fraught with problems. Technical translation requires a high level of translation skills. It is a well-known fact that there is a shortage of well-trained and qualified translators for the South African Bantu languages, a shortage that is even more serious when it comes to translators specializing in technical translation. Making use of a translator who does not have the necessary background in technical translation will undoubtedly result in terminological definitions of poor quality.

Utilizing an electronic special field corpus as a source of definitional information is not the solution to all of the problems identified above, but it does have distinct advantages. In the first instance, being in electronic format corpora can be updated much more easily than for example, a paper LSP dictionary. By querying a corpus that is regularly updated, the terminologist is more likely to find the most recent information on a particular term. Secondly, the corpus provides the terminologist with a wealth of usage and textual information. Should the terminologist wish to include examples of use as part of the definition of a term, authentic examples can be sourced directly from the corpus. Thirdly, the terminologist can obtain a higher degree of compatibility between the target user of the textual material incorporated in the electronic corpus and the target reader of the final terminological product. This can be done by making sure that the texts that are selected for inclusion in the special field corpus have the same target user as the terminological end product. Lastly, since the textual material used for the compilation of the corpus is already in the target language, the input of a technical translator is no longer necessary. Also, less time needs to be spent on consultation with special field experts, since they are only required to verify, supplement or reject the information gleaned from the corpus.

4.   Semi-automatic retrieval of definitional information — a case study

4.1   Compilation of an electronic special field corpus

For the purpose of this investigation, a special field corpus was compiled, based on a number of texts on Northern Sotho linguistics, which were kindly provided in electronic format by Prof. L.J. Louwrens of the Department of African Languages of the University of South Africa (UNISA). These texts were reverted to text-only format in order to make it compatible with the software *WordSmith Tools* (WST), which was to be used as a corpus-querying tool. After a simple count by making use of WST's *WordList* function, it was revealed that this special field corpus contains 74 251 tokens (running words) and 4 744 types (unique words). It has to be borne in mind that this particular corpus is unmarked and untagged, as are all the corpora which are currently available for Northern Sotho. Even so, making use of raw corpora, i.e. running texts, does

have certain advantages, as has been indicated by Sager (2001: 764, 765). In the first instance, he points out that "terminology extracted from running text or discourse offers a greater guarantee of thematic completeness and coherence". The linguistic behaviour of terms can be deduced from suitably selected contexts. Secondly, terminology is dynamic, in that new terms are continuously coined and added to the body of knowledge of a particular subject field, sometimes even replacing existing terms. Terms also become obsolete. Running text therefore gives a good indication of the actual existence and currency of terms. Thirdly, running text can assist the terminologist in checking the correctness of previously entered terminology, and to update the terminological database if necessary. However, even though utilising running text is a fruitful exercise, the ideal should always be to work towards establishing a more sophisticated tool in the form of a marked and tagged corpus.

4.2   Identification of 50 single word test terms

As a next step, the *KeyWord* tool was used to identify the 50 raw, i.e. unlemmatised single word terms which have the highest frequency in the corpus. This is done by comparing the frequency of each item in the *WordList* of the special purpose corpus with its frequency in a second, much larger reference corpus. As a reference corpus, the *Pretoria Sepedi Corpus* (PSC) of the University of Pretoria was used. The PSC is an organic corpus and the version used for this study contains roughly 5.9 million words. All items that display a great disparity in frequency are identified as keywords, since the disparity would imply that that specific item occurs with unusual frequency in the smaller corpus. This whole process of keyword identification is done automatically; the only required human intervention being to read through the suggested keyword list and to decide on term status[3]. The 50 test terms identified in this manner are listed in (1).

(1)   50 unlemmatised test terms, extracted semi-automatically from the special field corpus

N	Term	Translation	N	Term	Translation
1	babolediŝani	*interlocutors*	26	mantŝu	*(linguistic) words*
2	deiktiki	*deictic*	27	maŝala	*pronouns*
3	dikafoko	*phrases*	28	mediriŝo	*moods*
4	direwa	*topics*	29	mmolediŝwa	*addressee, 2nd pers.*
5	ditlaleletŝo	*complements*	30	modiriŝo	*mood*
6	kamano	*(inter)relationship*	31	modiriŝogore	*subjunctive mood*
7	kgatelelo	*emphasis*	32	modiriŝokanegelo	*consecutive mood*
8	kgokagano	*discourse, communication*	33	modiriŝopego	*indicative mood*
9	lebopikganetŝi	*negative morpheme*	34	modiriŝopegotlhaodi	*situative mood*

10	lediri	*verb*	35	modirišotaelo	*imperative mood*
11	lefoko	*sentence*	36	modiro	*function*
12	legoro	*(noun) class*	37	phetlekokgokagano	*discourse analysis*
13	lehlathi	*adverb*	38	poledišano	*dialogue*
14	leina	*noun*	39	sediri	*subject*
15	leinataodi	*head noun*	40	sedirwa	*object*
16	lekgokasediri	*subject concord*	41	seemotikologo	*discourse context*
17	lekgokedi	*agreement morpheme*	42	serewa	*topic*
18	lereo	*term*	43	tatelanontšu	*word order*
19	lešala	*pronoun*	44	tiro	*predicate*
20	lethuši	*auxiliary verb*	45	tlhološo	*meaning*
21	mabopi	*morphemes*	46	tlhološotheo	*basic meaning*
22	madiri	*verbs*	47	togaganyo	*cohesion*
23	mafoko	*sentences*	48	tšhupetšogotee	*coreference*
24	mafokofokwana	*complex sentences*	49	tswalane	*being related to*
25	mafokotheo	*basic sentences*	50	tswalano	*relationship*

4.3   Semi-automatic retrieval of definitional information — a case study

4.3.1   Isolating Concordance lines for the 50 test terms

As was stated earlier, the first aim of this investigation is to establish whether definitional information is indeed provided in the text. For this purpose, the *Concord* tool of WST is used. When the selected term is entered as a search node, a list of concordance lines or KWIC *(keywords-in-context)* lines in which that specific term appears, is automatically thrown up, thus placing the term within its textual context. Compare the example in (2) where the term **lediri** 'verb' has been used as a search node.

(2)   Concordance lines (KWIC lines) for the term **lediri** 'verb'

Leina le o tšogo le ngwalolla le tlaleletša	**lediri** a rwala ka tsela efe? Ke tlaleletšo
mantšu a mabedi, e lego lethuši o hlwa le	**lediri** a bala. Mehlala ye mengwe ke
a tlaleletšago modiro wo o hlalošwago ke	**lediri** a bitšwa didirwa, gomme mo
la Legoro 1 -thuš-: modu wa	**lediri** -a : mosela woo lediri le felelago
a mararo, e lego (i) sediri baithuti, (ii)	**lediri** ba rekile; le (iii) sedirwa dipuku
o ngwalolle lediri. Ge tiro e le lethuši le	**lediri** go nyakega gore o ngwalolle
Le gona re boletše gore	**lediri** le ka tlaleletšwa ka ditlaleletši tša

The number of concordance lines differs from one term to the next, depending on how many times the specific term appears in the text. By studying the KWIC lines thrown up for every term, one can then establish whether definitional information is provided within the text for that specific term. For these 50 test terms, the total number of KWIC lines that had to be perused, is 4 246. Definitional information was found in 292 of these KWIC lines, covering 45 of the selected 50 terms. These findings therefore support Pearson's hypothesis

that definitional information is provided by writers of technical texts. The authors of these specific technical texts do indeed, whether advertently or inadvertently, provide definitional information on the terms they use.

With regard to the terms on which no information was found, a few remarks can be made. The first term on which no information was found, is **baboledišani**. Although this term has the specialised meaning of 'interlocutors' within the linguistic context, its meaning is to a large extent self-evident, possibly because of its morphological structure. It is a deverbative noun, derived from the verbal root **bolel-** 'speak', affixed with a causative suffix **-iš-** and a reciprocal **-an-**. The prefix **ba-** of course indicates that the referents of this term are human beings. Due to its morphological make-up, it is therefore quite easy to derive the meaning of this term and explains why the author of this particular text did not deem it necessary to provide definitional information on the term. The next two terms on which no information was found, are **modiro** 'function' and **lereo** 'term'. These lexical items belong to a category which is known as non subject-specific-specialised vocabulary or subtechnical terms. Pearson (1998: 19) describes subtechnical words as general language words that have taken on specialised meanings in more than one domain, and due to their concomitant high frequency of use could be assumed as being known. With regard to the term **mabopi** 'morphemes' for which no information could be found, it could be argued that this is such a basic concept within the field of linguistics, that it is not surprising that no definitional information was provided, especially when it is borne in mind that the texts used in this particular study target senior university students. The last term for which no definitional information could be retrieved, is **tswalane** 'being related to'. This was the only verb that was found amongst the 50 test terms. In the study done by Taljard and De Schryver (2002), a total number of 350 terms were extracted from the same texts on which the current study is based. Of these, only 14, i.e. 4% were verbs, of which **tswalane** was one. A concordance search on the other 13 verbs revealed that definitional information is not provided for any of these terms. It therefore seems that retrieving definitional information on verbal terms does pose a problem.

4.3.2 Identification of textual markers of definitional information

Studying KWIC lines in search of definitional information obviously represents an improvement on the more traditional method of physically reading through all the textual material in search of definitional information. Ahmad and Rogers (2001: 740) point out that reading a text to extract any kind of terminological data is labour-intensive and potentially repetitive, because of the need to recover different kinds of terminological data. They furthermore indicate that the computational processing of relatively large quantities of text may allow patterns to emerge which are unlikely to be detected by manual scanning. The reference to the possible existence of patterns immediately suggests the possi-

bility of semi-automatic recognition of definitional information. Pearson (1998: 103) indicates that definitional information can be signalled either by syntactic or lexical devices. It therefore follows that if the strategies or patterns which are used to signal definitional information in texts can be identified, it would enable the terminologist to semi-automatically retrieve at least some definitional information from these texts. It should be kept in mind that the corpus utilized for the purpose of this case study is an untagged one. The initial aim was simply to identify lexical markers of definitional information, similar to the way in which Pearson (1998: 136 et seq.) identifies fillers for the slots **X**, **Y** and **=**, in what she terms a formal terminological definition. She identifies two variant formulae that characterise a formal definition:

(3)

<u>Formula 1</u>: **X = Y + distinguishing characteristic, whereby X is subordinate to Y**
Example: an adverb (X) is (=) a linguistic word (Y) which qualifies a verb (distinguishing characteristic). ['Adverb' being a subordinate to the superordinate 'linguistic word'.]

<u>Formula 2</u>: **Y + distinguishing feature = X, whereby X is subordinate to Y**
Example: a linguistic word (**Y**) which qualifies a verb (**distinguishing characteristic**) is (called) (=) an adverb **(X)**. ['Adverb' being a subordinate to the superordinate 'linguistic word'.]

In these formulae, **X** is obviously the term being identified, **Y** the superordinate, and **=** the linguistic device which links the term to its superordinate. She then proceeds by identifying, for English, typical lexical fillers for the slots **Y** and **=**. She indicates that **Y** must either refer to a term (which is the case in the illustrative example above) or to a class word. Typical class words which are to be found in this slot are: 'technique, method, process, function, property, system, class, device', whereas the **=** slot is filled by connectives such as 'comprise(s), consist(s) of, define(s), denote(s), is/are, is/are called, is/are known as', etc. In her study, Pearson focuses mainly on definitional information that is provided in the format of these formal definitions. However, with regard to the Northern Sotho case study, the objective is to retrieve as much definitional information as possible on each term, and not only information provided in the form of a formal definition. Furthermore, although the initial aim was to identify lexical markers that signal definitional information, it quickly became clear that even in an untagged corpus, certain syntactic patterns can be identified as strategies by means of which definitional information is marked. Both these strategies are of value to the terminologist, the main advantage of lexical markers being that definitional information signalled by these items, can be retrieved by using a basic word-processing tool such as *MSWord* or *WordPerfect*, simply by making use of the *Search* function. This makes the manual scanning of all textual material unnecessary, since the terminologist can automatically

pinpoint all occurrences of the particular term in the text where definitional information is provided on it. Syntactic patterns that signal definitional information, on the other hand, will be of much value when hopefully in the near future, Northern Sotho will have a tagged corpus, enabling the terminologist to automatically search for these patterns.

4.3.3 Lexical and syntactic markers of definitional information in Northern Sotho

The lexical items as well as the syntactic patterns found as markers of definitional information in the special purpose corpus on Northern Sotho linguistics are given in (4):

(4)   Table 1 (See attached spreadsheet)

The procedure followed to arrive at these results can be briefly explained: the 292 KWIC lines in which definitional information was found, were scrutinised in order to identify any textual markers which signalled the presence of definitional information on any of the 45 terms. These markers of definitional information function on a very non-theoretical level. They simply indicate that some form of definitional information on a specific term is to be found in the textual vicinity flagged by these markers. The definitional information can either precede or follow the term and its marker, and may even be spread across sentence boundaries. After the markers had been identified, these were then indicated on a spreadsheet by means of a cross next to the term on which it provided information. From the information provided under B (leftmost column), for example, it is clear that for 22 of the 45 terms, definitional information is textually marked by the presence of the identifying copulative particle **ke**, which follows the term (**T**) in question. Markers that appeared only once were not taken into account. Analysis of the marking strategies brought to light that two distinct kinds of strategies could be identified, i.e. general strategies and strategies which are suspected to be subject-specific. The strategies in columns B, C, D, E, G, H, J and K, as well as those in the unshaded columns under F, represent general strategies. This means that regardless of the subject field, these strategies would in all probability function as markers of definitional information in Northern Sotho texts. The strategies listed in the shaded columns under F and the one in column I are of a slightly different nature. Their function as marking strategies are probably directly linked to the lexical items that are found in these patterns and the relationship of these items with the particular subject field. Verbs such as **ganetšwa** 'is/are negated', **šuthišwa** 'is/are moved', **tlogelwa** 'is/are deleted', etc. form such an inherent part of the subject field of Northern Sotho linguistics, that it is highly improbable that these strategies would be useful in identifying definitional information on terms appearing in a corpus based on, for example, chemistry or music texts.

TABLE 1

A	B				C			D	E	F								G		H		I			J		K			
Strategies →	T + Identifying cop				Identifying cop + T			lereo le (la) + T 'this term (of)'	lereo le (la) + T + COP 'this term (of) is'	T (+dem) + s.c.(+ka) + PASS V								PASS V + T		PASS V + ke + T		T + s.c. + V			inchoative V + T		Orthographic strategies			
Lexical markers →	ke 'is/are'	e le(go) 'which is/are'	e ba 'become(s)'	e ka ba(go) '(which) can be/become'	ke 'is/are'	e le(go) 'which is/are'	e tla/ka ba 'can/will be'			hlolwa 'is/are caused'	šomišwa 'is/are used'	bitšwa 'is/are called'	ganetšwa 'is/are negated'	šuthišwa 'is/are moved'	tlaleletšwa 'is/are qualified'	tlogelwa 'is/are deleted'	tswalanywa 'is/are connected'	bitšwa 'is/are called'	bopša 'is/are formed'	hlolwa 'is/are caused'	laolwa 'is/are determined'	hlaloša 'mean(s)'	tšweletša 'produce(s), result(s) in'	-na 'has, have'	hlola 'cause(s)'	bopa 'form(s)'	(T)	T(*)	T:	
Terms																														
deiktiki	x																													
dikafoko		x√																	x							x√	xx√				
direwa		x																													
ditlaleletšo	x																														
kamano													xxx√						xx	x											
kgatelelo																										xxxx√					
kgokagano	xx					x√							x√													xx√					
lebopikganetši					xx																										
lediri	x				x	xx√						x				xxxxxx															
lefoko	xxxx√			x																				x		xx√					
legoro																				x			x								
lehlathi	x√	x		x															x√												
leina					xxx√		x																								
leinataodi																			xx√	x											
lekgokasediri	x√													x					xxx√												
lekgokedi						x												xxx				xxx√									
lešala												x	xx																		
lethuši						xx√																									
madiri																													x		
mafoko		x																								xx√	xx√				
mafokofokwana					x√																xx√						xx	x			
mafokotheo	x√																	x													
mantšu	x	x√				x√															x√				xx		xxxx√				
mašala	x√	x										x																			
medirišo						x																									
mmolediišwa	xx√																										xx				
modirišo	xxxxxx																														
mmodirišogore	x					xx		xxxx		x	x												xx						x	
modirošokanegelo								x					x										x							
modirišopego	xxx√				xx																			xx						
modirišopegotlhaodi	x												x																	
modirišotaelo	x√														x															
phetlekokgokagano	x																											x		
poledišano																												x		
sediri		x			xx			xx						x		x	x												x	
sedirwa	x			x										x		x													x	
seemotikologo		x				xx		xxxxxx													xx			x						
serewa	xxxx√		x√		x																							x		
tatelanontšu																												x		
tiro	xx√	xxxxx√		xxxxx√				xxxx								x														
tlhalošo	x																														
tlhalošotheo	xxxx√	x√																													
togaganyo						x		x	x	xx															xx			x		
tšhupetšogotee								x															x							
tswalano																								xx√					x	
	40	14	1	9	14	9	2	16	5	8	4	6	2	2	7	2	2	8	4	7	5	3	2	5	13	11	2	7	3	
Total no. of occurrences — General strategies			64				26	16	5						18				12		12					24			12		189
General and specific strategies			64				26	16	5						18		15		12		12				10	24			12		214
No. of new terms on which definitional info is provided			27					8	2	0						2				1		1			0		2		2		45

This implies that the strategies **T (+dem) + s.c. (+ka) + PASS V** and **T + s.c. + V** would probably function as markers of definitional information in any text, but containing in the VERB (**V**) slot, verbs which are relevant to the particular subject field. This would imply that these verbs would have to be identified anew for every subject field within which the terminological activity takes place.

4.3.4 Analysis of results

With regard to the statistical analysis of the table, two sets of figures are relevant. Firstly, at the bottom of each column, the total number of occurrences for each strategy is given, thus the strategy **Term + Identifying copula** in column B appears in a total of 64 out of a possible 292 KWIC lines. The total number of KWIC lines in which the general strategies appear, is 189. This implies that by using the patterns appearing in these general strategies as a search node in WST's *Concord* tool, roughly 65% of all concordance lines in which definitional information appears, can be retrieved automatically from the corpus. To illustrate this: by using the pattern **Term + ke** as a search node, a total of 40 concordance lines, providing definitional information on 22 different terms, can be culled from the corpus. Compare example (5) in this regard, where an excerpt of 10 concordance lines is given by way of illustration. Translations of the relevant concordance lines have been added in order to illustrate the definitional information to be found in these lines.
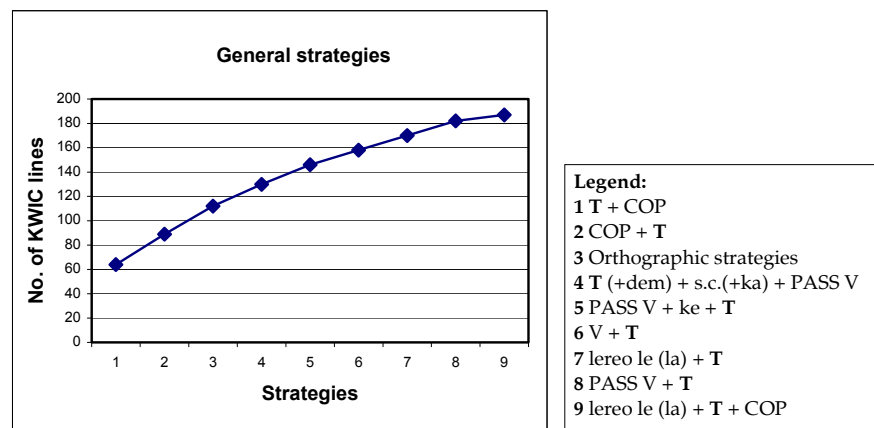
(5)   Definitional information marked by the pattern **Term + ke**

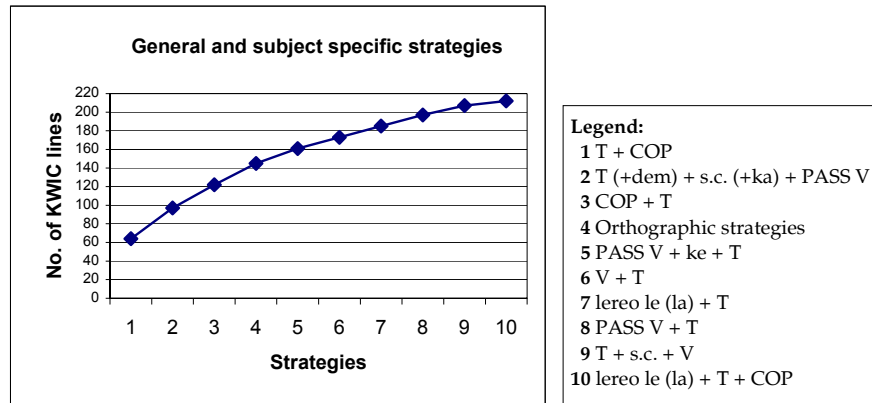|	Term	ke	
Mantšu a	**deiktiki**	**ke**	mantšu a a šupago motho yoo goba selo seo se bonalago mo seemotiko-logong sa kgokagano.
Deictic words *are* words that indicate a person or an object which is visible within the discourse context.			
mehuta ye e fapanego ya	**ditlaleletšo**	**ke**	gore mantšu ao a šomišwago mo dika-fokong go tlaleletša leina goba lediri
… different kinds of *complements is* that those words that are used in phrases to complement a noun or a verb …			
Malebišatheo a	**kgokagano**	**ke**	go fihliša molaetša wo o tšwago go mmoledi (goba mongwadi) go mmoledišwa (goba mmadi).
The basic function of *discourse is* to carry a message coming from the speaker (or writer) to the addressee (or reader).			
Re ka re	**lediri**	**ke**	kokwane ye e thekgilego mantšu a mangwe mo lefokong
We can say that a *verb is* the base which supports the other words in the sentence …			
Elelwa gore	**mafokotheo**	**ke**	mafoko a a nago le tiro e tee fela
Remember that *basic sentences are* sentences which have only one predicate …			

Mašala	**Mašala**	**ke**	mantšu a a šomago bjalo ka maina, gomme a ka bewa sebakeng sa maina mo lefokong:

Pronouns are words that function like nouns, and they can be placed in the place of a noun in a sentence:

kgethollwago mo lefokong.	**Sedirwa**	**ke**	motho yoo goba selo seo se angwago ke modiro wo o dirwago ke sediri.

… distinguished in a sentence. The *object is* the person or object which is influenced by the action carried out by the subject.

Ke ka fao ge re itše ra re	**tiro**	**ke**	motheo wa lefoko, ka ge tlhalošo ya yona e le kokwane ye e thekgilego ditlhalošo tša mantšu a mangwe mo lefokong.

Therefore we say that the *predicate is* the foundation of the sentence, since its meaning is the base which supports the meanings of the other words in the sentence.

yona ga e bonale. Re ka re	**tlhalošo**	**ke**	senaganwa goba boikakanyetšo bja mmoledi le mmoledišwa.

… is not visible. We can say that *meaning is* the idea or the intention of the speaker and the addressee.

le na le tlhalošo ya motheo.	**Tlhalošotheo**	**ke**	ke tlhalošo ye e rwelwego ke lentšu ge le eme le le nnoši.

… has a basic meaning. The *basic meaning is* the meaning which is carried by the word when it stands alone.

When the occurrence of the subject-specific patterns is added to those of the general strategies, the retrieval rate goes up from 189 concordance lines out of a possible 292, to 212, thus from 65% to 73%. These figures are represented in (6) and (7) respectively:

(6)   Retrieval rate of concordance lines using general strategies only

General strategies

No. of KWIC lines

Strategies

Legend:
1 **T** + COP
2 COP + **T**
3 Orthographic strategies
4 **T** (+dem) + s.c.(+ka) + PASS V
5 PASS V + ke + **T**
6 V + **T**
7 lereo le (la) + **T**
8 PASS V + **T**
9 lereo le (la) + **T** + COP

186   Elsabé Taljard

(7)   Retrieval rate of concordance lines using both general and subject-specific strategies

General and subject specific strategies

Legend:
 1 T + COP
 2 T (+dem) + s.c. (+ka) + PASS V
 3 COP + T
 4 Orthographic strategies
 5 PASS V + ke + T
 6 V + T
 7 lereo le (la) + T
 8 PASS V + T
 9 T + s.c. + V
10 lereo le (la) + T + COP

Although an increase of 8% seems almost negligible, it was found that the definitional information identified by means of these subject-specific strategies, provided the terminologist with highly relevant information, which would contribute much to the formulation of a good terminological definition of the particular term. Granted, from a statistical point of view, 73% retrieval does not seem to amount to much, but it has to be pointed out that there is a large amount of repetition to be found in the concordance lines, possibly due to the instructional nature of the text. Compare (8) below, in which five different KWIC lines provide definitional information on the term **mantšu** 'linguistic words', but the definitional information provided in these lines, is exactly the same, all of them basically indicating that sounds are combined to form words.

(8)   Definitional information on the term **mantšu** 'linguistic words'

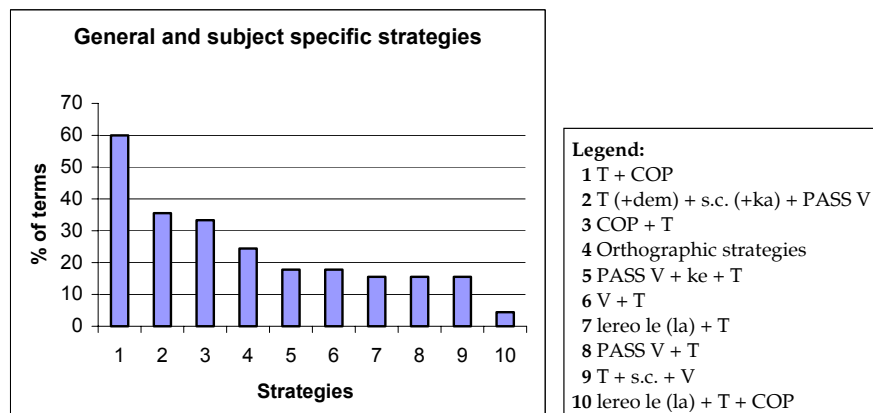|	**Term**	
gomme e re lemoša ge motheo wa	**mantšu**	e le medumo. Gape re lemoga ge mantšu a
… and it indicates to us that *the basis of words are sounds*. Furthermore we understand if words …		
medumo e tswalana go bopa	**mantšu**	mantšu a tswalana go bopa mafoko
… *sounds are combined to form words*, words are combined to form sentences …		
tswalanya medumo ye e itšego go bopa	**mantšu**	Ge a re mma, o tswalanya medumo
… *combine(s) certain sounds to form words*. If he/she says mma, he/she combines sounds …		
medumo e šomišwago ka gona go bopa	**mantšu**	Nepo ye bohlokwa ya fonolotši ke gona go
… *sounds which are combined to form words*. The important aim of phonology is to …		

Ge medumo e tswalanywa, go bopša	**mantšu**	Lekala la thutapolelo le le nyakišišago

When *sounds are combined, words are formed*. The section of grammar which investigates …

The second set of figures that is relevant to the table, concerns the relationship between a given strategy and the number of terms for which it provides definitional information. In (9), the retrieval rate of definitional information for each of the general strategies is given, whereas (10) reflects the retrieval rate for general and subject-specific strategies.

(9) Retrieval rate: general strategies (%)

General strategies only

Legend:
1 T + COP
2 COP + **T**
3 Orthographic strategies
4 T (+dem) + s.c. (+ka) + PASS V
5 PASS V + ke + **T**
6 V + **T**
7 lereo le (la) + **T**
8 PASS V + **T**
9 lereo le (la) + **T** + COP
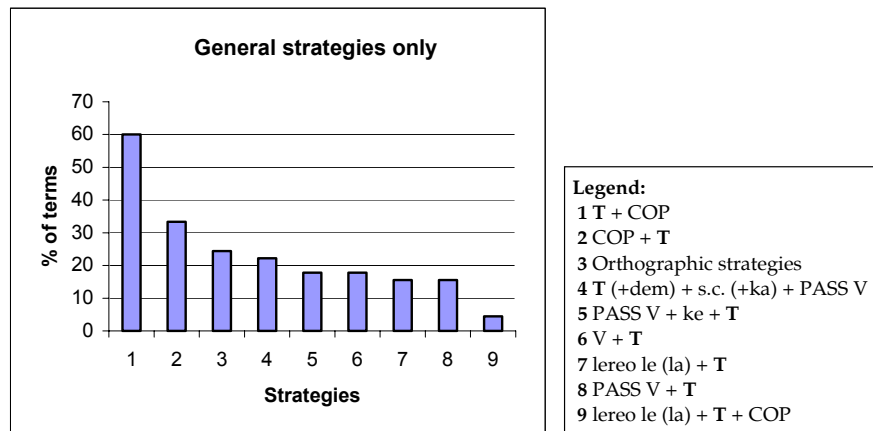
(10) Retrieval rate: general and subject-specific strategies (%)

General and subject specific strategies

Legend:
1 T + COP
2 T (+dem) + s.c. (+ka) + PASS V
3 COP + T
4 Orthographic strategies
5 PASS V + ke + T
6 V + T
7 lereo le (la) + T
8 PASS V + T
9 T + s.c. + V
10 lereo le (la) + T + COP

The strategy **Term + Identifying copula** provides information on 27 of the 45 terms; thus by employing this strategy as a search node in a KWIC line, one can retrieve information on 60% of the total number of terms. When the pattern **Identifying copulative + Term** is utilised, definitional information on a further 8 terms can be retrieved, pushing the retrieval rate up to 78%. Since these are both general strategies, one can expect that the use of these patterns as search nodes in WST's *Concord* tool would result in a high retrieval rate of definitional information, regardless of the subject field within which the terminologist may be working. This hypothesis will of course have to be tested by means of further research, making use of special field corpora based on a variety of special subject fields.

5.   Retrieval of information on conceptual relationships

It has already been pointed out that an analysis of a special field text not only provides general definitional information, but that it may also reveal the conceptual relationships in which terms appear. Information on these relationships is particularly important when the purpose of the retrieval of definitional information is writing of a terminological definition. An analysis of the definitional information which was retrieved from the corpus brought to light that conceptual information on quite a number of terms had in the process of definitional retrieval, also come to the fore. Although Sager (1990: 29) indicates that the "simplistic view of the past that concepts are adequately represented by three types of relationships (generic, partitive, other) has been generally abandoned", he nevertheless indicates that these relationships are still frequently used in terminology, together with a fourth category, labelled 'complex relationships'. Therefore, the conceptual relationships that were investigated for the purpose of this study are:

— Generic relationships, which deal mainly with the concepts 'superordinate' and 'subordinate' and which establish a hierarchical order of concepts;

— Partitive relationships, also called part-whole relationships, which indicate the connection between concepts consisting of more than one part and their constituent parts;

— Polyvalent relationships, which include polyhierarchical relationships, which occur when a concept is placed within more than one hierarchy within a given subject field; and

— Complex relationships, which include quite a large range of relationships such as causal relations, instrumental relations, production relations, functional relations, etc.

An analysis of the 292 KWIC lines that provided definitional information on the 45 test terms, revealed that the definitional information uncovered for these

terms also include information on the conceptual relationships between the particular term and other related terms. For 24 of the 45 terms, i.e. 53%, conceptual information could be retrieved from the text. The terms as well as the strategies that led to the recovery of the conceptual relationships have been marked by means of the symbol √ on the table in (4). A summary of the strategies and the number of terms for which conceptual information is provided by each strategy is given in (10):

(10)   Retrieval of conceptual information

Strategy	No. of terms
T + COP	13
COP + **T**	6
T (+dem) + s.c. (+ka) + PASS V	2
PASS V + **T**	5
PASS V + ke + **T**	2
Inchoative V + **T**	6

From the above, it is clear that the strategy **T** + COP which is responsible for the highest retrieval rate of definitional information, is also linked to the retrieval of conceptual information for the highest number of terms. Also noticeable is the fact that none of the subject-specific strategies seems to function as a marker of conceptual information. Some examples of conceptual information that were retrieved from the corpus appear in (11)–(14):

(11) Generic relationship

… go kgethollwa mehuta ye mebedi ye bohlokwa ya **kgokagano**, e lego (i) *kgokagano ka molomo*; le (ii) *kgokagano ka go ngwala*.
'… two important kinds of *discourse* are distinguished, which are (i) *spoken discourse*; and (ii) *written discourse*.
(**kgokagano** 'discourse' = superordinate; *kgokagano ka molomo* 'spoken discourse' and *kgokagano ka go ngwala* 'written discourse' = subordinates)

Ka fao *lekgokedi* le le bitšwa **lekgokasediri**.
'Therefore this *agreement morpheme* is called a ***subject concord**.'
(**lekgokedi** 'agreement morpheme' = superordinate; **lekgokasediri** 'subject concord' = subordinate)

(12) Partitive relationship

Dikarolo tše pedi tše bohlokwa tšeo di hlolago **lefoko** ke *sekafokoina* le *sekafokodiri*.
'The two important sections that form a ***sentence*** are the *noun phrase* and the *verb phrase*.'
(**lefoko** 'sentence' = whole; *sekafokoina* 'noun phrase' and *sekafokodiri* 'verb phrase' = parts)

Ngwana o thoma go tswalanya *medumo* ye e itšego go bopa **mantšu**.
'A child starts to combine certain *sounds* to form ***words***.'
(**mantšu** 'words' = whole; *medumo* 'sounds' = part)

(13) Polyvalent relationship

… lentšu la mathomo mo lefokong e tla ba **lehlathi**.
'… the first (linguistic) word in the sentence will be the ***adverb***.'

Tlaleletšatiro ya mohuta wo e bitšwa **lehlathi**.
'This kind of verbal adjunct is called an ***adverb***.'
(The term **lehlathi** is situated within two hierarchies: on the one hand it is a
    (linguistic) word, which — within the structuralist framework — implies
    that adverbs constitute a word class in Northern Sotho. Secondly, it is indi-
    cated that adverbs are also verbal adjuncts.)

(14) Complex relationships

Causal relationship
Togaganyo ka tlhološo e hlolwa ke **tswalano** ye e bonalago gare ga tlhološo ya
    mantšu a a fapanego.
'*Semantic cohesion* is caused/established by the ***relationship*** which exists be-
    tween the meanings of different words.'
(**tswalano** 'relationship' = cause; *togaganyo (ka tlhološo)* '(semantic) cohesion' =
    effect)

Instrumental relationship
Malebiša a … phetlekokgokagano … ke go nyakišiša le go hlatholla ka fao
    polelo e dirišwago ka gona go hlola **kgokagano** gare ga batho.
'The purpose of … discourse analysis … is to investigate and to explain how
    language is used to establish communication/discourse between people.'
(*polelo* 'language' = instrument; **kgokagano** 'communication' = process)

The same kind of investigation that was carried out in order to identify textual
markers of definitional information in corpora would have to be carried out
with regard to the identification of markers of conceptual information. It does
seem highly unlikely though that strategies for the marking of conceptual
information would differ radically from those used to mark definitional infor-
mation, since conceptual information is nothing but a specific kind of defini-
tional information.

6.   Using definitional information retrieved semi-automatically for the writing of a formal definition — a case study

Definitional information retrieved from a corpus can be used either as a rough
and ready tool by technical translators to assist them in decoding the text
which is to be translated, or it can be utilised for the writing of a formal termi-

nological definition. The actual content of the definition will naturally be determined by the usual factors that should be taken into account when any definition is compiled, i.e. target user, available space (in the case of paper dictionaries), etc. What follows is a brief illustration of how definitional information retrieved from the text can be used in the formulation of a terminological definition. The term selected for illustrative purposes is **lehlathi** 'adverb'.

In (15), the definitional information on **lehlathi** as culled from the text is cited.

(15) Definitional information on the term **lehlathi**

Ka ge boemotheo bja **lehlathi** e le mafelelong a lefoko, le swanetše go šutišwa gore le eme mathomong ge go nyakega gore le tšwelele ka tsela ya kgatelelo.
'Since the basic position of an *adverb* is at the end of a sentence, it must be moved to the beginning if it is to be emphasized.'

Lehlathi ke karolofoko yeo e hlathago ka moo modiro o phethagalago ka gona.
'An *adverb* is a sentence part which distinguishes/explains the way in which an action is carried out.'

Tlaleletšatiro ya mohuta wo e bitšwa **lehlathi** ka gobane lehlathi le hlatholla modiro wo o bolelwago ke tiro.
'A verbal adjunct of this kind is called an *adverb* since an adverb qualifies the action expressed by the predicate.'

Ge a nyaka go gatelela **lehlathi** (ke gore maabane), lentšu la mathomo mo mafokong e tla ba lehlathi.
'If he/she wants to emphasise the *adverb* (which is yesterday), the first word in the sentence will be the adverb.'

Based on the information provided in (15), the following terminological definition can be formulated:

(16) Formal definition of **lehlathi**

lehlathi: tlaleletšatiro yeo e hlathollago modiro wo o bolelwago ke tiro. Boemotheo bja lehlathi ke mafelelong a lefoko, eupša ge go nyakega gore le tšwelele ka tsela ya kgatelelo, le swanetše go šutišwa gore le eme mathomong. Mohlala: *maabane*.

'adverb: a verbal adjunct which qualifies the action expressed by the predicate. Its basic syntactic position is at the end of the sentence, but it can be moved to the beginning of the sentence for purposes of emphasis. Example: *maabane* 'yesterday'.'

It must again be stated very clearly that the process of retrieving definitional information semi-automatically does not make the involvement of the subject-

field expert redundant. There is no guarantee that this process will for example, succeed in retrieving the most salient features of a particular concept from the text, simply because they are not marked by any of the identified text markers, or because they are not mentioned in the text. A case in point is the definitional information retrieved for the term **leina** 'noun'. The following definitional information was retrieved for this term:

(17) Definitional information on the term **leina**

Sediri mo lefokong e ka ba **leina** go swana le Moithuti o a bala.
'The subject in a sentence may be a *noun* as in *The student reads/studies*.'

Mo sekafokoineng, lentšu le bohlokwa ke **leina** mola lentšu le bohlokwa mo
   sekafokodiring e le lediri.
'In a noun phrase, the important word is the *noun*, whereas the important word
   in the verb phrase is the verb.'

baithuti ke **leina** la Legoro 2
'*students* is a *noun* of class 2'

dipuku ke **leina** la Legoro 10
'*books* is a *noun* of class 10'

Leina le ka hlaolwa ke ditlaleletšo tše di fapanego.
'A *noun* can be determined by different adjuncts.'

… re lemoga ge lešala gagwe le laolwa ke **leina** le Makena.
… we understand that the pronoun *gagwe* is determined by the *noun* Makena.'

It is only a subject expert who will be able to point out that the most salient feature of nouns in Northern Sotho does not appear amongst the information retrieved, i.e. that nouns are morphologically characterised by a class prefix which is affixed to a stem. It is also not possible to deduce the correct super-ordinate, i.e. **lentšu** 'linguistic word' from the retrieved information. Although not foolproof, the process of semi-automatic retrieval of definitional informa-tion does however reduce the time that has to be spent on consultation with special field experts, which in turn, might make them more willing to partici-pate in terminology projects.

7.   Conclusion

The main aim of this article was to investigate whether authors provide defini-tional information for high frequency terms in technical texts, and if so, whether it would be possible to retrieve this information in a semi-automatic way. To this end, a case study was done for Northern Sotho, using a special field corpus on linguistics. It was found that definitional information on high frequency terms is indeed provided in these texts: definitional information could be identified for 45 of the 50 test terms, i.e. 90%. It was pointed out that

semi-automatic retrieval of this information could significantly reduce the dependency of the terminologist on the input of the special field expert. Consequently, a number of lexical and syntactic patterns that function as textual markers of definitional information were identified. When these patterns are entered as search nodes in WST's *Concord* tool, or for that matter in the *Search Box* of an ordinary *MSWord* document, the need to physically read through the textual material in search of definitional information is eliminated. The patterns **T + Identifying copula** and **Identifying copulative + T** distinguished themselves as highly effective markers of definitional information. By making use of these patterns as search nodes, definitional information could be retrieved for a surprising 78% of the test terms.

A close reading of the concordance lines containing definitional information revealed that important information regarding conceptual relationships existing between terms could also be retrieved. This information is particularly relevant in cases where the aim of information retrieval is the writing of a terminological definition.

Finally, it was stated that special field experts would always remain the final arbiters in evaluating the correctness and applicability of definitional information retrieved from texts. The value of the semi-automatic retrieval of definitional information is that the demands on the input of these experts are significantly reduced, since the software can isolate information, which could potentially be used in terminological definitions, leaving the expert the task of evaluating, and if necessary, supplementing the information proffered by the software.

Currently for Northern Sotho, terminologists do not have access to tagged and marked corpora, they do not have morphological parsers enabling them to retrieve all kinds of information from corpora with the proverbial pressing of a button, they do not have sophisticated software programmes, but they do have basic tools. This study hopefully serves to illustrate that by making innovative use of the basic tools that are available, one is indeed able to do terminological work of high quality.

Endnotes

1.   In this article, (South African) Bantu languages is used in accordance with its international application, i.e. as a purely linguistic term referrring to a specific family of languages.
2.   For more information on *WordSmith Tools*, see the home page of Mike Scott, the creator of the software: <http://www.lexically.net> (or else: <http://www. liv.ac.uk/~ms2928>).
3.   For a detailed discussion of this procedure, see Taljard and De Schryver (2002: 51-56).

Bibliography

Ahmad, Khurshid and Margaret Rogers. 2001. Corpus Linguistics and Terminology Extraction. Wright, Sue Ellen and Gerhard Budin (Eds.). 2001: 725-760.

Bowker, Lynne and Jennifer Pearson. 2002. *Working with Specialized Language: A Practical Guide to Using Corpora.* London: Routledge.

Pearson, Jennifer. 1998. *Terms in Context.* Amsterdam: John Benjamins.

Sager, Juan C. 1990. *A Practical Course in Terminology Processing.* Amsterdam: John Benjamins.

Sager, Juan C. 2001. Terminology Compilation: Consequences and Aspects of Automation. Wright, Sue Ellen and Gerhard Budin (Eds.). 2001: 761-771.

Shreve, Gregory M. 2001. Terminological Aspects of Text Production. Wright, Sue Ellen and Gerhard Budin (Eds.). 2001: 772-787.

Taljard, Elsabé and Gilles-Maurice de Schryver. 2002. Semi-Automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho. *Lexikos* 12: 44-74.

Wright, Sue Ellen and Gerhard Budin (Eds.). 2001. *Handbook of Terminology Management. Volume 2. Application-Oriented Terminology Management*. Amsterdam: John Benjamins.

Reflections on the Mediostructure in Special-Field Dictionaries. Also According to the Example of the *Dictionary for Lexicography and Dictionary Research* *

Herbert Ernst Wiegand, *Germanistisches Seminar, University of Heidelberg, Heidelberg, Germany (herbert.ernst.wiegand@gs.uni-heidelberg.de)*

> *One who runs alone cannot be outrun by another*
> (Ethiopian proverb)

Abstract: The central question posed in this contribution is: What constitutes the difference between a special-field dictionary article giving a cross-reference with a fixed textual reference position on the one hand, and a special-field dictionary article giving a cross-reference without a fixed reference position on the other, with regard to systematic cross-referencing? Because this very special type of question can only be answered systematically with reference to a theory of the mediostructure, some necessary terms are firstly introduced by means of illustrative analyses of special-field dictionary articles giving cross-references, such as e.g. *homogeneously externally orientated dictionary article*, *heterogeneously externally orientated dictionary article*, *item identifying the starting-point of the reference*, *item giving the reference*, *item giving a reference marker*, *mediostructural article network*, *the user's pre-knowledge relevant to the reference*, *position-bound reference*, *item-bound reference*, *single* and *expanded domain for the starting-point for the reference*, *mediostructural selection* and many more. On the basis of various analyses of different special-field dictionary articles with variable and fixed reference positions, a picture evolves which comes close to the situation where the question at the beginning of the contribution can be answered as follows: When the results of an appropriate mediostructural selection are presented by means of the application of the method of item-bound reference in special-field dictionaries with a variable reference position, and also when this happens by means of the application of the method of position-bound reference in special-field dictionaries with a fixed reference position, it is possible in principle that a systematic reference orientated to the dictionary functions can result. It can continuously be established, however, that in special-field dictionary articles with a fixed reference position, there is a higher degree of well-considered reference practices. This can be explained in so far that, in special-field dictionary articles with a fixed reference position, the article form already takes into account a genuine feature of special-field knowledge which consists of being systematic coherent knowledge, and therefore a

---

196   Herbert Ernst Wiegand

fixed reference position is provided, that means: not only defining and additional explanation is important in the formation of the article, but also the creation of connections which go beyond articles by means of references. Making references, therefore, is not a lexicographical action performed more or less concomitantly; rather, it obtains a relevant place because of being explicitly taken into account in the microstructural programme, and thus in the article form.

Keywords: SPECIAL-FIELD DICTIONARY ARTICLE GIVING A CROSS-REFERENCE; MEDIOSTRUCTURE; FIXED REFERENCE POSITION; VARIABLE REFERENCE POSITION; MEDIOSTRUCTURAL SELECTION; ITEM-BOUND REFERENCE; POSITION-BOUND REFERENCE

Abstract: Überlegungen zur Mediostruktur in Fachwörterbüchern. Auch am Beispiel des *Wörterbuchs zur Lexikographie und Wörterbuchforschung*. Die zentrale Fragestellung des Beitrages lautet: Worin besteht der Unterschied zwischen einem verweisvermittelnden Fachwörterbuchartikel mit festgelegter textueller Verweisposition einerseits und einem verweisvermittelnden Fachwörterbuchartikel ohne festgelegte Verweisposition andererseits im Hinblick auf eine systematische Verweisvermittlung? Da die sehr spezielle Fragestellung nur relativ zu einer Theorie der Mediostruktur systematisch bearbeitet werden kann, werden zunächst anhand exemplarischer Analysen von verweisvermittelnden Fachwörterbuchartikeln einige notwendige Termini eingeführt, wie z.B. *homogen außenorientierter Wörterbuchartikel*, *heterogen außenorientierter Wörterbuchartikel*, *Verweisausgangsangabe*, *Verweisangabe*, *Angabe mit Verweiskennzeichnung*, *mediostrukturelles Artikelnetz*, *verweisrelevantes Benutzervorwissen*, *positionsgebundene Verweisung*, *angabegebundene Verweisung*, *einfacher* und *erweiterter Verweisausgangsbereich*, *mediostrukturelle Selektion* und zahlreiche weitere. Anhand mehrerer Analysen unterschiedlicher Fachwörterbuchartikel mit variabler und fester Verweisposition ergibt sich ein Bild, das es nahelegt, dass die Ausgangsfrage des Beitrages wie folgt beantwortet werden kann: Sowohl wenn die Ergebnisse einer angemessenen mediostrukturellen Selektion durch eine Anwendung der Methode der angabegebundenen Verweisung in Fachwörterbuchartikeln mit variabler Verweisposition präsentiert werden, als auch wenn dies durch eine Anwendung der Methode der positionsgebundenen Verweisung in Fachwörterbuchartikeln mit fester Verweisposition geschieht, kann im Prinzip eine systematische, an den Wörterbuchfunktionen orientierte Verweisung erfolgen. Durchgehend ist jedoch festzustellen, dass in Fachwörterbuchartikeln mit festgelegter Verweisposition eine in höherem Grad reflektierte Verweisungspraxis angetroffen wird. Dies lässt sich so erklären, dass in Fachwörterbuchartikeln mit fester Verweisposition bereits in der Artikelform eine genuine Eigenschaft fachlichen Wissens, die darin besteht, systematisch zusammenhängendes Wissen zu sein, ausdrücklich dadurch berücksichtigt wird, dass eine feste Verweisposition angelegt ist, d.h.: Nicht nur Definieren und weiterführendes Erklären ist bei der Artikelgestaltung wichtig, sondern auch das Herstellen von artikelübergreifenden Zusammenhängen durch Verweisen. Verweisen ist damit kein lexikographisches Handeln, das mehr oder weniger nebenbei erfolgt; vielmehr erhält es dadurch einen relevanten Stellenwert, dass es im Mikrostrukturenprogramm und damit in der Artikelform ausdrücklich berücksichtigt wird.

Stichwörter: VERWEISVERMITTELNDER FACHWÖRTERBUCHARTIKEL, MEDIOSTRUKTUR, FESTE VERWEISPOSITION, VARIABLE VERWEISPOSITION, MEDIOSTRUKTURELLE SELEKTION, ANGABEGEBUNDENE VERWEISUNG, POSITIONSGEBUNDENE VERWEISUNG

1.   What it is about and what it is not about

Systematic analyses of the mediostructure of special-field dictionaries are rare in metalexicography. They will also not be given here. This contribution does *not* deal with an exhaustive presentation of the provision of reference data and furthermore *not* with a detailed description of the mediostructural programme of the *Dictionary for Lexicography and Dictionary Research* (DLDR); it also does *not* deal with a presentation of the abstract mediostructure of this special-field dictionary and its various partial structures (in the sense of Wiegand 2002: 222ff)[1]. On the contrary, the theme of my selected "reflections on the mediostructure in special-field dictionaries" results from the following question: Is there really a systematic difference between a special-field dictionary article giving a cross-reference **with** a fixed textual reference position on the one hand, and a special-field dictionary article giving a cross-reference **without** a fixed textual reference position on the other hand, from the perspective of giving cross-references in a systematic way? And if yes, what does it exist of? Even though these questions result from my experience with special-field lexicography, they are nevertheless very specific; they cannot be answered convincingly in an ad hoc manner by reaching for practical experiences, but only in relation to refined theoretical pre-set guidelines; these are put forward in my newer works on the theory of lexicographical text carriers, which consists of two partial theories, namely, a theory of dictionary structure and a theory of dictionary subject (cf. Wiegand 1998: 8ff). In what follows, I draw particularly on the following works in which above all central aspects of the first-mentioned partial theory are expressed: Wiegand 2000, 2001, 2001[2002], 2002, 2002a, 2002b, 2002c, 2003, 2003a. The notions introduced in these works cannot again be developed systematically here. Rather, the terms in the next section which are necessary for the argument by means of which the question can be answered, i.e. by contemplating some special-field dictionary articles which do not stem from the DLDR, are applied and only occasionally provided with exemplary explanations.[2]

2.   Special-field dictionary article giving cross-references: illustrative analyses for introducing terminology

Given here are two special-field dictionary articles giving secondary cross-references; dictionary article 1 (da_1) exhibits variable reference positions and a fixed reference position at the end of the article; accordingly da_1 belongs to the articles with variable and final reference position(s); da_2, on the other hand, exhibits only variable reference positions (cf. Figure 2-1).

da_1:

Essigsäure: Acidum aceticum glaciale, Azetsäure, Äthansäure, CH_3COOH. Wasserklare, stechend riechende, leicht brennbare, mit Luft explosive Gemische bildende Flüssigkeit, die unterhalb 16°C zu einer kristallinen Masse erstarrt (↑ Eisessig). E. kommt in zahlreichen Naturprodukten, z.B.

da_2:

grauer Markt
→ Markt, auf dem durch die Festsetzung eines → Mindestpreises, der über dem Gleichgewichtspreis (→ Preisbildung, →

198   Herbert Ernst Wiegand

äther, Ölen, in freier und gebundener Form vor. Salze und Ester werden als ↑ Azetate bezeichnet. Die E. hat in Form des ↑ Azetylkoenzym A im intermediärstoffwechsel große biochem. Bedeutung. Sie wird von zahlreichen Bakterien beim Abbau der Kohlenhydrate in den Vormägen und Dickdarmabschnitten der Pflanzerfresser gebildet, beim Rind in Mengen bis zu etwa 4 000 g pro Tag. Wird von den meisten Geweben der Pflanzenfresser verbrannt. Dient in der laktierenden Milchdrüse bes. zur Synthese der kurz- und mittelkettigen Fettsäuren. Bei der Fettbildung erscheinen etwa 33% der Energie der eingesetzten E. im Produkt. Reine E. wirkt stark ätzend auf Haut und Schleimhaut, resorptiv nephrotoxisch und hämolytisch. Etwa 20 ml wirken oral für den Menschen tödlich.   [6, 13]

Gleichgewicht) liegt, ein Angebotsüberschuß entsteht. Dieser induziert Geschäfte unterhalb des Mindestpreises, so daß ein bestimmtes Produkt zum offiziellen Mindestpreis, aber auch auf dem g. zu einem niedrigeren Preis gehandelt wird. Ein sinnentsprechendes Beispiel ist der Verkauf eines → Markenartikels auch als namenloses Produkt zu einem geringeren Preis.

Figure 2-1: da$_1$ from Wiesner and Ribbeck (1991) and da$_2$ from Woll (1992)

Special-field dictionary articles belong to the dictionary articles giving secondary cross-references when they exhibit at least one text segment giving a cross-reference. In the partially condensed article da$_1$, the following text segments giving cross-references are exhibited:

—   the habitual complete item giving a reference "↑Eisessig"

—   two habitual, morphosyntactically integrated reference markers, namely "↑Azetate" and "↑Azetylkoenzym A"

—   the homosegmental reduced item giving a reference "6, 13"

The latter numerical item giving a reference, orientated towards the outer text, stands in a fixed textual final reference position, which appears in all dictionary articles in Wiesner and Ribbeck (1991) as textual article positions. In the also partially condensed article da$_2$, the following text segments giving cross-references can be found:

—   three habitual, morphosyntactically integrated reference markers, namely, "→ Markt", "→ Mindestpreises" and "→ Markenartikels" (cf. under 3.)

—   two habitual reference markers in parentheses, namely, "→ Preisbildung" and "→ Gleichgewicht", with which the reference address is designated verbatim.

Special-field dictionary articles, which do not exhibit text segments giving cross-references, are not dictionary articles giving cross-references.

Both da$_1$ and da$_2$ are polydirectional dictionary articles: different external cross-reference addresses are designated in both. Whereas, however, only lemmatic external cross-reference addresses are designated in da$_2$, which consequently all lie in the same structure, namely, the macrostructural access structure, the numerical outer text external cross-reference addresses "6" and "13" are both designated in da$_1$ in addition to lemmatic external cross-reference addresses, and both lie within an external numerical outer text access structure belonging to the mediostructural access structures (in the sense of Wiegand 2003). The mediostructural orientation of both multiple externally orientated

special-field dictionary articles is therefore different: da$_2$ belongs to the homogeneously externally orientated dictionary articles, and da$_1$, in contrast, belongs to the heterogeneously externally orientated ones (cf. Wiegand 2003).

The special-field article giving a cross-reference da$_3$ (in Figure 2-2) is — as opposed to da$_1$ and da$_2$ — simply externally orientated, and in fact outer text orientated.

da$_3$:

> **Essig:** Verdünnte Essigsäurelösungen für Gewürz- und Konservierungszwecke. Der meiste E. wird durch Gärungsverfahren gewonnen. Nach den Rohstoffen bzw. der Gärungsmaische unterscheidet man Sprit- oder Branntweinessig, echten Weinessig, Weinessig 20% und 40%, Malzessig, Obstessig, Bieressig. Gärungsessig enthält noch Aromastoffe, die als Stoffwechselprodukte der Essigbildner aufgefaßt werden können. Die übl. Handelssorten enthalten 5 bis 10 g E.-Säure je 100 ml. E.-Essenz enthält 50 g E.-Säure in 100 g und wird durch Verdünnen von Syntheseessigsäure (Eisessig) hergestellt. Eine 80%ige E.-Essenz darf nur in Flaschen bestimmter Form und besonders deklariert gehandelt werden.   [13]

Figure 2-2: Dictionary article da$_3$ from Wiesner and Ribbeck (1991)

Because da$_1$ and da$_3$ are from the same special-field dictionary, the question arises why these special-field dictionary articles are not mediostructurally linked. This question concerning dictionary criticism, results from the fact that, whereas the reference prerequisites determined by the scope of the dictionary subject (in the sense of Wiegand 2002a: 172) are such that a text segment giving a cross-reference in the form of "↑Essigsäure" in da$_2$ can be expected, there are nevertheless dictionary-specific reasons why this expectation would be unjustified. These reasons which would be justifiable from a lexicographical-theoretical point of view, can only be explained with reference to the system of dictionary functions (in the sense of Wiegand 2001[2002]: 228ff) assigned to the dictionary. However, nothing is stated in the front matter texts of Wiesner and Ribbeck (1991) on this matter, and in fact also not in the informal language used in the lexicographical practice. In the accompanying metatext (the guidelines for the use of the dictionary), one finds the following guidelines on the mediostructure for the practice of referring under "3. Other" [!].

(Z 2-1: Wiesner and Ribbeck 1991)

> 3.1 The marker ↑ stands as guideline with the meaning that the special-field word is being described by the following keyword, e.g.
> Augenfliege ↑ Musca autumnalis
> or in the text with the meaning that the word following is explained as keyword, e.g. ... differentialdiagnostisch muß auch ↑ Tollwut beachtet werden.
> 3.2 At the end of a definition one or more numbers are presented, which refer to the author or authors (see list of authors).

In 3.1 in Z 2-1, text segments of pure reference articles are "explained" in the first instance, namely by means of the example of "Augenfliege ↑ Musca au-

tumnalis" and therefore by means of an incomplete reference article, because the last (right-aligned) text segment "14", which forms the fixed final reference position, is lacking. Although the "explanation" is intelligible, it is nevertheless obviously nonsensical in one respect. Then, of course, the "special-field word" *Augenfliege* is not described in the reference article, i.e. in the special-field dictionary article on *Musca autumnalis*, but specific features (of veterinary medical relevance) are described of a pastural fly which is called *Augenfliege* by farmers and of which the scientific name of the species is *Musca autumnalis*. The reference prerequisite stipulated by the dictionary subject is consequently the system-transcendental synonym relation of *Augenfliege* and *Musca autumnalis*. The same semantic relation is also the reference prerequisite stipulated by the dictionary subject in the following pure reference articles (cf. Figure 2-3).

da_4: **Ac. aceticum glaciale** $\uparrow$ Essigsäure   [14]

da_5: **Azetsäure** $\uparrow$ Essigsäure   [16]

da_6: **Äthansäure** $\uparrow$ Essigsäure   [16]

Figure 2-3: Pure reference articles da_4–da_6 from Wiesner and Ribbeck (1991)

The numbers "14" and "16" in the three reference articles da_4–da_6 as well as the numbers "6" and "13" in da_1 and da_3 deal with items of the numerical index of author names, which are at the same time outer text orientated items giving a reduced reference; they are outer text-orientated because, by means of them, a numerical outer text external cross-reference address is designated, which lies in the numerical external mediostructural outer text access structure ‹ 1. ‹ 2. ‹ 3. ‹ ... ‹ 72. ‹ 73. › (with "‹" standing for "*precedes*"). Each numerical outer text external cross-reference address opens, in the process of performing the external access actions, access to an outer text rubric of the outer text "list of authors". For example, "14" opens the external access to the following outer text rubric (= otr_1), which consequently functions as outer-textual reference goal rubric:

otr_1: 14. Prof. Dr. sc. med. vet. Regine Ribbeck, Leipzig
        (Arachno-Entomologie, Pharmakologie, Allgemeines)

The external access to otr_1 is, of course, only possible in the process of performing a procedure of following up a reference. A user who accesses from outside, can therefore only read in the list of authors; otr_1 is consequently an indirectly accessible outer text entry (in the sense of Wiegand 2003).
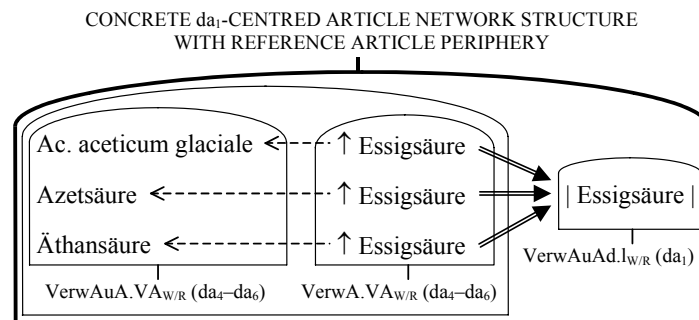
All reference articles in Wiesner and Ribbeck (1991) are heterogeneously externally orientated, because with them at least two external cross-reference addresses are designated at any given time, which lie in different outer access structures.

The three reference articles da_4–da_6 form the reference article periphery of a mediostructural article network internal to the word list, of which the centre of reference lies within da_1, so that the article network is centred towards da_1.

Wiesner and Ribbeck (1991) contains hundreds of da_n-centred article networks of the same article network type. All these article networks exhibit a concrete partial mediostructure which belongs to the same structural type. In order to give formally correct representations of the partially concrete mediostructures which belong to the article network in Wiesner and Ribbeck (1991), some classes of items giving cross-references and classes of reference addresses necessary for this purpose are introduced in such a way that their names (which function as class symbols) are listed and supplied with an explanation with regard to the contents, as follows:

— $VerwAuA_{W/R}$, the class of all items identifying the starting-point of the reference in Wiesner and Ribbeck (1991)

— $VerwAuA.VA_{W/R}$, the class of all items identifying the starting-point of the reference in reference articles in Wiesner and Ribbeck (1991)

— $VerwA_{W/R}$, the class of all items giving a reference in Wiesner and Ribbeck (1991)

— $VerwA.VA_{W/R}$, the class of all items giving a reference in reference articles in Wiesner and Ribbeck (1991)

— $VerwAuAd.l_{W/R}$, the class of all lemmatic (l) external cross-reference addresses in Wiesner and Ribbeck (1991)

— $VerwA.red.n_{W/R}$, the class of all numerical (n) reduced (red.) items giving a reference in Wiesner and Ribbeck (1991)

— $UmtAuAd.n_{W/R}$, the class of all numerical (n) outer text external cross-reference addresses in Wiesner and Ribbeck (1991)

The concrete da_1-centred article network structure with reference article periphery can now be represented in Figure 2-4.

Figure 2-4: Structural representation of a concrete partial mediostructure from Wiesner and Ribbeck (1991) with the name *da_1-centred article network structure with reference article periphery.* Notational conventions: "x--→y" means *x is article internally addressed to y*; "x⟹z" means *x is mediostructurally addressed to z*.

Apart from the word list-internal article networks, of which the centre of reference lie within a dictionary article, mediostructural article networks appear in Wiesner and Ribbeck (1991), of which the centre of reference forms an outer text rubric with numerical reference address. All dictionary articles of which the outer text orientated numerically reduced items giving a reference are equidirectional (so that by means of them the same numerical outer text external cross-reference addresses are designated in the fixed reference positions), belong to a mediostructural article network which exceeds the word list. In this way, for example, all dictionary articles, and in fact the reference articles as well as all other articles which exhibit a reduced item giving the reference "14" and therefore stem from one author, belong to such an article network.
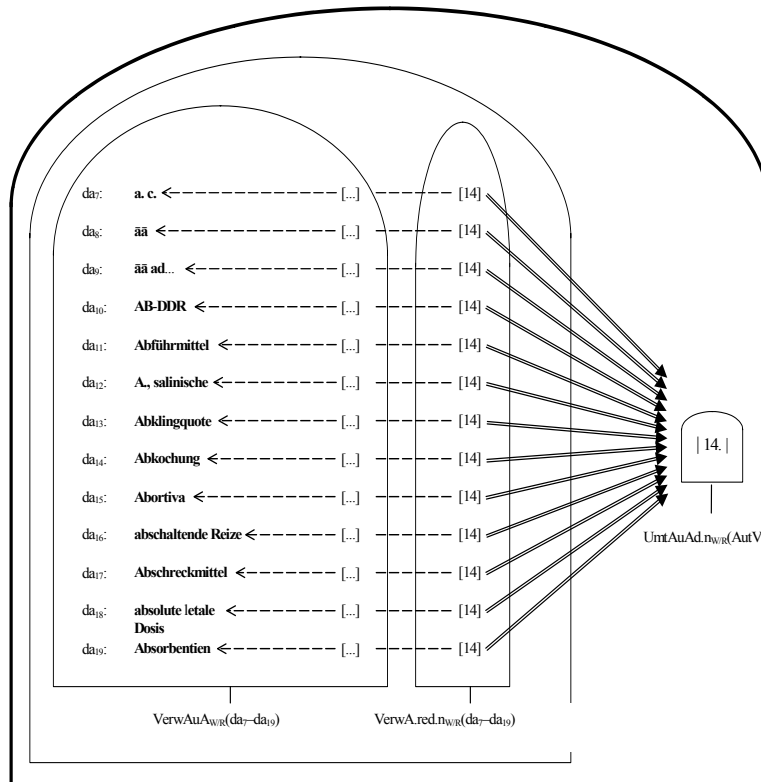
Figure 2-5: Structural representation of an excerpt from a concrete partial mediostructure from Wiesner and Ribbeck (1991) with the name *otr₁-centred article network structure with article periphery.* Notational conventions: "x--{...}-->y" means *x is article-internally non-adjacently addressed to y*; "x⟹z" means *x is mediostructurally addressed to z.* Abbreviation: AutV = list of authors

Figure 2-5 shows a small excerpt from the concrete partial mediostructure which belongs to the outer text rubric-centred article network exceeding the word list, of which the reference centre is the above-mentioned outer text rubric otr_1 with the numerical reference address "14". The excerpt only takes into account the non-functional article stretch, which ranges from the dictionary article with the lemma **A** to the article with the lemma **Abzeichen**; in this article stretch are thirteen articles (da_7–da_{19}) which exhibit a reduced item giving the reference "14".

The reference articles, which belong to the article periphery of an outer text-centred article network in Wiesner and Ribbeck (1991), are always also part of the reference article periphery of a da_n-centred article network.

After this exemplary insight into the functioning of special-field dictionary articles giving cross-references as well as into some of their text segments giving cross-references within the framework of mediostructural article networks, we also need some more precise knowledge about the differences between article-internal items giving references on the one hand, and article-internal items with reference markers on the other hand (cf. for the following Wiegand 2002a: 216 ff). We first look at some features and the functioning of items giving references; the terminological refining is withheld on an intermediate level.

Items giving a reference such as, for example, "↑Eisessig" in da_1 and "↑Essigsäure" in da_4–da_6, are monofunctional items giving cross-references of which the genuine purpose exclusively exists in the fact that a reference is made accessible. Items giving references are article constituents which stand in a particular textual reference position, in which, under certain circumstances, additional items giving references could appear (cf., for example, Figure 3-5). In the structural analysis of condensed or partially condensed dictionary articles, each textual reference position is in accordance with a microstructural reference position as excerpt from the concrete sequential order of the microstructure. Each item giving a reference, whether it is complete or reduced, is addressed twice. *Firstly*, it exhibits an article-internal relation address, which forms the item identifying the starting-point for the reference. *Secondly*, it is always mediostructurally addressed to a reference address; this reference address is designated with the item giving the reference address as a partial item of the item giving the reference (therefore in the reference position). In da_5, for example, the item giving the form of the lemma sign **"Azetsäure"** is the item identifying the starting-point for the reference and "Essigsäure" is the item giving the reference address; with the latter the guiding element ⌐Essigsäure⌐ of the lemma **Essigsäure** in da_1 is designated, so that it is possible for a user-in-action who has accessed the reference, to perform, knowing this guiding element, an external procedure of following up a reference, which would lead him/her to the designated address. That he/she should (or can) do this, can be inferred in the case of complete items giving the reference, such as, for example, "↑Essigsäure", by means of the habitual items giving the reference relation "↑". In da_5, the item giving the reference is adjacently addressed to the left, so that the address distance is zero and so that, by means of the item giving the reference,

together with the item identifying the starting-point for the reference, a single domain for the starting-point for the reference is identified. Accordingly, one talks of a reference position with a single domain for the starting-point for the reference. No pre-knowledge relevant to references is needed for a single domain for the starting-point for a reference. How one should understand this remark, is illustrated by means of a simple example (= E 2-1). A user-in-action — he/she is called B_1 — who has consulted and read, for example, the pure reference article da_5 with the search question of the type WHAT IS AZET-SÄURE?, cannot infer, by means of the item identifying the starting-point for the reference "Azetsäure", any knowledge which can be considered relevant for the solution whether or not to accept the references supplied by means of the item giving the reference "↑Essigsäure". Rather, if he/she wants to reach his/her original user's aim, B_1 takes this and opens up an obligatory reference leading straight to the article, getting lexicographical information of the contents and learning that he/she must perform an external procedure of following up a reference to reach the external cross-reference address ⎮Essigsäure⎮ which in turn opens up external access to the entire reference goal article with the lemma **Essigsäure** (that is to da_1), so that B_1 has the possibility to solve the search question by means of appropriate search question-specific selected lexicographical data in da_1, which then function as data for the reference goal. The knowledge gap which led to the use of the dictionary directly becomes the knowledge situation relevant to the reference for the user-in-action more or less simultaneously with the successful external access to the lemma **Azetsäure**, namely then, when the user-in-action establishes that he/she is being referred to another article. If one assumes that B_1 infers by means of the item giving the synonym "Azetsäure" in da_1 firstly that acetic acid (Azetsäure) is the same as Essigsäure (in German) and then by means of the following item giving the definiens which follows the formula "CH_3COOH" finds a solution to his/her search question, then the item giving the synonym and the item giving the definiens function as items giving the reference goal within the reference goal article. With this, it should be sufficiently clear what it means when one speaks about the fact that no pre-knowledge relevant to the reference is needed for a domain for the starting-point for a reference.

Next we look at an additional example (= E 2-2), which links up with (E 2-1). We assume that the interest of the user is aroused, on the basis of reading the item giving the definiens in da_1, for the fact that acetic acid solidifies below 16°C into a crystalline substance. For this reason, B_1 opens up, by means of the habitual item giving the reference "↑Eisessig", which directly follows the item giving the definiens in da_1, a reference going straight to the article, which has the result that he/she then performs a consultation procedure motivated by the reference (in the sense of Wiegand 1988: 412). The item giving the reference "↑Eisessig" is non-adjacently addressed to the left to the item giving the form of the lemma sign **"Essigsäure"** which functions as item identifying the starting-point for the reference. Between the item giving the reference and its non-adjacent reference address situated to the left, the following five elementary items

occur: the three items giving the synonyms "Acidum aceticum glaciale", "Azetsäure", and "Äthansäure"; the item giving the formula "CH_3COOH" and the item giving the definiens "Wasserklare [...] erstarrt". The distance in addressing which exists between the addressed item and its relation address, is calculated (in the sense of Wiegand 2002b: 144ff) according to the number of elementary items which stand between the item giving the reference "↑Eisessig" and **"Essigsäure"** (cf. also da$_{29}$); it consequently adds up to five, so that a five-fold expanded addressing relation exists between "↑Eisessig" and **"Essigsäure"**. An expanded domain for the starting-point for the reference thus belongs to the reference position in which "↑Eisessig" stands; this is in accordance with the domain of expansion of the addressing relation multiplied by the item identifying the starting-point for the reference. In clear contradistinction to a reference position with a single domain for the starting-point for the reference, a reference position with an expanded domain for the starting-point for the reference needs user pre-knowledge relevant to the reference. This is all the more extensive, as more propositional contents are opened up by means of the expanded domain for the starting-point for the reference. The user's pre-knowledge relevant to the reference is therefore precisely that propositional knowledge a user-in-action can have by cognitively working successfully through the items and/or item texts in the expanded domain for the starting-point for the reference, and which functions when determining whether a reference provided should be accepted or not. The user's pre-knowledge relevant to the reference should be looked at again.

Next we look at the second type of item giving a cross-reference, the item giving a marker for the reference. Reference markers are functional supplements to the item, therefore it is not possible to isolate them by means of the method of functional-positional segmentation. Items with reference markers are items giving cross-references which are always at least bi-functional. Their primary genuine function is determined by the class of items with the same general genuine purpose to which they belong. The function of cross-referring is never their primary function, but always a subordinate one (mostly secondary). For example, the first item in da$_2$ belongs to the items with reference markers; it is a bi-functional item giving the cross-reference and the definiens, namely, one with four reference markers. Its primary function lies in the fact that it has to present the definiens as part of the definition; its secondary function is the function of referring. Numerous classes of items with the same general genuine purpose have a subclass with corresponding items giving cross-references, for example:

— an item giving the paraphrase of the meaning/an item giving the cross-reference and the paraphrase of the meaning

— an item giving a competence example/an item giving a cross-reference and a competence example

— an item giving a phrase/an item giving a cross-reference and a phrase.

In da$_{20}$, "bei ↗ Licht b." is a condensed item giving a cross-reference and a phrase.

da$_{20}$:

> **bes<u>e</u>hen** (er besieht), besah, hat besehen *jmdn., etw.*
> *genau ansehen* (1.3): ich besah ihn (mir) aufmerk-
> sam, näher, von oben bis unten; sich den Schaden
> b.
> + bei ↗ Licht b.

Figure 2-6: Dictionary article da$_{20}$ from HWDG (1984)

In da$_{20}$, the item giving the cross-reference and the phrase is addressed to the item giving the form of the lemma sign **"bes<u>e</u>hen"**. This does not function, however — as in the case of items giving the reference — as an item identifying the starting-point for the reference. In da$_{20}$, the lexicographer does not refer the user-in-action from **"bes<u>e</u>hen"** to the external cross-reference address ⏐Licht⏐; rather, the user-in-action is referred from "bei ↗ Licht b." *to* ⏐Licht⏐ and therefore *to* the article on *Licht*. This means: the item identifying the starting-point for the reference is at the same time the condensed item giving the cross-reference and the phrase. The same is true of the item giving the cross-reference and the definiens in da$_2$. In principle, the following is true: each item with a reference marker functions at the same time as an item identifying the starting-point for the reference. An additional substantial difference with items giving a reference lies in the fact that the items with reference markers do not stand in their own textual reference positions. For example, "bei ↗ Licht b." in da$_{20}$ stands in the postcomment to the phraseology, which is clearly no reference position. Accordingly, the domain giving the starting-point for the reference in items with reference markers is linked to the particular item and not to the fixed article position; therefore one speaks of an item-bound (short: bound) domain giving the starting-point for the reference and of an item-bound reference. On the basis of the expandability of the addressing relation, items giving a reference are variable with regard to their position; that means one can fix reference positions in the microstructure programme totally independently from the textual positions of other items. Items giving a reference can therefore be inserted in a more versatile way than items with reference markers, where this flexibility is not possible. For example, in the same dictionary, an item giving the cross-reference and the paraphrase of the meaning can stand in the same textual position as an item giving the paraphrase of the meaning. Accordingly, also in da$_{21}$, the condensed item giving the phrase "vor jmdm., etw. auf der H. sein" stands in the same textual position as the item giving the cross-reference and the phrase in da$_{20}$.

da$_{21}$:

> **²Hut,** die; -, ⏐o. Pl.⏐; *in den Verbindungen*⏐ geh.
> in guter H. (*Obhut*) sein; jmdn., sich in guter H.
> wissen

+ vor jmdm., etw. auf der H. sein (*sich vor jmdm.,
etw. in acht nehmen)

Figure 2-7: Dictionary article da_{21} from HWDG (1984)

After the illustrative analyses of special-field dictionary articles giving cross-references undertaken so far in 2., in which not all mediostructurally relevant aspects were taken into account at all, we have at our disposal the necessary theoretical tools to try to answer the question posed in 1. about the systematic difference between special-field dictionary articles giving cross-references with a fixed textual reference position and those without a fixed reference position with regard to the systematic provision of references — now also taking into account the DLDR.

3.   Special-field dictionary articles with versus special-field dictionary articles without reference position

Special-field dictionary articles without a fixed reference position, and consequently those with one or more variable reference positions, are, for example, all the articles giving cross-references in Glück (2000) which are no pure reference articles; an example is da_{22} in Figure 3-1.

da_{22}:

Diastratisch (griech. διά (dia) ›durch‹, lat. strātum ›Decke‹, d.h. ›sich über verschiedene Schichten erstreckend‹) Auf verschiedene Sozialschichten gleichzeitig bezogene (ling.) Untersuchungsmethodik oder Sichtweise; Ggs. *synstratisch*. D. Forschungsfragen fallen per definitionem in den Aufgabenbereich der ↗ Soziolinguistik. Dazu gehören z.B. Untersuchungen der restringierten und ↗ elaborierten ↗ Kodes, der Verwendung von ↗ Dialekten und Standardvarietäten, von ↗ Bildungssprache, von ↗ Sprachbarrieren und Möglichkeiten ↗ kompensatorischer Erziehung; ↗ diachronisch, ↗ diaphasisch, ↗ diatopisch.   AM

Figure 3-1: Dictionary article da_{22} from Glück (2000)

As opposed to Wiesner and Ribbeck (1991), the item giving the abbreviation of the author names "AM" at the end of da_{22} is not an item giving a cross-refence, because it is not introduced in the accompanying metatext as a means of referring. Therefore "AM" also does not stand in a fixed reference position as, for example, "13" in da_3. The consultation act of a user-in-action who wants to know who hides behind the "AM" and therefore consults, in the front matter of Glück (2000), the list of authors, should for that reason not be understood — as

demonstrated in Wiegand (2002a: 198ff) in an argument with Kammerer and Lehr (1996) — as follows: A user infers, by means of "AM", an outer text-orientated reference, subsequently performing a procedure of following up a reference. "AM" therefore is not a potential item constituting a reference (in the sense of Kammerer and Lehr 1996). Rather, the user performs a habitual consultation procedure with a consultative search question (for example, questions such as: Who compiled the article? or: What does "AM" mean?). The motive for the user action is therefore not to open up a reference, but rather, by using a dictionary, to solve a (relatively harmless) text reception disturbance, for which elimination had been provided by the compilation of the list of authors.

In da_{22}, a partially condensed dictionary article, the item giving the antonym "*synstratisch*", is followed by an item text. This contains ten morphosyntactical integrated, habitual reference markers, from "↗ Soziolinguistik" to "↗ diatopisch". It should besides be mentioned that with morphosyntactical integrated, externally orientated reference markers, the external cross-reference addresses are often not designated literally, but instead with inflectional forms of those lemma signs which belong to the reference goal article. This is also the case in da_{22}, namely in the five cases combined in the following overview (cf. Figure 3-2).

Reference markers		Lemma-internal leading element as external cross-reference address	Lemma of the reference goal article
Reference relation markers	Reference address markers		
↗	elaborierten	|elaboriert|	**Elaboriert**
↗	Kodes	|Kode|	**Kode**
↗	Dialekten	|Dialekt|	**Dialekt**
↗	Standardvarietäten	|Standardvarietät|	**Standardvarietät**
↗	Sprachbarrieren	|Sprachbarriere|	**Sprachbarriere**

Figure 3-2: Morphosyntactical integrated reference markers, which designate inflectional forms

German users who master the practice of general dictionary use (in the sense of Wiegand 1998: 370), can probably infer the external cross-reference address from the designated inflectional forms without problems; for non-native speakers, it could already be different. The accumulation of morphosyntactical integrated reference markers in item texts, to which reference markers in parentheses are also added in many dictionary articles in Glück (2000), can by no means be considered user-friendly (cf. also Tarp 1999: 128). The practice of referring certainly appears to be really problematic when one questions its significance. If one asks about the "significance" of a part of the lexicographical practice (to which the accompanying practical results also belong, cf. Wiegand

1998: 34ff), then one asks about their specific usefulness and the textual organi-
sation of the user relation. This, however, always also means: one asks about
the types of situations of dictionary use to which those situations of use also
belong where a dictionary should be usable with success, and consequently
also about the types of users (e.g. laypersons, semi-experts, experts) who
belong to the circle of addressees. In other words, this means that one poses the
following question: Which dictionary functions (in the sense of Wiegand
2001[2002]) are supported by which results of lexicographical description prac-
tice?

In Glück (2000), there are only some very general and rather insignificant
explanations for user reference in the (reprinted) preface to the first edition,
which creates the impression that they came out of the economic pen of a pub-
lishing-house reader and not out of the head of a scientist who knows what
he/she writes and who has to be committed to the truth. There is no explana-
tion about which user groups of the circle of addressees were particularly fa-
voured and how the dictionary functions attached to the linguistic special-field
dictionary are structured hierarchically. If one assigns the dictionary function
of text reception (short: the reception function) as primary function to Glück
(2000), as it were as an experiment, then the presented section in da_{22} from the
mediostructural selection (as the selection from the set of all possible refer-
ences) can be considered completely meaningless. Because a user who comes
across a usage of *diastratisch* unknown to him/her in a linguistic text, in such a
way that he/she does not understand the text section, and, for example, con-
sults the linguistic special-field dictionary with the search question of the type
WHAT DOES THE TERM X MEAN?, does not need anything more than an
understandable, but nevertheless correct (according to the special-field) item
giving the definiens or an item text giving the definition by means of which the
meaning of the term can be inferred. If additional terms are found in the de-
fining text parts, the mediostructural selection has to be organised in such a
way that references are made to these terms. This does not already state how
this should best be done, namely, whether in a fixed reference position or by
means of items or item texts with reference markers or by means of a combina-
tion of both methods of referring. Corresponding text segments giving cross-
references are then article constitituents or parts of them, which strengthen the
reception function of the dictionary and supports the user in those instances of
dictionary use which belong to the situation type covered by the dictionary
function supporting the text reception (cf. Wiegand 2001[2002]: 238ff). In addi-
tion, if one assigns, as an experiment, the function of conveying specific special-
field information to Glück (2000), which belongs to the non-conflict related
dictionary functions, then, for example, the following argument could result:
Let us assume that a student (who can be categorised as a semi-expert) works
on a major assignment about the notion of "diasystem", and in the process of
data collection consults da_{22}. For him/her, it can be absolutely useful to accept
the references supplied in da_{22}. In order to strengthen the specific special-field
information in a special-field dictionary, it certainly is not necessary to write

item texts decorated with reference arrows such as the one in da$_{22}$. It namely suffices if there is a single sentence in a prominent place in the user's guidelines; this could, for example, be: "All linguistic terms used in dictionary articles are included as lemmata and are explained in the accompanying dictionary article." Such a sentence, which can also be understood as an expression of applying the method of omission of references, only explicitly emphasises what could be considered as the most important principle for making references in special-field lexicography for specific types of special-field dictionaries, to which particularly dictionaries for one specific field or part of a field, aimed at semi-experts, belong.

It is unnecessary and not useful to document the so-called permanent adherence to this principle with thousands of unnecessary reference arrows which obstruct reading. This practice in special-field lexicography, as one finds it, for example, in Glück (2000), also does not improve when it can be found in different variants in numerous special-field dictionaries. It looks like a contagious special-field lexicographical textual illness which is healed by means of a mostly arrow-formed virus from the family of mediostructural print viruses (which are related to certain electronic linking viruses), a virus which attaches itself to all items designating special-field terms when these special-field terms are also designated by means of a lemma.

Of course, the item-bound reference, i.e. the method of reference which leads to items with reference markers functioning as items identifying the starting-point for a reference, or which leads to item texts giving cross-references, can also be applied in a meaningful way. Whether such a reference without a fixed reference position can be considered meaningful or not, depends above all on three factors: (a) on the special-field scope of the dictionary subject, (b) on the type of dictionary, and (c) on the functions of the dictionary.

If one considers, for example, Wiesner and Ribbeck (1991) with regard to (a)–(c), the following argument arises: in Z 2-1 we read, for example, that the reference arrow "↑" in the dictionary text signifies that the word which follows it, has been explained as a keyword; with that, nothing wrong has been said, but also nothing very important, while it remains open in which cases the reference arrow had been used. The accompanying practice of reference is as follows: in the articles, the upward arrow is used only very occasionally; in addition, terms which have their own dictionary article, are relatively often used without there being a reference to these articles. That is, for example, the case in da$_{23}$ in Figure 3-3.

da$_{23}$:

Chalazion, n. ⟨griech.⟩ (Hagelkorn): Etwa erbsengroße, chronisch granulierende Entzündung mit Riesenzellausbildung einzelner Meibom' Talgdrüsen am Tarsus palpebrae, häuftig durch Hyperkeratose und Sekretstauung ausgelöst. In fortgeschrittenen Fällen bleibt eine Verunstaltung der Lider nicht aus, die dann zum chirurg. Eingriff zwingt.

[26,71]

Figure 3-3: Dictionary article da$_{23}$ from Wiesner and Ribbeck (1991)

For the following terms in da$_{23}$, there are independent dictionary articles to which this article does not refer: *Entzündung*, *Meibom' Drüsen*, *Tarsus*, *Palpebra* and *Hyperkeratose*. In addition, there is an article for *Riesenzelle*. One could have organised the mediostructural selection in such a way that an article such as the following in Figure 3-4 could have resulted.

Chalazion, n ‹griech.› (Hagelkorn): Etwa erbsengroße chronisch granulierende ↑Entzündung mit Riesenzell-ausbildung (↑Riesenzelle) einzelner Meibom' Talgdrüsen (↑Meibom' Drüsen) am Tarsus palpebrae (↑Tarsus, ↑Pal-pebra), haüfig durch ↑Hyperkeratose und Sekretstauung ausgelöst. In fortgeschrittenen Fällen bleibt eine Verun-staltung der Lider nicht aus, die dann zum chirurg. Ein-griff zwingt.   [HEW]

Figure 3-4: Dictionary article da'$_{23}$

The fact that this method was not used as in da'$_{23}$ (and thus not as in da$_{22}$), clearly shows that it was presupposed that the potential users of Wiesner and Ribbeck (1991) would be experts with academic knowledge, namely, above all veterinary surgeons and perhaps advanced semi-experts (for example, advanced students in veterinary medicine). This interpretation is also supported by the observation that in many dictionary articles various terms are often used in the items giving the definiens and in the encyclopedic item texts which are not included as a lemma in Wiesner and Ribbeck (1991). This is, for example, the case in da$_{24}$ in Figure 3-5.

da$_{24}$:

Cetaceum (Walrat): Aus den Kopfhöhlen vor allem von *Phy-seter macrocephalus* (Physeteridae) und anderer Wale ge-wonnenes festes Wachs, das hauptsächlich aus Palmitin-säure-Zetylester ($C_{15}H_{31}CO–O–C_{16}H_{33}$) besteht. Dient zur Veränderung der Dichte des Wales: Beim Tauchen wird die Dichte des Walrates erhöht, indem seine Durchblutung gedrosselt wird, so daß es erstarrt. Relativ selten verwen-det als ↑ Salbengrundlage.   [6, 14]

Figure 3-5: Dictionary article da$_{24}$ from Wiesner and Ribbeck (1991)

Physeter macrocephalus and *Palmitinsäure-Zetylester* are not included as lemmata. In order to appropriately interpret this element of lexicographical practice, one also has to consider the following, in my view (cf. (a)): The scope of the dictionary subject in Wiesner and Ribbeck (1991) is explicitly heterogeneous. In the selection of lemmata, the following fields of study, for example, were taken into account: history of veterinary medicine, animal protection, animal insurance, cytology, histology, endocrinology, disinfection, feeding principles, foot-and-mouth disease, internal veterinary medicine, veterinary hygiene, poultry diseases and about 80 additional fields. Not all of the terms stemming from

these fields and used in the articles can also be included and explained primarily, because that would cause an obligation to explain them once again in turn; rather, one should presuppose a knowledge of them or at least the knowledge where to find them in a reference work. According to the positioning of Wiesner and Ribbeck (1991) so far, it is consequently a comprehensive special-field dictionary specifically for experts. If one presupposes now that the primary dictionary functions are the function of reception and the function of conveying special-field information, then the following positive evaluation of the referring practice can be made: References are only made to articles when there is a good chance that a user-in-action with a genuine veterinary medical search question which emerges during text reception in a situation of study can broaden his/her veterinary medical knowledge by means of the data in these articles, if he/she accepts the references provided.

Even though therefore the application of the method of bound reference, which leads to items and item texts with reference markers, can result in a meaningful reference practice if the mediostructural selection of the parameters (a)–(c) are taken into account, is it nevertheless striking that in special-field dictionaries in which the method of position-bound reference resulted in fixed reference positions, one can mostly find a higher measure of well-considered reference practices. The articles da$_{25}$–da$_{28}$ exhibit, for example, a fixed reference position.

da$_{25}$:

ampicillin resistance *resistencia* f *a la ampicilina*
 If bacteria can survive and divide in the presence of ampicillin they are said to be ampicillin resistant or to show ampicillin resistance.
 = Apr
 ≠ ampicillin sensitivity
 → ampicillin; § 21

da$_{26}$:

colchicina f *colchicine*
 La colchicina es un alcaloide venenoso, que interfiere con las divisiones celulares. La colchicina se une a los microtúbulos e impide la formación de fibras del huso funcionales. Como resultado, los cromosomas no se separan y permanecen en metafase. Mediante el tratamiento con colchicina, de los extremos de las raíces u otras partes de las plantas en crecimiento, es posible, en algunos casos, producir plantas en las que se ha duplicado el número de cromosomas.
 • Quizás la primera prueba directa de una alteración específica de la membrana plasmática, en la resistencia a múltiples

da$_{27}$:

opholdstilladelse SUB **Aufenthaltsgenehmigung** ⟨F⟩
 ▲ *ansøge om* ~ Aufenthaltsgenehmigung beantragen; *foreløbig ~ på 12 måneder* vorläufige Aufenthaltsgenehmigung für 12 Monate (AKK); *forlænge en* ~ eine Aufenthaltsgenehmigung verlängern; *få ~ i et land* eine Aufenthaltsgenehmigung in einem Land (DAT) (ereilt) bekommen*
 ⇒ Aufenthaltsberechtigung, Aufenthaltserlaubnis

da$_{28}$:

svind SUB **Schwund** ⟨M; GEN SING -s; kun SING⟩
 ▲ ~ *i vægt* Gewichtsverlust; ~ *ved lækage er medforsikret* Schwund durch Leckage ist mitversichert; *der er* ~ *i beholdningen* der Vorrat schwindet*; *der må regnes med et* ~ *under lagringen* bei der Lagerung (DAT) ist mit (einem) Schwund zu rechnen; *et* ~ *i ladningen på 10%* ein Schwund der Ladung von 10% (DAT)
 ⇒ Gewichtseinbuße, Gewichtsverlust

medicamentos, la obtuvimos con nues-
tros estudios sobre las células de hamster
chino, que eran resistentes al medica-
mento colchicina.
→ cromosoma; meiosis; mitosis; aparato
del huso.

Figure 3-6: Dictionary articles da_{25} to da_{28} from Kaufmann and Bergenholtz (1998) and from Lauridsen et al. (1999)

In all four special-field dictionaries in Figure 3-6, the textual reference position stands at the end of the article, so that dictionary articles with final reference position are given. The right arrow is at times the habitual item giving the reference relation, which at the same time functions as the item giving the identification for the reference position. In da_{25}, the heterogeneously outer-orientated habitual item giving the reference "→ampicillin; §21" stands in the reference position, which designates with its first item giving the external cross-reference address "ampicillin", a lemmatic external cross-reference address, and with its second item giving the reference address "§21", an alpha-numerical outer text external address of which the place to find the reference address is the outer access structure of the systematic introduction "Introduction to Molecular Biology" in the front matter of Kaufmann and Bergenholtz (1998). In the reference position of da_{26}–da_{28}, there are at times homogeneously outer-orientated, habitual items giving a reference, because all designated reference addresses belong to the lemmatic external cross-reference addresses and at times lie in the same macrostructural access structure. All four articles in Figure 3-6 exhibit an expanded domain for the starting-point for the reference. With regard to the organisation of the reference position, articles such as those in Figure 3-6 were examples for the special-field dictionary articles in DLDR, on which we are now going to focus.

The DLDR is a polylingualised terminological special-field dictionary; on the level of the outer texts and in the definitorial section of the dictionary articles, it is bilingual (German and English). For the German lemmata, equivalents from ten languages are presented.[3]

The primary dictionary functions which are assigned to the DLDR are the reception function and the function of conveying special-field information. The secondary dictionary functions are the translation supporting dictionary function (short: translation function) and the production supporting dictionary function (short: text production function). The circle of addressees consists of semi-experts and experts. In what follows, only the primary functions are taken into account. The fixed reference position is located at the place in the dictionary article where lexicographical statements about the meaning of the term *Adressenentfernung* and those about special-field features of "distancing of address" are already dealt with, so that the definition position, the position for additional explanations and the position for an excerpt from the terminological network including the item giving the form of the lemma sign **"Adressenentfernung"** which functions as item identifying the starting-point for the ref-

erence, form the expanded domain for the starting-point for the reference. That knowledge which a user-in-action can acquire about the term *Adressenentfernung* and about distancing(s) of address, forms the user's pre-knowledge relevant to the reference. In what follows, the explanation is continued by means of an example.
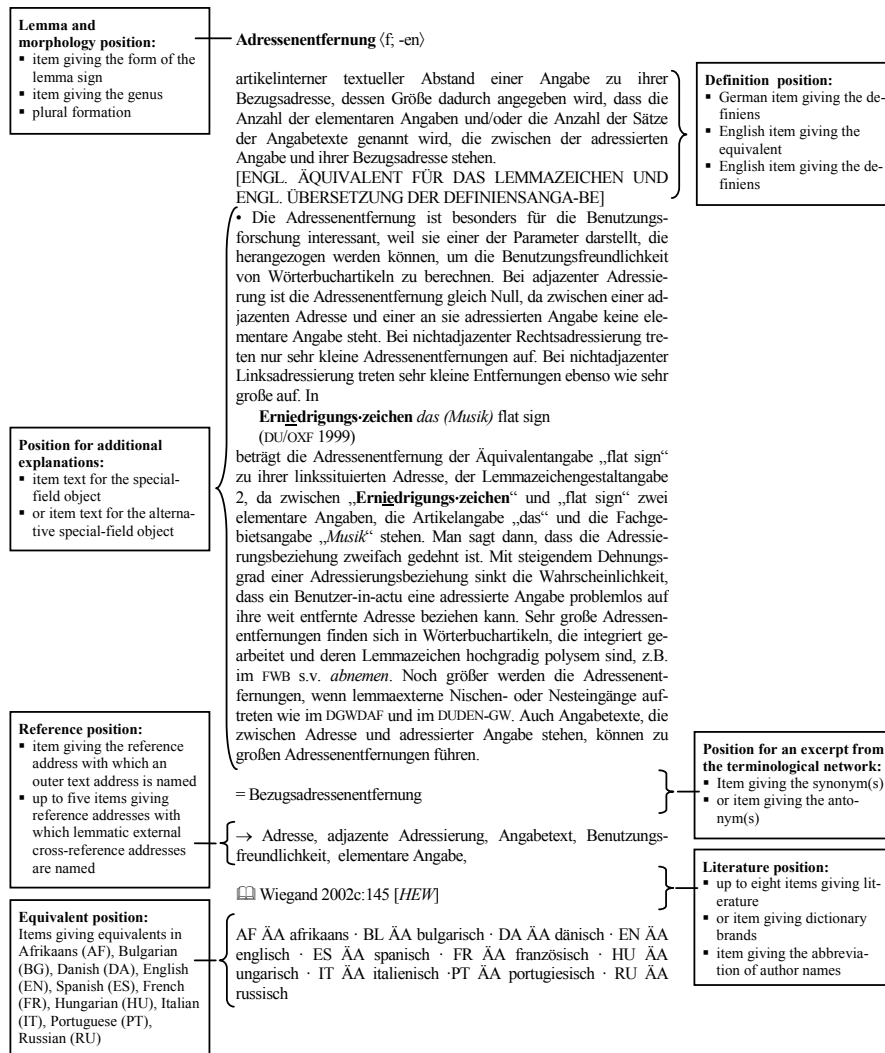

Figure 3-7: Illustration of da$_{29}$, a (middle-sized) DLDR-article, which deals with a reduced single article, since the article positions for the item text for the alternative special-field object and the antonym position are empty.

E 3-1: Our user-in-action, B_1, while reading a special-field text, comes across the term *Adressenentfernung*, which is unknown to him/her, so that it cannot be guaranteed that he/she will understand the metalexicographical special-field text. Therefore a text reception disturbance exists. If B_1 should turn to the DLDR, he/she would be in a conflict-conditioned consultation situation, which belongs to the type of situations of dictionary use covered by the text reception supporting function assigned to the DLDR. Accordingly, when B_1 has reached da_{29} after having made an external access procedure with the search question of the type WHAT DOES ONE UNDERSTAND UNDER ADRESSENENTFER-NUNG?, he/she has to be supported by the article text in such a way that the solution to the search question can be found. Depending on the way in which the knowledge situation particularly takes shape for the user-in-action, the required answer can already be found after reading the German item giving the definiens, which should be valid in E 3-1. In this case, one part of the user's pre-knowledge relevant to the reference, which belongs to the expanded domain for the starting-point for the reference, closes the knowledge gap for the user-in-action. Here, in the last instance, the question arises about the motive for designation and the coining of the term *user's pre-knowledge relevant to the reference*, because in example E 3-1, B_1 does not actually infer a reference by means of the item in the fixed reference position. In the rest of the argument, one should therefore take into account that this question of a terminological nature should be answered. With regard to both primary functions, namely, the reception supporting dictionary function and the function of conveying specific special-field related information, specifically that knowledge is rele-vant to the reference (from the viewpoint of a lexicographer of a terminological dictionary such as the type the DLDR represents) which tunes in with the user-in-action when reading the item giving the definiens. Because, understanding the item giving the definiens is, in the *first* place, the presupposition of the fact that all consultative search questions should be answered according to the meaning of the term, and in addition, also for the fact that, in the *second* place, the special-field explanations are being understood. The knowledge of the definition which comes into being when reading the item giving the definiens, is special-field encyclopedic object-constituting meaning knowledge and de-termines that the definiendum is something specific in the special field and not something unspecific or something else (cf. Wiegand 1988: 772ff). A special-field lexicographer of a terminological dictionary of the type of the DLDR should therefore set up the textual assumptions in such a way that the knowl-edge of the definition can be obtained by the user-in-action. Because of the fact that, in a worthwhile special-field dictionary for experts and semi-experts, terms related in a lexicographically-external manner to the term which forms the particular lemma sign, are also used in the items giving the definiens, there arises an obligation for reference related to the knowledge of the definition with regard to both dictionary functions, in case a user-in-action does not know a term which appears in the item giving the definiens. This obligation for refer-ence which exists particularly for the sake of semi-experts, can be fulfilled by

the lexicographer in various different ways, in that he/she could apply either one of the following methods of reference mentioned, or a combination of methods. The already mentioned methods of lexicographical reference are: the item-bound reference, the position-bound reference, and the omission of a reference.

In the application of the method of item-bound reference, the item giving the definiens in da$_{29}$, for example, would have been given as an item giving the cross-reference and the definiens. It could then, for example, have both the following forms:

(a)   article-internal textual distance of an ↗ item to its ↗ relation address, of which the value is given in such a manner that the number of ↗ elementary items and/or the number of sentences of ↗ item texts which stand between the ↗ addressed item and its relation address, are named.

(b)   article-internal textual distance of an *item* to its *relation address*, of which the value is given in such a manner that the number of *elementary items* and/or the number of sentences of the *items texts* which stand between the *addressed item* and its relation address, are named.

(a)   is an item giving the definiens with habitual, non-demarcated reference markers; the reference markers are integrated morphosyntactically.

(b)   is an item giving the definiens with typographical reference markers; here also, the reference markers are integrated morphosyntactically.

The method of position-bound reference is applied in da$_{29}$. In the reference position, reference is made to, amongst others, the articles to the *address* of the *terms* designated in the item giving the definiens (as superordinate of *Bezugsaddresse* — "relation address"), the *item text* and the *elementary item*.

In the application of the method of omission of reference, the following sentence would, for example, be found in the accompanying metatext: "All terms used in the position of the definition appear as headword and are explained in that particular article." In this case, the omission of reference is restricted to the position, and the user has to be informed separately in the metatext what the reference practice in the rest of the articles will look like. Item-bound reference and position-bound reference can be combined. Hartmann and James (1998) worked with this combination of methods. The application of the method of omission of reference does not, of course, automatically have the result that every term becomes a potential item constituting a reference (in the sense of Kammerer and Lehr 1996). Rather, with this method, it should be taken into account that each term used in items and item texts can potentially create conflict in the sense that it could cause a disturbance in the text reception during the process of dictionary use.

The user's pre-knowledge relevant to the reference, which belongs to the expanded domain for the starting-point for the reference, is not only relevant from the viewpoint of the lexicographer, but also from the perspective of the

user-in-action. Whoever has understood the entire explanatory dictionary text after reading da[29], is faced with an offering of references in the closing section of the article. The user can ignore them or (partially) accept them: the appropriate decision is, however, determined (if not exclusively) by the user's pre-knowledge relevant to the reference. Particularly in the context of non-communicative usage (in the sense of Wiegand 1998: 531), items giving a reference can be used by means of which expletive or additional references can be opened up. Such items giving references strengthen the specific subject-related information function of the dictionary. In da[29], it is particularly the item giving the lemmatic external cross-reference address "Benutzerfreundlichkeit" that provides for the opening up of additional references.

In order to take into account another aspect, a section of a second article from DLDR is presented.

Angabe der semantischen Lemmatisierung

⟨f; -n⟩

Angabe, anhand derer erkannt werden kann, wie der semantische Teil der Lemmatisierungsaufgabe gelöst wurde.

[ENGL.   ÄQUIVALENTANGABE,   ENGL. DEFINIENSANGABE]

• Die Angabe der semantischen Lemmatisierung ist entweder als Polysemieangabe oder als Homonymieangabe realisiert. Hierbei können die semiotischen Details der Realisierung verschieden sein. – Symbol für die Angabeklasse: A.semL

→ § 14, Angabe, Homonymieangabe, Polysemieangabe, semantische Lemmatisierung
📖 Reichmann 1989: 72f; Wiegand 1991: 398, 419, 2003; Zöfgen 1998   [*HEW*]
[ÄQUIVALENTPOSITION]

Figure 3-8: da[30] from DLDR

In the reference position of da[30], five items giving the reference address follow the habitual item giving the reference relation "→". With the first, namely, "§14", an alphanumerical outer text reference external address is designated which lies in the external access structure of the integrated outer text "Systematic Introduction" and which directly opens up access to §14 "Die Bauteile von Wörterbuchartikeln" ("The components of dictionary articles"). This reference particularly strengthens the specific subject-related information function of the dictionary. Whoever accepts the offering of references as a semi-expert, has a good chance to get a quick overview of the other components of the dictionary. Because an outer text reference position is given at the end of §14, from which reference is made to articles, and in fact especially to synoptic articles (e.g. the

articles on *Angabe*, *Mikrostrukturanzeiger*, *lexikographische Bearbeitungseinheit*), the user can purposefully continue the action of reference on the basis of questions which come up. The article-heterogeneous list of words of the DLDR is consequently mediostructurally linked to the Systematic Introduction by means of the method of bi-directional reference. In this way, for example, all synoptic and single articles in which reference position the item giving the outer text external address §14 appears, form the article periphery of an article network, of which the centre is §14.

In the reference position of da_{30}, four alphabetically ordered items of lemmatic outer addresses follow on "§14". These were conceptualised as follows: reference is made to the synoptic articles in which the superordinate to the lemma sign **"Angabe der semantischen Lemmatisierung"**, namely, *Angabe*, has been treated lexicographically. Because *Angabe der semantischen Lemmatisierung* is itself a generic term, reference is made to both the hyponym terms *Homonymieangabe* and *Polysemieangabe*. In order to support the compilation of the definition knowledge, reference is finally made to the article with the lemma **semantische Lemmatisierung**. With the four items of lemmatic outer addresses, a small coherent part of the network is consequently "cut" out of the terminological-semantic network to which the terms belong.

Here we break off the illustrative explanations and try to answer the question at the beginning in the summary in the next section.

4.   Summary

Subject-field knowledge such as, for example, that from lexicography and dictionary research, does not consist of isolated elements of knowledge. Rather, it is something which is in many ways connected with regard to its contents. Whoever compiles an alphabetical printed dictionary is forced to distribute the data to be presented along the provided text compound constituents. The close connection with regard to the contents can at the same time only be taken into account in a very restricted way. In order to uncover the connection, which is necessarily concealed by the data distribution along the different access positions, a mediostructural network of the lexicographically distributed data has to ensue. The networking is inadequate if everything which belongs together extra-lexicographically is also linked together by means of mediostructural addressing. Rather, a mediostructural selection should result which depends decisively on the primary dictionary functions and consequently on the intended user relation. The question which was formulated in section 1. deals with whether the results of the mediostructural selection (in as far as dictionary articles are concerned when presented by means of the application of the method of bound reference so that special-field dictionary articles without a fixed reference position occur) are present in such a way that these special-field dictionary articles can be distinguished from those (with regard to a systematic

way of giving cross-references) in which the results of the mediostructural selection are presented according to the method of position-bound reference.

Firstly, it has to be established that both types of special-field dictionary articles giving cross-references allow for appropriate articles with systematically conceptualised reference. Apart from the many smaller differences which we have already come across and which should not be repeated here, a substantial systematic difference exists, in my view, in the fact that in articles giving cross-references with a fixed reference position, a genuine feature of the subject-knowledge, namely, coherent knowledge carried over by means of a theory, should be explicitly showed and brought to the fore in such a way that it can be seen clearly in the article structure as part of the dictionary structure! Not only the definiens position, the position for additional explanation, and additional positions are provided according to the microstructural programme, but also a reference position. This is an obligation to a theory-driven approach and the systematic method of conveying coherent subject-field knowledge. The fixed reference position enforces, to a great extent, a systematically conceptualised reference, which can at the same time be considered a successful immunisation against the "arrow virus".

Concordance

external cross-reference address	Verweisaussenadresse	(VerwAuAd)
item giving the reference	Verweisangabe	(VerwA)
item identifying the starting-point of the reference	Verweisausgangsangabe	(VerwAuA)
mediostructure	Mediostruktur	
outer text external cross-reference address	Umtextaussenadresse	(UmtAuAd)
reference article	Verweisartikel	(VA)
special-field dictionary article giving a cross-reference	verweisvermittelnder Fachwörterbuchartikel	
reference marker	Verweiskennzeichnung	

Endnotes

1. For the DLDR cf.: <www.fabulex.de>; furthermore: Kammerer (2001[2002]); Wiegand (2002, 2003a); also compare footnote 3.
2. I shall also only occasionally refer to the mentioned works, and not in the case of every term.
3. The equivalents are researched by the following collaborators: Afrikaans: Rufus H. Gouws (Stellenbosch); Bulgarian: Pavel Petkov (Sofia); Danish: Jens-Erik Mogensen (Copenhagen); English: Rufus H. Gouws; French: Franz Joseph Hausmann (Erlangen); Italian: Giovanni

Rovere (Heidelberg); Portuguese: Stefan Schierholz (Göttingen); Russian: Dimitri Dobrovol'skij (Moscow); Spanish: Reinhold Werner (Augsburg); Hungarian: Regina Hessky (Budapest). — The editors of the DLDR are: Michael Beißwenger (Dortmund), Rufus H. Gouws (Stellenbosch), Matthias Kammerer (Karlsruhe), Angelika Storrer (Dortmund), Herbert Ernst Wiegand (Heidelberg), and Werner Wolski (Paderborn). — The article structures of the special-field dictionary articles are described in Kammerer 2001[2002]: 263ff, and in more detail in Wiegand (2003a); a general characterisation can be found in Wiegand (2002) with an overview of the entire text compound and a view into the systematic introduction. — The DLDR will be printed in two volumes of about 6 000 lemmata including the reference lemmata; it is being written on the internet. At the moment, about 1 400 articles are on the internet (protected by password), although without English translation of the item giving the definiens and without the equivalents. These parts will later be added automatically. — The editorial staff of the DLDR who had been in Heidelberg until June 2004, has since July 2004 been located in Dortmund.

Bibliography

Brdar-Szabé, Rita and Elisabeth Knipf-Komlósi (Eds.). 2004. *Lexikalische Semantik, Phraseologie und Lexikographie: Abgründe und Brücken. Festgabe für Regina Hessky*. Duisburger Arbeiten zur Sprach- und Kulturwissenschaft 57. Frankfurt am Main/Bern/New York/Paris: Peter Lang.

Glück, Helmut (Ed.). 2000. *Metzler Lexikon Sprache 2*. Revised and expanded edition with 70 illustrations, 17 of which maps. Stuttgart: Verlag J.B. Metzler.

Hartmann, R.R.K. and Gregory James. 1998. *Dictionary of Lexicography*. London/New York: Routledge.

HWDG = **Kempcke, Günter et al. (Eds.).** 1984. *Handwörterbuch der deutschen Gegenwartssprache*. In two volumes A–K and L–Z. Berlin: Akademie-Verlag.

Kammerer, Matthias. 2001[2002]. XML-getaggte Wörterbuchartikel. Ein Bericht aus der Praxis des Wörterbuchs zur Lexikographie und Wörterbuchforschung. *Lexicographica* 17: 249-301.

Kammerer, Matthias and Andrea Lehr. 1996. Potentielle Verweise und die Wahrscheinlichkeit ihrer Konstituierung. Wiegand, Herbert Ernst (Ed.). 1996. *Wörterbücher in der Diskussion II. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 311-354. Lexicographica. Series Maior 76. Tübingen: Niemeyer.

Kaufmann, Uwe and Henning Bergenholtz. 1998. *Encyclopedic Dictionary of Gene Technology/Diccionario Encyclopédico Ingeniaría Genética. Vol. I: English (With Spanish Equivalents); Vol. II: Español (Con equivalentes en inglés)*. Toronto: Lugus Libros LatinAmerica Inc.

Lauridsen, Ole, Sven-Olaf Poulsen, Tore Gadegård and Sven Tarp. 1999. *Dansk–tysk erhvervsordbog*. Viborg: Forlaget Systime.

Tarp, Sven. 1999. Theoretical Foundations of the So-called Cross-reference Structures. *Lexicographica* 15: 114-137.

Wiegand, Herbert Ernst. 1988. Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachlichen und enzyklopädischen Wissen. Munske, Horst Haider, Peter von Polenz, Oskar Reichmann and Reiner Hildebrandt (Eds.). 1988. *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern*: 729-790. New York: Walter de Gruyter.

Wiegand, Herbert Ernst. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband.* With 159 illustrations in the text. Berlin/New York: Walter de Gruyter.

Wiegand, Herbert Ernst. 2000. Wissen, Wissensrepräsentation und Printwörterbücher. Heid, Ulrich, Stefan Evert, Egbert Lehmann and Christian Rohrer (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress. EURALEX 2000. Stuttgart, Germany, August 8–12, 2000*: 15-38. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Wiegand, Herbert Ernst. 2001. Sprachkontaktwörterbücher. Typen, Funktionen, Strukturen. Igla, Birgit, Pavel Petkov and Herbert Ernst Wiegand (Eds.). 2001. *Theoretische und praktische Probleme 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/Zurich/New York: Georg Olms.

Wiegand, Herbert Ernst. 2001[2002]. Was eigentlich sind Wörterbuchfunktionen? Kritische Anmerkungen zur neueren und neuesten Wörterbuchforschung. *Lexicographica* 17: 217-248.

Wiegand, Herbert Ernst. 2002. Zur Makrostruktur und zu den äußeren Zugriffsstrukturen im *de Gruyter Wörterbuch Deutsch als Fremdsprache*. Wiegand, Herbert Ernst (Ed.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des* de Gruyter Wörterbuch Deutsch als Fremdsprache: 413-442. Lexicographica. Series Maior 110. Tübingen: Max Niemeyer Verlag.

Wiegand, Herbert Ernst. 2002a. Altes und Neues zur Mediostruktur in Printwörterbüchern. *Lexicographica* 18: 168-252.

Wiegand, Herbert Ernst. 2002b. Adressierung in zweisprachigen Printwörterbüchern. Wiegand, Herbert Ernst (Ed.). 2002. *Studien zur zweisprachigen Lexikographie mit Deutsch VIII*: 111-175. *Germanistische Linguistik* 166. Hildesheim/Zurich/New York: Georg Olms Verlag.

Wiegand, Herbert Ernst. 2002c. Über textuelle Strukturen der Wörterbuchartikel und Artikelnischen im *de Gruyter Wörterbuch Deutsch als Fremdsprache*. Zugleich ein Beitrag zur Weiterentwicklung einer Theorie der Wörterbuchform. Wiegand, Herbert Ernst (Ed.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des* de Gruyter Wörterbuch Deutsch als Fremdsprache: 497-595. Lexicographica. Series Maior 110. Tübingen: Max Niemeyer Verlag.

Wiegand, Herbert Ernst. 2003. Überlegungen zur Typologie von Wörterbuchartikeln in Printwörterbüchern. Ein Beitrag zur Theorie der Wörterbuchform. *Lexicographica* 19 [2004]: 169-313.

Wiegand, Herbert Ernst. 2003a. Wörterbuch zur Lexikographie und Wörterbuchforschung/Dictionary of Lexicography and Dictionary Research. Städtler, Thomas (Ed.). 2003. *Wissenschaftliche Lexikographie im deutschsprachigen Raum. Im Auftrag der Heidelberger Akademie der Wissenschaften*: 417-437. Heidelberg: Winter.

Wiesner, Ekkehard, and Regine Ribbeck (Eds.). 1991. *Wörterbuch der Veterinärmedizin*. Compiled by 73 scientists. 2 volumes. Third revised edition. Jena/Stuttgart: Gustav Fischer Verlag.

Woll, Artur (Ed.). 1992. *Wirtschaftslexikon*. Sixth revised and expanded edition. Munich/Vienna: R. Oldenbourg Verlag.

Basic Problems of Learner's Lexicography [*]

Sven Tarp, *Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark (st@asb.dk)*

Abstract: In general, learner's dictionaries suffer from the fact that lexicography has been treated as a subdiscipline of linguistics and not as an independent discipline and that, correspondingly, no serious efforts have been made to develop a general theory of learner's dictionaries within the general theory of lexicography. Such a theory must undoubtedly be developed on the basis of the theory of lexicographic functions. Foreign language learning is a complex process and the very concept of a learner is in no way unambiguous. Dictionaries can be conceived to assist the learner in different aspects of the language-learning process. Accordingly, at least four main functions of a learner's dictionary can be distinguished corresponding to the following user situations: text reception in the foreign language, text production in the foreign language, conscious study of the foreign-language lexis and conscious study of the foreign-language grammar. Learner's dictionaries are, thus, not to be considered a special type of dictionaries, but constitute a more general category including various types of learner's dictionaries according to their functions or combination of functions. The great challenge to learner's lexicography is to conceive and compile dictionaries that assist the learner in as many aspects of the language-learning process as possible.

Keywords: LEXICOGRAPHY, LEARNER'S LEXICOGRAPHY, LEARNER'S DICTIONARIES, GENERAL THEORY OF LEARNER'S LEXICOGRAPHY, LEXICOGRAPHIC FUNCTIONS, COMMUNICATION-ORIENTATED FUNCTIONS, KNOWLEDGE-ORIENTATED FUNCTIONS, FOREIGN-LANGUAGE TEXT RECEPTION, FOREIGN-LANGUAGE TEXT PRODUCTION, ASSIMILATIONS OF THE FOREIGN-LANGUAGE SYSTEM, LEARNERS' CHARACTERISTICS, USER NEEDS, USER SITUATIONS, LEXICOGRAPHIC DATA, LEXICOGRAPHIC STRUCTURES

Opsomming: Basiese probleme van die aanleerdersleksikografie. Oor die algemeen ly aanleerderswoordeboeke onder die feit dat die leksikografie as 'n subdissipline van die linguistiek behandel is en nie as 'n onafhanklike dissipline nie en dat, ooreenkomstig, geen ernstige pogings aangewend is om 'n algemene teorie van aanleerderswoordeboeke binne die algemene teorie van die leksikografie te ontwikkel nie. So 'n teorie moet ongetwyfeld ontwikkel word met die teorie van leksikografiese funksies as basis. Die aanleer van 'n vreemde taal is 'n komplekse proses en selfs die konsep van 'n aanleerder is geensins duidelik nie. Woordeboeke kan bedink word om die aanleerder met verskillende aspekte van die taalaanleerproses te help. Gevolglik kan ten minste vier hooffunksies van 'n aanleerderswoordeboek, wat ooreenstem met die volgende gebruikersituasies, onderskei word: teksresepsie in die vreemde taal, teksproduksie in die vreemde taal, bewuste studie van die leksis van die vreemde taal en bewuste studie van die grammatika van die vreemde taal. Aanleerderswoordeboeke moet dus nie as 'n besondere soort woordeboeke

---

beskou word nie, maar vorm 'n meer algemene kategorie wat verskillende soorte aanleerders-woordeboeke volgens hul funksies of kombinasie van funksies insluit. Die groot uitdaging vir die aanleerdersleksikografie is om woordeboeke te bedink en te maak wat die aanleerder help met soveel aspekte van die taalaanleerproses as moontlik.

Sleutelwoorde: LEKSIKOGRAFIE, AANLEERDERSLEKSIKOGRAFIE, AANLEERDERS-WOORDEBOEKE, ALGEMENE TEORIE VAN AANLEERDERSLEKSIKOGRAFIE, LEKSIKOGRA-FIESE FUNKSIES, KOMMUNIKASIEGEORIËNTEERDE FUNKSIES, KENNISGEORIËNTEERDE FUNKSIES, VREEMDETAALTEKSRESEPSIE, VREEMDETAALTEKSPRODUKSIE, ASSIMILASIE VAN DIE VREEMDETAALSISTEEEM, AANLEERDERSEIENSKAPPE, GEBRUIKERSBEHOEF-TES, GEBRUIKERSITUASIES, LEKSIKOGRAFIESE DATA, LEKSIKOGRAFIESE STRUKTURE

1.   Introduction

In a world that is becoming ever more internationalised and where intercultural communication is of growing importance for anybody who does not want to be overtaken by this development, there is a growing need for different sorts of dictionaries that can support the necessary communication. Among the basic dictionaries needed are the so-called learner's dictionaries that support or at least are supposed to support the foreign-language learner in the whole process of studying, learning and assimilating a given foreign language. Although the number of learner's dictionaries is not superabundant compared to other types of dictionaries, it is, nevertheless, possible to find various such dictionaries or at least one such dictionary for all of the most important international languages and also for a growing number of languages with a more regional or local importance. However, the big and ever bigger growing interest in learner's dictionaries is first of all reflected in the comprehensive literature covering this special field of lexicography. With a few exceptions, e.g. Shcherba (1940), Zöfgen (1994) and Tarp (1999a), this literature is mainly dealing with special problems of specific dictionaries and suffers from a lack of a theoretical lexicographic approach. The fact is that no general theory of learner's dictionaries has been developed so far and that only very few attempts have been made to remedy this shortcoming. It goes without saying that both the theoretical literature and the lexicographic practice are seriously affected by this regrettable lack of a general theory, although it must be admitted that some of the dictionaries are far better than the corresponding "theoretical" literature. This article therefore calls for the need to develop a general theory of learner's lexicography and indicates some of the ways that can be followed in order to reach this goal. This is all done under the motto that there is nothing as practical as a good theory.

2.   The concept of lexicographic theory

As mentioned above, the discussion in this article suggests that lexicographic theory — and especially the theory of learner's dictionaries — must be further developed. The first factor to be dealt with is, however, what is meant by theo-

224   Sven Tarp

ry. In fact, there are two competing ways of understanding lexicographic theory, the one contemplative and the other transformative (see Bergenholtz and Tarp 2003). The contemplative theory is a result of the observation of actual existing dictionaries and their use. Its relation to practice is limited to the continuous study of still more dictionaries in order to examine if the theory is valid or has to be adjusted. On the contrary, the purpose of the transformative theory of lexicography is not only to gain a good understanding of existing dictionaries and their use, but also, and first of all, to influence the lexicographic practice by means of indications and guidelines for future dictionary conception and making. Its relation to practice is a relation of interaction. It does not only study the finished product but also the human activity of conceiving, planning and making dictionaries, an activity in which it is heavily engaged itself. The contemplative theory has made a number of important contributions to lexicography but it often tends to submerge itself in the definition of a huge number of items and structures that are relevant to a linguistic and not to a lexicographic approach to dictionaries, and which alienates many lexicographers from this theory. Hence, there is little doubt that the desired general theory of learner's dictionaries should be developed as an integrated part of the transformative theory of lexicography.

3.   The relation between lexicography and linguistics

A major weakness in the theoretical literature dealing with learner's dictionaries is that the large majority of authors still regard lexicography as a subdiscipline of linguistics. They treat the problems of lexicography from the point of view of linguistics and, consequently, they do not see the need to develop an independent theory of lexicography and even less a theory of learner's lexicography. However, the fact that lexicography is considered an independent discipline with its own subject matter, i.e. dictionaries — and not language as in the case of linguistics — does not mean that there is no relation between lexicography and linguistics. According to the modern theory of dictionary functions (see Bergenholtz and Tarp 2002, 2003, 2004a, 2004b and Tarp 2003a), lexicography has a big interdisciplinary vocation and, as such, it treats — or ought to treat — linguistics in the same way as it treats any other discipline, i.e. it utilises the results from these disciplines whenever they are needed in order to conceive high-quality utility products that meet the specific needs of specific groups of users in specific user situations. In order to do this, lexicography studies human activities with the purpose of detecting specific needs that can be satisfied by means of dictionaries. In the case of learner's dictionaries, it studies the language-learning process in order to determine when and how dictionaries can assist this process. This study embraces a thorough analysis of the learner as a potential dictionary user (the lexicographic concept of a learner), of the general circumstances of the language-learning process as such, of the specific user situations where dictionaries can provide assistance and, finally, of the specific complex of user needs that arise for each specific type of

users in each specific type of user situation. On this basis, it is possible to determine the various functions of learner's dictionaries. This is the aim of the following discussion.

4.   The learner as a potential dictionary user

The functional theory of lexicography is a user-orientated theory and, as such, it considers it highly important to know the characteristics of the dictionary users in question, in this case foreign-language learners. In its study of these characteristics, lexicography has much in common with the branch of linguistics dedicated to the learning process, i.e. foreign-language didactics. However, the learner profile made by lexicography differs in many ways from that made by linguistics, as lexicography is only interested in those learners'/users' characteristics relevant for dictionaries. From the point of view of lexicography (and linguistics), a learner is a person in the process of learning a language not his/her native language. This concept, however, is a very broad one that covers a wide range of non-native speakers trying to learn a particular language. The following variables must, at least, be taken into account in order to make the concept less abstract and produce a more detailed profile of the potential dictionary user:

(a) Proficiency level
(b) Adult or child
(c) Emigrant or other type of learner
(d) "Joint" or separate culture
(e) Level of general culture

(a)   Language learning is a continuum where quantitative changes are transformed into qualitative changes. It is evident that this overall process can be subdivided into as many phases as necessary for each purpose. Linguistics usually distinguishes between three main levels: beginners, intermediate learners and advanced learners. This reduction of the number of proficiency levels is very convenient, but it is often based on schematic criteria such as the number of hours of instruction that do not contribute to a further understanding of the learner's special problems or needs. Wingate (2002), for example, defines an "intermediate learner" as a foreign-language student who has had more than 450 hours of instruction and an "advanced learner" as a student with more than 800 hours of instruction. Such a mechanical criterion makes it difficult to draw comparisons between Hong Kong learners (who are the subjects of Wingate's study) and learners who follow another educational system. It is, for instance, not the same to be 4 or 40 students in the classroom, to have much or little homework, to follow a study program that is teacher-orientated or one that is case-orientated outside the classroom and without the presence of the teacher, etc. And, apart from this, each and every student is different and develops his/her foreign-language skills in an individual way. Wingate herself illustrates

the problem. From a qualitative analysis based upon a think-aloud study of high verbal and low verbal "intermediate" Chinese students of German using two different types of German explanations — the so-called "new explanations" inspired by the principles of *Collins Cobuild English Dictionary* and the "traditional" explanations in *Langenscheidts Großwörterbuch Deutsch als Fremdsprache* (LGDAF) respectively — she (Wingate 2002: 209-210) concludes:

> High verbal ability students are on the threshold of being able to use a monolingual dictionary in the style of the new definitions, which is linguistically not as difficult as the LGDAF, while the weaker students at this proficiency level are not yet ready to use even this more user-friendly type of monolingual dictionary.

The conclusion that Wingate (2002) draws here is that in terms of user needs for explanation of foreign-language words there is actually no fundamental difference between beginners and the majority whom she considers "intermediate" learners. This conclusion might be interesting from the point of view of language didactics that has to recommend a specific type of dictionary for students of a specific proficiency level. But from the point of view of lexicography, it is hardly relevant at all. What is interesting for lexicography are those qualitative changes in the user profile that might modify the user needs and the character of the corresponding lexicographic data. This question still has to be analysed and discussed more profoundly by lexicographers. But at this point, it is already possible to distinguish some nodal points of interest in the curve of the learner's growing proficiency level in the foreign language:

(i)   When the learner starts to think in the foreign language and produce some easy phrases directly in this language;

(ii)   When the learner reaches a level where he/she is able to understand foreign-language explanations based on a reduced vocabulary of, for instance, 2 000–3 000 words;

(iii)   When the learner reaches a level where he/she mainly thinks and produces phrases directly in the foreign language, i.e. without a written or mental outline in the mother tongue; and

(iv)   When the learner reaches a level where he/she is capable of understanding complex explanations in the foreign language.

To each of these nodal points in the language-learning process correspond different — or the possibility of different — lexicographic solutions. When the learner starts to produce phrases directly in the foreign language, it is, for instance, possible (or necessary) to conceive a dictionary based on a list of foreign-language words in order to assist him/her with this kind of text production (see Tarp 2004). When the learner is able to understand foreign-language explanations based on a controlled vocabulary, it is possible to conceive dictionaries using this kind of explanations in order to assist him/her with text reception (see Tarp 2003b), etc. These considerations make it both possible and

necessary to draw a distinction between at least three different levels of language learning: beginners, experienced learners and advanced learners:

(i) *Beginners* are those who are taking the first steps into the foreign language and trying to assimilate the basic vocabulary and grammar.

(ii) *Experienced learners* are those who have mastered the basic vocabulary and the most important grammatical rules and are beginning to think and express themselves relatively freely in the foreign language, but sometimes have to pass through their mother tongue in order to produce more complex phrases.

(iii) *Advanced learners* are those who have acquired a considerable active and passive vocabulary, have assimilated the general grammar of the foreign language and are thinking and expressing themselves freely in this language, although they are still not "native".

There is hardly any doubt that the lexicographic concepts of "experienced learner" and "advanced learner" presuppose a relatively higher proficiency level than the corresponding linguistic concepts of "intermediate learner" and "advanced learner", as the two sets of concepts are interpreted above. Furthermore, future research might modify the lexicographic three-step graduation and even add further in-between levels. But if the aim is to improve learner's dictionaries, this should, as mentioned, always be done in the light of lexicography and not from the point of view of linguistics.

(b)   The proficiency level is by far the most important factor to distinguish between different learner characteristics, but there are also other factors that should be taken into account before conceiving or reviewing a dictionary. One of these is the distinction between learners who are adults and children respectively. A child is, for instance, still in the process of learning his/her mother tongue and reaching the adults' level of general culture and instruction and is supposed to have little experience in dictionary use. The learner's dictionary should be adapted to this reality and this would, among other things, imply that the selected vocabulary could be more reduced, that the definitions and the metalanguage should be easier, that explanations of some words would be needed instead of equivalents, that the lexicographic structures and layout should be simpler, etc.

(c)   Another relevant distinction is between what could be called "normal" learners on the one hand and immigrants and refugees on the other. This distinction is particularly important for countries where the consequences of neoliberal globalisation during the last decades has resulted in a growing influx of people who have flown from war, political repression, hunger or disagreeable social conditions in general. These people have a variety of backgrounds and are, in principle, not especially interested in learning the language spoken in the country in question, but they have to as it is part of their new life. Learning the new language is part of the integration process which includes learning

about the country's culture and the institutions and laws relevant for them. The authors of learner's dictionaries designed for immigrants and refugees must be aware of their special lexicographic needs which include, among others, the selection and explanation of words referring to the relevant laws, institutions and cultural phenomena. In some countries, e.g. the Scandinavian countries, a whole new branch of learner's dictionaries, the so-called immigrant's dictionaries, has actually developed during the last decades (see, e.g. Bergenholtz and Malmgren 1999). With a few exceptions such as the Swedish *Lexin* (see Gellerstam 1999), most of these dictionaries are noted for their poor lexicographic quality and it is an obvious task for professionally trained lexicographers to intervene in this field in order to assist people who have already suffered enough.

(d)   The compilers of a learner's dictionary must also have at least some knowledge about the target group's cultural background and whether this is close to or far from the culture related to the language that is being learned. The eleven official languages in a multilingual country like South Africa correspond, for instance, to different cultures but have, at the same time, many common cultural references. Such a reality must be reflected in learner's dictionaries. If the learner's culture is different from that of the foreign-language community, notes on cultures should be added and explanations written in such a way that they do not presuppose a cultural knowledge which first has to be acquired together with the language learning. One of the many problems of the famous British learner's dictionaries is, for example, that they are compiled for foreign-language learners in general without taking the very different cultural backgrounds of the target users into account. In many ways, the explanations in these dictionaries are very "British". This might not constitute a great problem for users coming from countries culturally and geographically close to Britain, but for learners with a completely different cultural background it sometimes creates insurmountable problems.

(e)   A last factor that must be taken into account in order to generate a user profile for a concrete dictionary is the learners' general cultural level. This aspect has, at least, consequences for the lemma selection, the character of the explanations, the possible cultural notes, the metalanguage and symbols and the structures of the dictionary. There is no reason to take a too high cultural level for granted as it will only make the users' access to the relevant data difficult or even impossible. The German learner's dictionary *De Gruyter Wörterbuch Deutsch als Fremdsprache* is an example of a dictionary that uses a condensed and codified metalanguage which presupposes a level of general linguistic knowledge the average learner cannot be expected to possess (see Tarp 2002).

5.   General circumstances of foreign-language learning

The user characteristics treated in the previous section are all important in order to construct a reliable user profile for any dictionary. However, apart

from these characteristics, there are also a number of user-related factors that are not user characteristics as such. These factors have to do with the general circumstances in which the language-learning process takes place and are, as such, relevant in order to determine the user needs and the corresponding lexicographic data to be included in the dictionary. These circumstances are:

(a) Conscious or spontaneous learning
(b) Learning inside or outside the language area
(c) Learning being exposed or not exposed to the mother tongue
(d) Learning with or without contrastive methodology
(e) Learning related or not related to a specific textbook or didactic system
(f) Learning related or not related to a particular topic (economy, history, culture etc.)

(a)   The distinction between conscious and spontaneous foreign-language learning was originally introduced by Shcherba in order to distinguish between the language-learning situation in Russia before and after the October Revolution in 1917 (see Mikkelsen 1992). Conscious learning is supposed to take place either as an individually planned and conducted study or related to a specific didactic system or text book. Spontaneous learning, on the other hand, takes place without these characteristics and is, as such, supposed to be more unsystematic. For many learners, the two processes might be interwoven with periods of conscious study followed by periods of spontaneous learning. Nevertheless, the characteristics of the two different processes have a number of important consequences for the design of learner's dictionaries. By definition, the only user situations where the spontaneous learner might need the assistance of a dictionary are text reception and text production (in their 007-inspired style, the British call them "decoding" and "encoding"). The conscious learner, on the other hand, might need assistance in other types of user situations as conscious study does not only consist of text production and text reception but also of other processes that are normally not supported by the existing learner's dictionaries and which constitute a big theoretical challenge for lexicography. The spontaneous learner might only consult the dictionary sporadically and cannot be expected to remember complex structures, symbols and abbreviations in the dictionary, whereas the conscious learner who might consult the dictionary several times every day would probably grow familiar with such complexities, and so on.

(b)   A second general circumstance of foreign-language learning that has serious consequences for learner's dictionaries is the geographical situation in which the learning takes place, i.e. inside or outside the area where the language in question is spoken. If it takes place outside the area — e.g. studying German in Hong Kong, French in South Africa or English in South America — it is much easier to know and control the vocabulary to which learners are exposed. On the contrary, if it takes place inside the language area, learners

would be exposed to thousands of words every day, on the streets, in shopping centres, restaurants, through television and radio channels, etc. If lexicographers want to satisfy learners' multiple and complex needs in this last case, they would have to select a large number of lemmata, even for dictionaries for text reception at beginner's level, whereas this might not be necessary for learners studying outside the language area.

(c)   A third general circumstance of foreign-language learning is related to the previous one, but quite the opposite, as it has to do with exposure to the mother tongue during the learning process. Such an exposure may have a number of negative consequences and can be observed among learners both outside and inside the foreign-language area. The discussion, however, is especially important in multilingual countries like South Africa where language learning frequently takes place in daily contact with the "foreign" language in question, although the learner in most cases will not only use this language but also his/her mother tongue, e.g. in the family, among the friends, etc. This is probably one of the more important explanations of the undesired mixed bilingualism that can be observed. Therefore, an effort should also be made to promote a conscious foreign-language study instead of spontaneous learning. Learner's dictionaries cannot solve these problems, but they can make a contribution in the right direction and support efforts to avoid mixed bilingualism by concentrating on the foreign-language system instead of a bilingual comparison between mother tongue and foreign language. This has, above all, relevance to some strategic decisions to be taken when planning a dictionary (see Gouws 1993, 1996, 2000).

(d)   Lexicography is not fanatic or extremist. Lexicographers might have their personal opinions, but the purpose of dictionaries is first of all to assist users however inconveniently these might behave. One of the questions discussed for years among linguists and language teachers is whether language teaching should use the contrastive methodology or not. There are different opinions although the majority are strongly against this methodology. Many lexicographers are obsessed with this discussion and many recommendations are to be found in the theoretical literature. However sympathetic their opinions might be, it is, nevertheless, not the task of lexicography to make such recommendations; it is the task of linguistics and language didactics. The task of lexicography is — or should always be — to study the real language-learning process followed by intended users and to support this process by means of dictionaries. This implies that if the target group uses the contrastive method of language learning, lexicography must assist them nevertheless, following the well-known cliché from many Hollywood films: "Don't take it personal, it's just a job."

(e)   Frequently, foreign-language learning is related to a specific textbook or didactic system during a certain period of the overall learning process. If a learner's dictionary is conceived to cover the users' needs during this specific period, then the vocabulary of the dictionary should be adapted to that of the

textbook or didactic system. It could, for example, as has already been done in a number of cases, be placed in the back matter of the textbook.

(f)   In countries with a multilingual composition of school children in the class rooms, the learning of a second — "foreign" — language is sometimes combined with the learning of a specific topic, e.g. dialects, history or geography. In such cases, the learner's dictionary might include "informations d'ordre dialectal, historique et géographique" (Mabika Mbokou 2001: 217).

6.   User situations

The learning of a foreign language is a complex process. Very schematically, it demands first of all reception of oral and written texts. Then it presupposes assimilation of what Shcherba called the foreign-language system, i.e. its lexis and grammar. And finally, it includes the training of the capacity to produce texts in this language, a process primarily carried out on the basis of the assimilation of the foreign-language system. These three partial processes — text reception, assimilation of the language system and text production — should, however, not be viewed mechanically as three stages where one stage necessarily precedes the others. On the contrary, they are in mutual interaction and in a certain way take place simultaneously, as one process always conditions the others.

During the reception phase, words are picked up and learned incidentally, and similarly, bit by bit, the lexical and grammatical structures of the language. The reception and understanding of still more complex texts presuppose the gradual assimilation of the foreign-language system as such. During the production phase, the acquired vocabulary is activated and confirmed and the same with the acquired general knowledge of the lexical and grammatical structures. Text production presupposes, up to a certain point, knowledge of the foreign-language system, but it can also be a means to assimilate this system if it is based on the consultation of outside sources, e.g. dictionaries. As regards assimilation of the foreign-language system, this is both a result — or by-product — of text reception and text production and a precondition for such a reception and production at a still higher level. The lexical and grammatical structures of the foreign language can, consequently, be assimilated in a natural or "passive" way through the processing (reception and production) of a large number of texts in this language. These structures can, however, also be assimilated in an "active" way through conscious study, as has been mentioned in the previous section.

Text production and text reception (encoding and decoding) have been discussed extensively in theoretical literature on lexicography, although there is still much research to be done in this respect. The study or conscious assimilation of the foreign-language system has, on the contrary, only been treated sporadically, in passing practically, in this literature, e.g. in Hausmann (1977), Mikkelsen (1992) and Tarp (1999a). As already indicated in the previous sec-

tion, this conscious study constitutes a big theoretical challenge for lexicography. It embraces a separate, or to be more precise, two separate user situations not analysed so far, i.e. conscious study of the foreign-language grammatical system and conscious study of the foreign-language lexical system. The decisive question for lexicographic theory and practice is whether dictionaries can assist these two user situations — and if they can, how and to what degree. The hypothesis of this article is that dictionaries actually can provide assistance to learners engaged in such a conscious study. This gives a total of four user situations relevant for learner's lexicography:

— text reception in the foreign language,
— text production in the foreign language,
— conscious study of the foreign-language grammatical system, and
— conscious study of the foreign-language lexical system.

Consequently, in order to prove the above hypothesis, it is necessary to analyse if there are any user needs related to the two user situations that can actually be assisted by means of lexicographic data. This is the aim in the following sections.

7.   User needs

User needs in terms of text reception and production have been thoroughly treated by Tarp (2004), and so have the corresponding lexicographic data. It has been clearly demonstrated that dictionaries can to a large extent satisfy the user needs in this respect. This and the following section will consequently concentrate on problems related to the conscious study of the foreign-language system. The first aspect to be noted is that the two component parts of this system, the lexis and the grammar, do not exist independently but are mutually interwoven. The lexis cannot be expressed without the grammar, and the grammar does not exist outside the lexis. However, through the method of abstraction, it is quite possible to isolate each of the two concepts and make them subject to a conscious study and teaching.

As for the assimilation of the lexis, it is not enough to learn the vocabulary and the meaning of its words and their pragmatic and stylistic properties. At least as important are the semantic fields of and the semantic relations between the words. Although few authors emphasise this last aspect, a word has not been learned until its semantic relations to other words have been assimilated as well (see Gouws 1989). It is in mutual confrontation with other words that the true meaning of a word is displayed. And this implies knowledge about semantic fields, synonymy, antonymy, hyperonomy, hyponomy, etc. Hence, in order to perform a conscious study of the lexis of a given language the learner needs information about:

— vocabulary,
— meaning of individual words,
— pragmatic and stylistic properties of individual words, and
— semantic relations between words.

As regards the assimilation of the grammatical system of a given language, the learner first of all needs information about:

— inflectional morphology,
— word formation (affixes, compounds and derivatives),
— syntax, and
— some formal rules of oral and written texts such as pronunciation, orthography, hyphenation and punctuation.

However, for all these categories — with the possible exception of punctuation — it is necessary to distinguish between at least two different levels of knowledge, i.e. the universal (general) and the individual, and sometimes also a further in-between level, the particular, mediating between these two by being universal in relation to the individual and individual in relation to the universal. This at least is the case in the majority of Indo-European languages. With respect to inflectional morphology it is, for example, not sufficient to learn the general principles of the inflection of nouns, verbs and adjectives, it is also necessary to know how these principles are applied at the individual level of each word, i.e. whether the inflectional pattern is regular or irregular and to what extent. The same applies to word formation. In a number of Germanic languages, compounds with nouns can, for instance, be made adding nothing, an "s" or even the medieval "e" to the noun. The general rules can be learnt but even native speakers would often be in doubt. As for syntax, each language has its general principles of word combinations whereas individual words of a language frequently display only part of these combinations and sometimes also some that have not been foreseen at the general level and which are often related to collocations, idioms and proverbs where atavistic forms may be found. And as for the formal rules of oral and written texts, in some languages it may occur that some words are not pronounced, spelled or hyphenated according to the general rules, a phenomenon often related to foreign words co-opted by the language.

It goes without saying that in order to assimilate the foreign-language grammar the learner must not only achieve an abstract knowledge of the system, but first of all, also a concrete knowledge, which is, in effect, the highest form of knowledge.

8.   Lexicographic assistance: data and structures (1)

In order to look for lexicographic data that can meet the above-mentioned user needs in terms of a conscious study of the foreign-language lexis and grammar,

the first step must be to study the existing dictionaries and find out whether they already provide data that can, at least partially, satisfy the relevant needs. The next step can then be to generalise these experiences and, on that basis, develop new solutions.

As regards lexicographic data that can assist the *assimilation of the foreign-language lexis*, this refers — as mentioned in the previous section — to data on vocabulary, meaning, pragmatic and stylistic properties of individual words and semantic relations to other words. As for the first type of data, it is obvious that existing dictionaries already provide data on foreign-language vocabulary. However, the main problem in a learner's dictionary is not the vocabulary as such, but to gain access to it as the user cannot be expected to read an alphabetically organised dictionary from one end to the other in order to increase his/her foreign-language vocabulary. Generally, there are two main ways of access: through the mother tongue and through the foreign language in question. The first is normally followed by learners at a beginner's level and, to a lesser degree, at an intermediate proficiency level. This will be their procedure when they have absolutely no other way of accessing the needed word or semantic field in the foreign language. The second way of access is followed by advanced learners and, to a larger and larger extent, by intermediate learners as they improve their foreign-language skills. The precondition for this procedure is that they already have access — through their knowledge of specific words — to the corresponding semantic fields from which they can continue increasing their foreign-language vocabulary. It is obvious that the two ways of access call for different lexicographic solutions. The first requires a bilingual word list or index, which can lead the user to the foreign-language words in question. Bilingual word lists are already available in existing dictionaries. Moreover, Tarp (2004) has presented a detailed proposal for a bilingual index especially conceived for foreign-language learners. It is therefore clear that there already exist lexicographic solutions that permit this first way of access to foreign-language vocabulary. The second way of access requires, on the other hand, solutions that link known foreign-language words with semantically and grammatically related words that may contribute to enrich the learner's foreign-language vocabulary. Several solutions of this sort can be found in existing dictionaries although they may have been developed for other specific purposes:

(a) References from individual articles to other individual articles where semantically and grammatically related words are treated.

(b) References from individual articles to special sections where related words are treated in a systematic way. Such sections can be directly attached to individual articles (synopsis articles) or they can be integrated in the dictionary as middle matter or outside matter.

(c) Systematically arranged dictionaries with an alphabetic index.

There are many examples of type (a) in existing dictionaries, although the majority of these are not learner's dictionaries, and this seems to be the most frequent of the three types.

As for the solutions of type (b), the *Cambridge International Dictionary of English* constitutes an example of a learner's dictionary that uses references to synopsis articles, e.g. from the lemma *button* to the section "Dressing and undressing", which is attached to the lemma *dress*. Here the user finds illustrations and explanations of various nouns and, especially, verbs related to the process of dressing and undressing. Macmillan's *English Dictionary for Advanced Learners* also contains articles with special sections where synonymous words are explained, although there are no references to these sections from other articles in the dictionary.

> **famous** […]
>   **1** if someone or something is famous, a lot of people know their name or have heard about them: *He dreamed of becoming a famous footballer*. […]
>   **2** *old-fashioned* very good: *a famous victory* […]
>
> ---
> **Other ways of saying famous**
> **well-known** fairly famous: *a well-known local reporter*
> **legendary** very famous and greatly admired: *Laurel and Hardy, the legendary comedy duo*
> **eminent** famous, and respected for their achievements, for example in science or academic work: *Professor Graham is an eminent brain surgeon*
> **notorious** famous for being a bad person or a criminal: *a notorious murderer / drugs baron / gang leader*
> **celebrity** someone who is famous and often talked about, especially an entertainer or sports player: *a magazine that shows photos of celebrities' homes*
> **star** a popular actor, musician, entertainer, or sports player who is very famous and successful: *a movie / football star*

Example 1:   Dictionary article and attached section from Macmillan (2002)

This dictionary article provides a good example of an article that makes it possible for the user who accesses the lemma *famous* to increase his/her English vocabulary through a conscious study of the attached section. The usefulness of the dictionary is, however, limited by the fact that it only includes a few articles of this sort and that no references are given to these enlarged articles. The *Oxford Advanced Learner's Dictionary* and the *Longman Dictionary of Contemporary English* constitute examples of learner's dictionaries where the user may find references from individual articles to special sections incorporated in the central word list as middle matter whereas the Swedish Lexin and the *De Gruyter Wörterbuch Deutsch als Fremdsprache* incorporate the same kind of sections in the

back matter. In the first three of these dictionaries, illustrations are used to support the systematic presentation of the semantic fields.

As for solutions of type (c), the only learner's dictionary of such a type known to the author of this article is the famous *Longman Lexicon of Contemporary English*. This dictionary contains a large number of semantic fields structured in a systematic way and a monolingual index giving access to these fields. One example of such a semantic field is «K107» that covers words related to winning, losing, etc. in terms of sport:

<div align="center">

K107 *verbs*: **winning and losing, etc**

[ALSO → C280-82]

</div>

---

win 1 […] to be the best or first in (a contest): *He won the race. Who won?* **2** […] **a** to gain (esp a prize): *He won a prize/cup/shield/a hundred pounds.* **b** to gain (the stated place) in a contest: *He won first place in the competition.* **3** […] to gain by effort or ability: *I can't win his friendship, though I've tried. By her hard work she won a place for herself on the school team.*

beat […] (esp in sports, games) to do better than (another): *He beat her at tennis. She beat him* (*in*) *running to the end of the road.* (*fig*) *This problem beats me.*

lose […] not to win: *He lost the race. Did he win or lose?*

defeat […] to beat; win a victory over (a person or group): *Our team defeated theirs.*

stump […] to be too difficult for: *I don't know the answer to that question; I'm stumped.*

tie […] to be equal to an opponent or his or her result in a competition: *He/They tied for second place in the game.*

draw […] to end (a game etc) without either side winning: *They drew* (*the game*) *5 points to 5 (= 5 all). We've had several drawn matches this year.*

result in […] to have as a result; cause: *The game resulted in a win for our side.*

Example 2:   Semantic field from Longman (1981)

The structure and presentation of this semantic field are similar to those of the field shown in example 1. Both enable the user to access and assimilate semantically related words. However, as Macmillan's *English Dictionary for Advanced Learners* only includes a few fields of this type whereas Longman covers a total of 15 000 systematically organised words, the usefulness of the latter is, consequently, far greater in terms of lexicographic assistance for increasing the vocabulary.

Word meaning and pragmatic and stylistic properties of individual words

constitute the second and third type of data needed to assist the assimilation of the foreign-language lexis. In this respect, it is evident that existing learner's dictionaries written for the purpose of text reception or text production already furnish this sort of data. An example of this is *Collins Cobuild New Student's Dictionary*:

> **consider** […] ⬚1⬚ VERB If you **consider** a person or thing to be something, this is your opinion of them. *I consider him a coward… Others consider the move premature… He considers that this is the worst recession this century.* ⬚2⬚ VERB If you **consider** something, you think about it carefully. *You do have to consider the feelings of those around you.* […]

> **USAGE** Note that when you are using the verb **consider** with a **that** clause in order to state a negative opinion or belief, you normally make **consider** negative, rather than the verb in the **that** clause. For instance, it is more usual to say '*I don't consider that you kept your promise*' than 'I consider that you didn't keep your promise'. The same pattern applies to other verbs with a similar meaning, such as believe, suppose, and think. *He didn't believe she could do it … I don't suppose he ever saw it… I don't think he saw me.*

Example 3:   Dictionary article and corresponding note on usage from Cobuild (2002)

Although a detailed research might show that these data or their presentation have to be somewhat modified in order to fully assist the user with the assimilation of the foreign-language lexis, at a general level it is nevertheless a fact that lexicographic practice has already found solutions providing for the above components of this particular function.

The fourth kind of data needed to assist the assimilation of the foreign-language lexis are those conceived to expose the semantic relations between words. However, as the above discussion of the procedure to increase the foreign-language vocabulary indicated, ways of access following the semantic relations between words, various methods to expose such relations have already been shown. What should be added here is that learner's dictionaries, at least for advanced learners, should not only provide synonyms, antonyms, etc. to single words but also to collocations, idioms and proverbs as Jónsson (1999) has pointed out. Jónsson himself has proposed a solution for this problem that includes an index leading the user to a systematic presentation of "phrases" (see Jónsson 2003) — a solution resembling the solution given in example 2.

To summarise: The above discussion has proven that dictionaries are perfectly capable of assisting the learner with the assimilation of the foreign-lan-

guage lexis. It has also been shown that if this was the only function of the dictionary, the best solution would probably be a systematically organised dictionary like the *Longman Lexicon of Contemporary English*. However, the discussion also indicated other solutions that could— at least partially — provide for this function if it, for one reason or other, had to be combined with other lexicographic functions. This article is not the place to go into details regarding this problem, but it should be noted that once the data are made available, the point at issue is an appropriate data distribution structure and a corresponding system of references that can guide the user to the relevant data. Although a long search path is not as user-friendly as a short one, it is nevertheless a way to solve a complex problem and provide assistance to the user in question.

9.   Lexicographic assistance: data and structures (2)

The lexicographic data needed in order to assist the user with the *assimilation of the foreign-language grammar* embrace — as mentioned in section 7 — data on inflectional morphology, word formation, syntax and formal rules of oral and written texts. However, the question immediately arises why it is altogether necessary to discuss the assimilation of the foreign-language grammar in relation to learner's dictionaries. There are already tools, e.g. grammars, especially conceived to assist or, at least partially assist, the learner in this respect. The answer has to do with the schism actually existing between grammars and dictionaries (see for instance Bergenholtz and Mugdan 1985). Grammars first of all deal with the universal grammatical rules of a given language and although they normally provide a number of individual examples in order to illustrate the general system, they are nevertheless not able to deal with all individual expressions of a universal rule. The problem is especially great when the general pattern does not apply completely at the individual level of each word, for instance in the case of regular and irregular inflectional morphology. Dictionaries, on the other hand, have the capacity, at least theoretically, to deal with all individual expressions, i.e. words and their grammatical properties. However, until now they have not been able to establish the desired bridge to the universal system combining the data included in the dictionary articles with the data located in the grammar or, to put it differently, combining the information retrieved from the dictionary articles with the information retrieved from the grammar. Bergenholtz (1984) was very much aware of this problem and suggested the conception of special grammars to be incorporated as independent component parts in dictionaries — a solution that can be found in various dictionaries (see also Mugdan 1989b). In many of these dictionaries, however, the result is only partial or formal as the schism between the universal and the individual frequently has been reproduced at the level of each dictionary. The solution to this problem is not only the conception of an advanced system of references but also the incorporation of data that can mediate between the two extremes.

Inflectional morphology is perhaps the most interesting area in this respect because a number of dictionaries have already conceived lexicographic solutions in terms of text production or text reception, which can also provide assistance to the assimilation of the foreign-language grammar. In general, inflectional paradigms of languages such as English, Afrikaans or Danish do not constitute a big lexicographic challenge. There are few inflected forms of nouns, verbs and adjectives and, in practice, all the forms, or at least all irregular forms, can easily be presented in individual dictionary articles without creating a too complex structure. This presentation can then be combined with a presentation of the general inflectional patterns in a special dictionary grammar as has already been done in several dictionaries. Although a number of learner's dictionaries for these languages have still not come up with a convincing solution — as Gouws and Tarp (2004) have shown for Afrikaans — it is far more interesting and relevant to discuss the problems related to languages such as German, Dutch, Russian, Icelandic, Italian, Portuguese, Spanish, etc., which all have a considerable number of inflected forms and where other lexicographic solutions must be found (see Mugdan 1989a).

Spanish verbs, which have up to 58 relevant forms to be treated, can be taken as an example. A recently published dictionary, Spansk–dansk erhvervsordbog, which is no learner's dictionary, has provided a solution that, nevertheless, could be used as a paradigm for Spanish learner's dictionaries. In a separate dictionary grammar, it provides the three regular inflectional patterns of Spanish verbs, a list of 16 slightly irregular inflectional patterns common to a large number of verbs and, finally, a list of completely irregular verbs. In the list with irregular inflectional patterns and the one with completely irregular verbs, only the irregular tenses are exposed, whereas the user's guide instructs the user to consult one of the three regular inflectional patterns for assistance when he/she has any problems with the remaining tenses. Moreover, all the regular and irregular inflectional patterns are illustrated by means of concrete words and, in this way, a relation is established between the universal and the individual, although only for a very limited number of words. In the individual dictionary articles, on the other hand, the verbs are treated in three different ways. As regards the regular verbs there is no data whatsoever on the inflectional pattern and, according to the instructions in the user's guide, this should be interpreted as an implicit reference to the regular inflectional paradigm presented in the back matter as part of the dictionary grammar. In the articles with completely irregular verbs, an explicit reference is provided to the list of irregular verbs in the dictionary grammar. And in the articles with verbs belonging to one of the 16 groups of irregular inflectional patterns common to a number of verbs, a "minirule" is provided:

complacer VB [c → zc foran a, o] […]

Example 4:   Dictionary article from Spansk–dansk erhvervsordbog (2003)

The minirule [c → zc foran a, o] indicates that *c* becomes *zc* before *a* and *o*. If the user does not know how to interpret this rule, it is at the same time a reference

to the dictionary grammar where he/she can find the following explanation and presentation of the inflectional pattern to which the verb *complacer* belongs:

> **§15. [c → zc foran a, o]**
>   i verber på *-acer*, *-ecer*, *-ocer* og *-ucir* ændres *c* til
>   *zc* foran *a* og *o*
>   eksempel: **conocer**

Præsens indikativ	*Præsens konjunktiv*	*Imperativ*
1. cono**zc**o	1. cono**zc**a	
2. conoces	2. cono**zc**as	2. conoce
3. conoce	3. cono**zc**a	3. cono**zc**a
1. conocemos	1. cono**zc**amos	1. cono**zc**amos
2. conocéis	2. cono**zc**áis	2. conoced
3. conocen	3. cono**zc**an	3. cono**zc**an

Example 5:   Excerpt from the dictionary grammar in Spansk–dansk erhvervs-
ordbog (2003)

In this paragraph, it is explained that the minirule [c → zc foran a, o] is valid for verbs terminating in *-acer*, *-ecer*, *-ocer* and *-ucir*, and it is also explained how this rule should be interpreted. The verb *conocer* is then used to illustrate the tenses and persons where this pattern is applicable. For the remaining tenses, an implicit reference is once more given to the list of regular verbs. In this way, the minirule mediates between the individual verb and the partially irregular inflectional pattern whereas the latter mediates between the minirule and the universal pattern for the inflection of Spanish verbs. With the above data and structures it is quite possible to use a dictionary in order to enhance one's knowledge of the foreign-language grammar in terms of inflection.

Word formation, which is the second component in the assimilation of the foreign-language grammar, embraces affixes, compounds and derivatives. The problem with regard to this varies from language to language and has been discussed in detail by Bergenholtz (2000), among others. In a language like Afrikaans, the use of diminutives is very frequent and this has also been reflected in some learner's dictionaries, e.g. *Basiswoordeboek van Afrikaans*:

> **oester (oesters; oestertjie)** *nw.* 'n Oester is 'n eetbare
> skulpdier met een plat en een hol skulp. […]

Example 6:   Dictionary article from Basiswoordeboek (1994)

In this article as in many others in Basiswoordeboek, the diminutive form *oestertjie* is provided immediately after the lemma. This is important in order to assist the user in understanding the lemma and expose the different ways it can be used. However, nowhere in the dictionary, neither in the word list nor in the outside matter, the user will find any introduction to the use of affixes and diminutives in Afrikaans and even less to the general meaning of the diminu-

tive *-tjie*. Therefore and because of the lack of mediations, the dictionary does not help the user to move from the individual to the universal, i.e. from the individual word to the universal system, and the conclusion is, accordingly, that a dictionary like Basiswoordeboek cannot in this respect be used to assist the assimilation of the language.

The Madagascan–German Rakibolana Malagasy–Alema provides another type of solution. This dictionary includes a Madagascan and a German grammar, written in both languages, and here the user will find, among much other information, a short introduction to the general use of affixes in Madagascan. At the same time, the most important affixes are selected as independent lemmata in the word list and then explained:

> **hatra** *ppk.* Präfix, das an Lokaladverbien angefügt
> wird; es drückt den Endpunkt einer örtlichen
> Strecke aus, z.B. *Handeha hatrery aho* bis da-
> hin/dorthin Tovona Atavone ny mpamaritra milaza
> toerana ilazana faritra tratrarina.

Example 7:   Dictionary article from Rakibolana Malagasy–Alema (1991)

In complicated cases, the explanation is even enhanced with a reference to the dictionary grammar, as is the case with the verbal morpheme *-a*:

> **-a** Verbalmorphem, das einem Wortstamm beigefügt
> wird, es dient **1)** zur Bildung der agissiven verben,
> z.B. *a-daboka* → § 1.5.4; **2)** zur Bildung der
> instrumentalen Verben, z.B. *a-haratra* → § 1.5.6.

Example 8:   Dictionary article from Rakibolana Malagasy–Alema (1991)

Another dictionary, Duden (1985), which has the subtitle "Wortbildung und Wortschatz" (Word Formation and Vocabulary) and is presented as a "Lern-wörterbuch" (learning dictionary), includes a section in the front matter where the user can find a list of all the affixes and affixoids treated and explained in the lemma list. In all these ways, the user gets an easy overview of the affixes and affixoids and a general introduction to their meaning and use in German.

The above method can also be used to treat compounds from both the universal and the individual point of view. Many dictionaries already contain and treat compounds in various ways. In order to solve the problems relevant to this article, a possibility would be to conceive a special dictionary grammar where the universal rules of compound words are explained and then, in all the relevant articles, concrete examples, minirules and even references to the grammar are provided as in the following proposal for a Danish dictionary:

> **barn** SUB ⟨et; -et, børn, børnene⟩ *child* […]
> -barn (§ 76.1) adoptivbarn, enebarn, spædbarn
> barn- (§ 76.2) barndom, barnløs
> barns- (§ 76.3) barnsben

> barne- (§ 76.4) barnebarn, barnedåb, barnegråd,
>   barnevogn
> børne- (§ 76.5) børnehave, børnehjem, børnesang

Example 9:   Proposal for the treatment of compounds in a dictionary article

First the minirules are provided (*-barn*, *barn-*, *barns-*, *barne-* and *børne-*), then the references to the relevant paragraphs in the dictionary grammar and, finally, the individual examples of how to apply the minirule. The choice could, of course, have been made that references should only be provided in more complicated cases, i.e. *barne-* and *børne-*. Nevertheless, the above proposal once more establishes a relation between the individual and the universal through a number of mediations, thus showing that dictionaries can also be designed to assimilate the foreign-language grammar in terms of forming compounds.

As regards derivatives, this problem has been thoroughly discussed by Bergenholtz (2000) who also makes a number of suggestions that can be adapted to the purpose of this article. First of all, the problematic of derivatives should be briefly treated in the dictionary grammar, as is, for instance, done in Rakibolana Malagasy–Alema (1991). Then all the derivatives — or at least all the relevant ones — should be selected as independent lemmata with a system of explicit or implicit references between them (or between the most important of them):

> **exclude** VERB […] → excluding PREP; exclusion
>   NOUN; exclusive ADJ; exclusivity NOUN
> **excluding** PREP […] → exclude VERB
> **exclusion** NOUN […] → exclude VERB
> **exclusive** ADJ […] → exclude VERB; exclusivity NOUN
> **exclusivity** NOUN […] → exclude VERB; exclusive
>   NOUN

Example 10:   Proposal for the treatment of derivatives in dictionary articles

It is obvious that explicit references could also be made to the corresponding paragraphs in the dictionary grammar, if this were considered necessary. Nevertheless, the above proposal shows that dictionaries can also provide for the assimilation of the grammar of derivatives. It should moreover be noted that the above data — just as the data on affixes and compounds — also support the assimilation of the foreign-language vocabulary and thereby indicate that it may be possible to integrate the two functions.

Syntax, together with inflectional morphology, undoubtedly constitute the most important and complex part of the grammar of a given language. If one for instance consults the Big Five English learner's dictionaries — Oxford (2000), Cambridge (1996), Longman (1995), Cobuild (2001, 2002) and Macmillan (2002) — it emerges that different methods have been used in order to expose the syntactic properties of individual words. This is hardly surprising as British lexicography is not based on functional theory and decisions taken by the edi-

tors frequently are more influenced by linguistics than well-defined user needs.
Three examples will show the variations:

> **suggest** *obj* SHOW/EXPRESS […] [T] to communicate
> or show (an idea or feeling) without stating it di-
> rectly or giving proof. *Recent polls seem to suggest*
> ***that*** *the government's popularity is at an all time*
> *low.* [+ *that* clause] • *Evidence suggests* ***that*** *expo-*
> *sure to lead may cause mental damage in children*
> *before they show any adverse physical symptoms.* [+
> *that* clause] • *She's applied for a lot of jobs recently,*
> *which suggest* ***that*** *she's not altogether happy with*
> *her present situation.* [+ *that* clause] • *She made one*
> *or two remarks which suggested* ***that*** *the company*
> *was in financial trouble.* [+ *that* clause] • *Now, I'm*
> *not suggesting* ***that*** *you were flirting with Adrian,*
> *but you were certainly paying a lot of attention to*
> *him.* [+ *that* clause] • *Are you suggesting* ***(that)*** *I*
> *look fat in these trousers?* [+ (*that*) clause] • *Some-*
> *thing about his manner suggested a lack of interest*
> *in what he were doing.*

Example 11:   Dictionary article from Cambridge (1996)

> **suggest** […] *verb* **1** ~ **sth (to sb)** to put forward an
> idea or a plan for other people to think about SYN
> PROPOSE. [VN] *May I suggest a white wine with this*
> *dish, Sir? A solution immediately* ***suggested itself to***
> ***me*** (= I immediately thought of a solution). [V
> (**that**)] *I suggest (that) we go out to eat.* [V **-ing**] *I*
> *suggested going in my car.* [VN **that**] ***It has been***
> ***suggested that*** *bright children take their exams*
> *early.* (*BrE* also) *It has been suggested that bright*
> *children should take their exams early.* **2** ~ **sb/sth**
> **(for sth)** ⁄ ~ **sb/sth (as sth)** to tell sb about a suitable
> person, thing, method, etc. for a particular job or
> purpose SYN RECOMMEND: [VN] *Who would you*
> *suggest for the job? She suggested Paris as a good*
> *place for the conference. Can you suggest a good*
> *dictionary?* help You cannot 'suggest somebody
> something': ~~Can you suggest me a good dictionary?~~
> [V **wh-**] *Can you suggest how I might contact him?*
> **3** ~ **sth (to sb)** to put an idea into sb's mind, to make
> sb think that sth is true: [V (**that**)] *All the evidence*
> *suggests (that) he stole the money.* [VN] *The symp-*
> *toms suggest a minor heart attack. What do these*
> *results suggest to you?* **4** to state sth indirectly SYN
> IMPLY: [V (**that**)] *Are you suggesting (that) I'm*
> *lazy?* [VN] *I would never suggest such a thing.*

Example 12:   Dictionary article from Oxford (2000)

suggest […] verb [T]

1 to offer an idea or a plan for someone to consider: *The report suggested various ways in which the service could be improved.* • **+ (that)** *I suggest we have dinner first, and then watch the film.* • **+ what/why/where etc** *Can anyone suggest what we should do to increase sales?* • **suggest doing sth** *If there is a mechanical problem, we suggest contacting the manufacturer directly.*

2 to tell someone about something that may be useful or suitable for a particular purpose: REC-OMMEND: *Can you suggest an inexpensive restaurant? • The doctor suggested rest as an alternative to surgery.*

3 to say that something is likely to exist or be true: IMPLY: *The cause of the problem, as Hutton suggests, is the leadership's refusal to accept any form of criticism.* • **+ (that)** *Are you seriously suggesting that she did this on purpose? • I'm not suggesting that giving up smoking will be easy.* **3a.** to make you think that something is likely to exist or be true: *As the name suggests, the Carlton Beach Hotel is situated near the sea. • Evidence suggests a link between asthma and pollution.* • **+ (that)** *Recent research suggests that the drug may be beneficial to people with muscular disorders.*

4 *formal* to remind you of something: *The architecture suggested a chapel.*

Example 13:   Dictionary article from MacMillan (2002)

None of the above dictionaries are conceived to assist the user in the assimilation of the foreign-language syntax, but it is evident that they all provide some kind of help in this respect as they all combine the individual examples with small notes that have the character of minirules. Cambridge (1996) provides the minirule after the text example(s), probably because it gives priority to text reception, but this solution nevertheless paves the way for some sort of incidental assimilation of the minirule. Oxford (2000) and Macmillan (2002) provide the minirules before the text examples, probably because they give priority to text production (Oxford even provides them just after the definition number). This solution sets the scene for a more systematic introduction to the syntactic properties of *suggest*. However, as the dictionaries are not conceived for this purpose the microstructure chosen makes the syntactic overview rather complicated. It is not a problem as such that the articles are structured according to the various meanings and senses, but it is clearly a problem that access to the minirules is rather difficult. Oxford is probably the most informative of the three dictionaries but its highly codified items (*sb*, *sth*, *wh-*) put strong limitations on its user-friendliness. A proposal for a solution based on the article from Oxford could be as follows:

suggest […] *verb*

1 to put forward an idea or a plan for other people to think about

⇒ *propose.*

suggest something: May I suggest a white wine with this dish, Sir?

suggest something to somebody: A solution immediately suggested itself to me.

suggest + (that) clause: I suggest (that) we go out to eat.

suggest + -ing phrase: I suggested going in my car.

suggest + (that) clause: It has been suggested that bright children should take their exams early.

2 to tell somebody about a suitable person, thing, method, etc. for a particular job or purpose

⇒ *recommend*

suggest something: Can you suggest a good dictionary?

suggest somebody for something: Who would you suggest for the job?

suggest something as something: She suggested Paris as a good place for the conference.

suggest + how clause: Can you suggest how I might contact him?

3 to put an idea into somebody's mind, to make somebody think that something is true:

suggest + (that) clause: All the evidence suggests (that) he stole the money.

suggest something: The symptoms suggest a minor heart attack.

4 to state something indirectly

⇒ *imply*

suggest + (that) clause: Are you suggesting (that) I'm lazy?

suggest something: I would never suggest such a thing.

Example 14:   Proposal for the treatment of syntax in a dictionary article (based on Oxford 2000)

It is obvious that this proposal may be modified in various ways (e.g. in terms of layout) but it provides, nevertheless, an example of how the syntactic properties of a word can be accentuated and made easily accessible without destroying the semantic presentation of the word. The latter also shows that the lexicographic assistance to assimilation of the syntax can be combined with the assistance to text reception, and as the syntactic minirules, at the same time, are essential for text production, this function can also easily be performed by the same dictionary (see Tarp 2004). Moreover, if the lexicographer wants to improve the assistance given to the assimilation of the syntax, the data provided in the individual articles should be combined with a more universal presenta-

tion in a dictionary grammar. The only English learner's dictionary mentioned above that includes a dictionary grammar is Cobuild (2001) and although this 222-page grammar is well structured and easy to understand, it is not integrated with the lemma list and the access in terms of concrete problems of the above mentioned types is also rather difficult.

Nevertheless, in the end the above analysis of the English Big Five shows that all the necessary ingredients — individual words, minirules, individual text examples and dictionary grammars — are already available in learner's dictionaries and that it is only a question of designing appropriate lexicographic structures and access (search) paths in order to conceive a learner's dictionary that can assist the user in the assimilation of the syntactic system of the language being learned.

The last problem to be discussed in this section has to do with some formal rules of oral and written texts such as pronunciation, orthography, hyphenation and punctuation. These formal rules can easily be exposed in a separate dictionary grammar. However, in many languages not all of these rules are strictly applied at the individual level of each word and sometimes there are also different options where the user may need some guidance.

As regards pronunciation, the dictionary grammar can fully expose the general rules of pronunciation of a given language. But in some languages, like English and Danish, there will frequently be various hypothetical ways of pronouncing a concrete word and the learner would therefore never know the pronunciation if it is not explicitly given. Furthermore, in most languages some individual words (e.g. proper names and foreign words recently co-opted in the language) will differ from the general pronunciation pattern. If the learner's dictionary is conceived for the assimilation of the spoken language, indications of pronunciation should therefore be given at the individual level of all the words in question. This is particularly relevant if the dictionary is an electronic one with sound as the written indications of pronunciation are often extremely difficult to interpret.

As regards orthography, it is evident that the general rules can be explained in the dictionary grammar and that each dictionary article automatically reflects how these rules are applied at the individual level of each word, also if they for one reason or other break the universal rules.

This also applies to hyphenation, which frequently is very difficult to explain and assimilate at the general level, for which reason it is recommended that all learner's dictionaries dealing with written language should indicate the hyphenation related to each and every word.

Punctuation is a slightly different topic. In this case, the general rules should first of all be explained in the dictionary grammar as punctuation is a phenomenon related to the phrase and text level and is difficult to treat at word level. However, in a few cases such as some conjunctions and insertions — *however*, *nevertheless*, *in effect*, etc. — it could be useful for the assimilation of the general punctuation rules to include a note in this regard on the use of commas.

To summarise: The discussion in this section has proven that dictionaries are quite capable of assisting the learner with the assimilation of the foreign-language grammar. In line with Mugdan (1989a, 1989b), it has also been shown that the best solution is to integrate a dictionary grammar, which explains the universal grammatical rules of the language in question, with the articles in the word list where the individual expressions of these rules can be exposed and where so-called minirules can play the role of the particular mediating between the individual and the universal. The concrete proposals put forward in this section suggest a solution to the abyss that actually exists between the language system described in grammars and the grammatical properties of the individual words treated in dictionaries.

The analysis in this and the previous section also showed that a number of dictionaries already, at least up to a certain point, assist the learner in the assimilation of the foreign-language lexis and grammar. It showed that the majority of the lexicographic data needed to assist the conscious study of the foreign-language system are already available in existing dictionaries and that the main challenge is to conceive an appropriate data distribution structure and a corresponding system of explicit and implicit references together with easy access and search paths (see Tarp 1995, 1999b and Lauridsen 1998). In this respect, the two sections also indicated that the discussion about dictionary structures necessarily has to be based on the respective lexicographic functions and not taken in isolation from the real user needs. It goes without saying that the proposals put forward have to be elaborated and adapted to each language, user group and type of learner's dictionary.

10.   Functions of learner's dictionaries

From the discussions in this article it appears that learner's dictionaries can have a variety of different lexicographic functions. They can be designed to assist three different types of users (beginners, intermediate and advanced learners) in four different types of user situations, of which two — text reception and text production — are communication-orientated and the other two — assimilation of the foreign-language lexis and grammar — are knowledge-orientated. This gives a total of 12 basic lexicographic functions of learner's dictionaries:

—   communication-orientated functions

(a) to assist learners at beginner's level with L_2 text production
(b) to assist intermediate learners with L_2 text production
(c) to assist advanced learners with L_2 text production
(d) to assist learners at beginner's level with L_2 text reception
(e) to assist intermediate learners with L_2 text reception
(f) to assist advanced learners with L_2 text reception

—   knowledge-orientated functions

(g) to assist learners at beginner's level with the assimilation of L$_2$ grammar

(h) to assist intermediate learners with the assimilation of L$_2$ grammar

(i) to assist advanced learners with the assimilation of L$_2$ grammar

(j) to assist learners at beginner's level with the assimilation of L$_2$ lexis

(k) to assist intermediate learners with the assimilation of L$_2$ lexis

(l) to assist advanced learners with the assimilation of L$_2$ lexis

As mentioned in the previous sections, when a dictionary is conceived and planned these functions must be described in further detail according to the specific foreign language, the specific user profile, the specific circumstances of foreign-language learning and the specific user needs that may be expected in each user situation. If the dictionary in question is going to have more than one function, the lexicographer must also make decisions regarding the principles of combination and the possible priorities in case of conflicting solutions. All this is a topic for future research.

The discussion in this article also suggests that the very concept of a learner's dictionary has to be redefined. A learner's dictionary does not refer to just one type of dictionaries, but to a whole branch of dictionary types. It is a superordinate concept that embraces a whole range of dictionaries with different functions and combinations of functions. This should be reflected in the whole process of dictionary planning and making, in dictionary reviewing and also in the presentation of each dictionary so as to enable the potential user to choose the dictionary that exactly fits his/her specific needs.

11.   How to use a learner's dictionary

Hausmann (1977) made a distinction between a "learning dictionary" (Lernwörterbuch) and a "consultation dictionary" (Konsultationswörterbuch). According to this typology, either the learning dictionary is to be studied from beginning to end (primary learning dictionary) or its individual articles are to be studied in this way (secondary learning dictionary), whereas the consultation dictionary is to be used for consultations related to text reception and text production. Hausmann himself notes that learning dictionaries of the above type are rare. He only puts forward a very reduced proposal as to how they could be conceived and he apparently considers learning dictionaries and consultation dictionaries to be exclusive to each other.

It is difficult to imagine a learner who would study the foreign language word by word and/or semantic field by semantic field. Such a method is contrary to all present knowledge about language learning. Although language learning could and in many cases should be a conscious process, this does not mean the learner has to study the dictionary in the above-mentioned way. It is

well known that back in the 19th century when comparative linguistics dominated and little progress had been made concerning language-didactic methods, many foreign-language students studied and learned the foreign language by means of a bilingual dictionary. The result, however, was an extreme degree of mixed bilingualism. Whether a learner's dictionary is bilingual or monolingual, alphabetic or systematic, no learner is recommended to learn the foreign language word by word or semantic field by semantic field. So the learner's dictionary should be used as a reference work and conceived accordingly. But this does not exclude that it could be used and conceived to be used in order to support a conscious study of the language system.

A learner engaged in a conscious study of a foreign language will many times consult the learner's dictionary for assistance to solve problems related to text reception or text production. Once the learner has found the lemma in question, he/she will frequently become interested in knowing more about the word and will then proceed to a conscious study of other parts of the same dictionary article or even other articles within the same semantic field. The consultation process begins with one purpose — for assistance in terms of communication — and is then transformed into a conscious study with another purpose — for knowledge about the foreign language in question. The two different ways of using a learner's dictionary are therefore not exclusive to each other. In this way, the traditional opposition between the dictionary as a consultation tool and as a learning tool, has also been dissolved. Instead, lexicographic theory should distinguish between monofunctional and multifunctional dictionary use.

12.   Conclusion

This article has concentrated on some basic problems of learner's lexicography. It has laid the foundation of a new theory of learner's dictionaries and has indicated some ways through which this theory can be further developed. The main challenge now is to transform the present knowledge into practical dictionary concepts for guiding the production of a new generation of learner's dictionaries that, hopefully, will meet the learner's needs even better than the existing ones.

Literature

Dictionaries

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

Cambridge 1996 = **Procter, Paul (Ed.).** 1996. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press.

Cobuild 1995 = **Sinclair, John (Ed.).** 1995. *Collins Cobuild English Dictionary*. London: HarperCollins.

Cobuild 2001 = **John Sinclair (Ed.).** 2001. *English Dictionary for Advanced Learners*. London: Harper-Collins.

Cobuild 2002 = **Sinclair, John (Ed.).** 2002. *Collins Cobuild New Student's Dictionary*. London: HarperCollins.

De Gruyter 2000 = **Kempcke, Günter (Ed.).** 2002. *De Gruyter Wörterbuch Deutsch als Fremdsprache*. Berlin/New York: Walter de Gruyter.

Duden 1985 = **Müller, Wolfgang (Ed.).** 1985. *Das Bedeutungswörterbuch. Wortbildung und Wortschatz. Ein Lernwörterbuch mit Bedeutungsangaben, Anwendungsbeispielen und Abbildungen, mit sinn- und sachverwandten Wörtern und den Bausteinen des Wortschatzes*. Mannheim/Vienna/Zurich: Duden.

Langenscheidt 1993 = **Götz, Dieter, Günther Haensch and Hans Wellmann (Eds.). In collaboration with Vincent J. Dochertz and Günther Jehle.** 1993. *Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Das neue einsprachige Wörterbuch für Deutschlernende.* Berlin/Munich: Langenscheidt.

Lexin 1993 = **Gellerstam, Martin (Ed.).** 1993. *Lexin. Språklexikon för invandrare. Svenska ord — med uttal och förklaringar*. Uppsala: Norstedts Förlag.

Longman 1981 = **McArthur, Tom.** 1981. *Longman Lexicon of Contemporary English*. Harlow: Longman.

Longman 1995 = **Summers, Della.** 1995. *Longman Dictionary of Contemporary English*. Harlow: Longman.

Macmillan 2002 = **Rundell, Michael (Ed.).** 2002. *English Dictionary for Advanced Learners*. London: Macmillan.

Oxford 2000 = **Wehmeier, Sally (Ed.).** 2000. *Oxford Advanced Learner's Dictionary*. Oxford: Oxford University Press.

Rakibolana Alema–Malagasy 1994 = **Bergenholtz, Henning (Ed.).** 1994. *Deutsch–Madagassisches Wörterbuch/Rakibolana Alema–Malagasy.* Antananarivo: Tsipika.

Rakibolana Malagasy–Alema 1991 = **Bergenholtz, Henning (Ed.).** 1991. *Rakibolana Malagasy–Alema/Madagassisch–Deutsches Wörterbuch.* Antananarivo: Leximal.

Spansk–dansk erhvervsordbog 2003 = **Tarp, Sven (Ed.).** 2003. *Spansk–dansk erhvervsordbog*. Copenhagen: Gads Förlag.

Other literature

Bergenholtz, Henning. 1984. Grammatik im Wörterbuch: Syntax. Wiegand, Herbert Ernst (Ed.). 1984. *Studien zur neuhochdeutschen Lexikographie V. Germanistische Linguistik*: 1-46. Hildesheim/Zurich/New York: Olms.

Bergenholtz, Henning. 2000. Lexikographie und Wortbildungsforschung. Barz, Irmhild, Marianne Schröder and Ulla Fix (Eds.). 2000. *Praxis- und Integrationsfelder der Wortbildungsforschung*: 19-30. Heidelberg: Winter.

Bergenholtz, Henning and Sven-Göran Malmgren (Eds.). 1999. Temanummer om indvandrerordbøger. *LexicoNordica* 6.

Bergenholtz, Henning and Joachim Mugdan (Eds.). 1985. *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28.–30.6.1984*. Lexicographica. Series Maior 3. Tübingen: Max Niemeyer.

Bergenholtz, Henning and Sven Tarp. 2002. Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen. *Lexicographica* 18: 253-263.

Bergenholtz, Henning and Sven Tarp. 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 171-196.

Bergenholtz, Henning and Sven Tarp. 2004a. The Concept of "Dictionary Usage". Dollerup, Cai (Ed.). *Worlds of Words. A Tribute to Arne Zettersten. Nordic Journal of English Studies* 3(1): 23-36, 2004.

Bergenholtz, Henning and Sven Tarp. 2004b. Wörterbuchfunktionen. Barz, Irmhild, Henning Bergenholtz and Jarmo Korhonen (Eds.). 2004. *(Lerner-)Leksikographie: Ein- und zweisprachig für Deutsch und Finnisch*. Frankfurt am Main/Bern/New York/Paris: Peter Lang.

Gellerstam, Martin. 1999. LEXIN — lexikon för invandrare. *LexicoNordica* 6: 3-18.

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

Gouws, Rufus H. 1993. Afrikaans Learner's Dictionaries for a Multilingual South Africa. *Lexikos* 3: 29-48.

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

Gouws, Rufus H. 2000. Towards the Formulation of a Metalexicographic Founded Model for National Lexicographic Units in South Africa. Wiegand, H.E. (Ed.). 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 109-133. Lexicographica. Series Maior 100. Tübingen: Max Niemeyer.

Gouws, Rufus H. and Sven Tarp. 2004. Wie leer wat uit Afrikaanse (aan)leerderwoordeboeke. *Tydskrif vir Geesteswetenskappe*.

Hausmann, Franz J. 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Max Niemeyer.

Hausmann, Franz J., Oskar Reichmann, Herbert E. Wiegand and Ladislav Zgusta (Eds.). 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.

Jónsson, Jón Hilmar. 1999. Fraseologiens plass i ordbøker for innvandrere. *LexicoNordica* 6: 65-78.

Jónsson, Jón Hilmar. 2003. Fraseologien i forgrunnen — fraseologisk register som ledd i ordbokens tilgangsstruktur. Hansen, Zakaris Svabo and Anfinnur Johansen (Eds.). 2003. *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 21.–25. august 2001*: 151-167. Skrift 7. Torshavn: Nordisk forening for leksikografi i samarbejde med Nordisk Sprogråd og Føroyamáldeild Fróðskaparseturs Føroya.

Lauridsen, Carsten. 1998. Fordelingsstrukturen af de faglige informationer og dens betydning for henvisningsstrukturen. Tarp, Sven (Ed.). 1998. *Leksikografi som speciale. Bind I*: 61-99. Århus: Spansk Institut og Center for Leksikografi, Handelshøjskolen i Århus.

Mabika Mbokou, Ludwine. 2001. Le rôle du dictionnaire dans le système éducatif. Emejulu, James Duplessis (Ed.). 2001. *Éléments de lexicographie gabonaise. Tome I*: 206-222. New York: Jimacs-Hillman.

Mikkelsen, Hans Kristian. 1992. What Did Shcherba Actually Mean by "Active" and "Passive" Dictionaries? Hyldgaard-Jensen, Karl and Arne Zettersten (Eds.). 1992. *Symposium on Lexicography V. Proceedings of the Fifth International Symposium on Lexicography May 3–5, 1990, at the University of Copenhagen*: 25-40. Tübingen: Max Niemeyer.

Mugdan, Joachim. 1989a. Information on Inflectional Morphology in the General Monolingual Dictionary. Hausmann, Franz J., Oskar Reichmann, Herbert E. Wiegand and Ladislav Zgusta (Eds.). 1989–1991: 518-525.

Mugdan, Joachim. 1989b. Grundzüge der Konzeption einer Wörterbuchgrammatik. Hausmann, Franz J., Oskar Reichmann, Herbert E. Wiegand and Ladislav Zgusta (Eds.). 1989–1991: 732-749.

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

Tarp, Sven. 1995. Distribution Structure. Bergenholtz, Henning and Sven Tarp (Eds.). 1995. *Manual of Specialised Lexicography*: 188-190. Amsterdam/Philadelphia: John Benjamins.

Tarp, Sven. 1999a. Lørnerordbøger for indvandrere og andet godtfolk. *LexicoNordica* 6: 107-132.

Tarp, Sven. 1999b. Theoretical Foundations of the So-called Cross-reference Structures. Wiegand, H.E. (Ed.). *Mediostrukturen in Fachwörterbüchern. Lexicographica* 15: 114-137, 1999.

Tarp, Sven. 2002. Functions in *De Gruyter Wörterbuch Deutsch* als Fremdsprache. Wiegand, H.E. (Ed.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutchen II. Untersuchungen anhand des* De Gruyter Wörterbuch Deutsch als Fremdsprache: 609-619. Lexicographica. Series Maior 110. Tübingen: Max Niemeyer.

Tarp, Sven. 2003a. Basic Elements of Lexicographic Theory/Éléments de base de la théorie lexicographique. Emejulu, James Duplessis (Ed.). 2003. *Éléments de lexicographie gabonaise. Tome II*: 7-35. New York: Jimacs-Hillman.

Tarp, Sven. 2003b. The Usefulness of Different Types of Articles in Learner's Dictionaries. Review of "Ursula Wingate: The Effectiveness of Different Learner Dictionaries: An Investigation into the Use of Dictionaries for Reading Comprehension by Intermediate Learners of German. Tübingen: Niemeyer 2002". *Hermes, Journal of Linguistics* 30: 215-234, 2003.

Tarp, Sven. 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.

Wingate, U. 2002. *The Effectiveness of Different Learners Dictionaries: An Investigation into the Use of Dictionaries for Reading Comprehension by Intermediate Learners of German*. Lexicographica. Series Maior 112. Tübingen: Max Niemeyer.

Zöfgen, E. 1994. *Lernerwörterbuch in Theorie und Praxis. Ein Beitrag zur Metalexikographie mit besonderer Berücksichtung des Französischen*. Lexicographica. Series Maior 59. Tübingen: Max Niemeyer.

The Use of Dictionaries in the Light of Language Teaching Methods[*]

Refilwe M. Ramagoshi, *Department of African Languages, University of Pretoria, Pretoria, Republic of South Africa (refilwe.ramagoshi@up.ac.za)*

Abstract: This article takes a critical look at the role played by second language teaching methods used in teaching African languages as first languages. The traditional method of teaching, used before 1984, encouraged rote learning and teacher-centred lessons. This did not foster the use of dictionaries and, if they were used, the lack of a dictionary culture made their use ineffective. Teachers should be well versed in the functional approach, introduced in 1984, to ensure a smooth transition to Outcomes Based Eductation (OBE). Although its inception since 1994 offers a greater scope for using dictionaries, teachers are still unsure about how the syllabi should be implemented. This, together with the lack of a dictionary culture, results in an inadequate and ineffective dictionary use. Some suggestions are made how this situation could be rectified.

Keywords: AFRICAN LANGUAGES, FIRST AND SECOND LANGUAGE TEACHING, TRADITIONAL EDUCATIONAL METHOD, ROTE LEARNING, FUNCTIONAL EDUCATIONAL APPROACH, OUTCOMES BASED EDUCATION, INTERPRETATION OF SYLLABI, AVAILABILITY OF DICTIONARIES, DICTIONARY USAGE, DICTIONARY CULTURE

Opsomming: Die gebruik van woordeboeke in die lig van taalonderrigmetodes. Hierdie artikel werp 'n kritiese blik op die rol wat tweedetaalonderrigmetodes speel by die onderrig van Afrikatale as eerste tale. Die tradisionele metode van onderrig wat voor 1984 gebruik is, het meganiese leer en onderwysergesentreerde lesse aangemoedig. Dit het nie die gebruik van woordeboeke bevorder nie, en, indien hulle gebruik is, het die gebrek aan 'n woordeboekkultuur hul gebruik oneffektief gemaak. Onderwysers moet goed onderleg wees in die funksionele benadering wat in 1984 ingevoer is, om 'n vlot oorgang na Uitkomsgebaseerde Onderrig (UGO) te verseker. Alhoewel die instelling daarvan sedert 1994 'n groter geleentheid vir die gebruik van woordeboeke gebied het, is onderwysers steeds onseker hoe uitvoering aan die sillabusse gegee moet word. Dit, saam met 'n gebrek aan 'n woordeboekkultuur, het 'n ontoereikende en oneffektiewe woordeboekgebruik tot gevolg. 'n Aantal voorstelle word gemaak hoe hierdie situasie verbeter kan word.

Sleutelwoorde: AFRIKATALE, EERSTE- EN TWEEDETAALONDERRIG, TRADISIONELE ONDERRIGMETODE, MEGANIESE LEER, FUNKSIONELE ONDERRIGBENADERING, UITKOMSGEBASEERDE ONDERRIG, INTERPRETASIE VAN SILLABUSSE, BESKIKBAARHEID VAN WOORDEBOEKE, WOORDEBOEKGEBRUIK, WOORDEBOEKKULTUUR

---

[*]   This article is a revised version of a paper read at the Seminar on Learners' Dictionaries presented by the African Association for Lexicography at the University of Pretoria, Pretoria, Republic of South Africa, 11–12 November 2002.

1.   Introduction

The high Matriculation failure rate among Black students is blamed on the fact that the majority of learners, teachers and parents find it difficult to communicate, read or write in English. In the *Sunday Times*, 30th January 2000, page 2, John Burmeister, National Director: The Molteno Project, states that Matriculation results could be improved by "accelerating the acquisition of competence in English — it being the popular choice of medium".

This addresses the symptoms, not the cause. At the then College of Continuing Training in Soshanguve, which closed down in 1994, it was found that teachers who came to the in-service training, of which the main aim was the improvement of their teaching skills, struggled not only with the methods of teaching an African language, but also with the content of different aspects of grammar and literature, because they could not interpret the syllabi.

The causes for this situation in language education and its impact on a dictionary culture will be dealt with in the following discussion.

2.   Three Phases in Education

In South Africa, three phases of education can be distinguished, as shown in the following diagram:

PHASE ONE	PHASE TWO	PHASE THREE
Traditional	Functional or	Outcomes Based
Method or	Communicative	Education
Rote Learning	Approach	

Prior to 1984, most African language teachers used the audio-lingualism theory, which encouraged the drill method, repetition and habit formation, resulting in no distinction between first and second language teaching.

If the first language were used as the language of learning in Grades 0–7 this, at least, would serve as a bridge before acquiring a second language. As a result, at present, both teachers and learners are struggling to learn. Sand and Forthering (1999: 42) suggest that a learner should at least continue with mother-tongue literacy while learning another language although problems might be caused by time constraints. On the same note, Afolayan (1999: 1) states: "Colonial experience has made Africa the continent where the child's mother-tongue is alienated within the educational system."

A teacher not well versed in using a first language approach tends to resort to drill work and rote learning. He/she teaches what he/she remembers from being a learner, sometimes without a clear understanding of the subject matter. This turns into a vicious circle as the teacher literally forces learners to rote learn what he/she was also forced to rote learn by his/her previous teachers. Sand and Forthering (1999: 47) reinforce this observation when they say:

"Learning to read and write in a language you don't know is a very slow and difficult thing to do."

Gravett (2001: 17) distinguishes between rote learning which does not require active thinking, and meaningful learning which entails an understanding of what is being learnt. Instead of proceeding from the known to the unknown, most teachers do not take into consideration what learners already know from home before coming to school. Ausubel (1978[2]: 163-164) expresses this notion as follows:

> Existing cognitive structure … is the principal factor influencing meaningful learning and retention … Logically meaningful material is always, and can only be, learned in relation to a previously learned background of relevant concepts, principles, and information … It is evident … that the substantive and organizational properties of this background crucially affect both the accuracy and the clarity of these emerging new meanings and their immediate and long-term retrievability.

Construction of an integrated knowledge is important and should be linked to something familiar, as the basis of meaningful learning. It is important that learners should first recognise that the new knowledge relates to what they already know.

Teachers assume they have to start from the beginning when teaching learners. For example, in a lesson on qualificatives, the teacher will make the learners repeat sentences that indicate colours. Yet learners come to school already familiar with colours as they have clothes or toys of different colours. Lessons then become teacher-centred and formulaic in nature as learners are forced to drill the examples as given by teachers. Goodman (1986: 7) summarises it well when he states: "Many school traditions seem to have actually hindered language development. In our zeal to make it easy, we've made it hard."

In 1984, the Department of Education introduced the functional or the communicative approach in first language teaching. By *functional approach* is meant a method where the living language forms the basis for the study of linguistic phenomena. It does not permit unnatural forced language usage even though the language may be grammatically correct. Leading questions play an important role. The teacher exposes learners to a specific grammatical form and helps them discover the relevant linguistic rules. Learners are lead from what they already know to the unknown — the important point here is that this new knowledge and insight are acquired by means of discovery (Malimabe 1997: 34-35).

Each province or region sent only one representative for a specific African language to the Department of Education. On arriving back at school, the representative was expected to report back to all first language teachers and duplicate handouts on the functional approach for each teacher at the inspector's offices. This did not happen. Thus only a few teachers were familiar with the system.

Another problem this educational approach faced was resistance to change — teachers just seemed to be too comfortable in their traditional method of teaching. This same attitude can also be seen with the introduction of the Outcomes Based Education (OBE) or Curriculum 2005, the new educational system envisaged after the 1994 elections.

OBE is a system built around the process and the outcomes, or the results of learning. According to the White Paper of October 1997, it represents a teaching approach that moves away from the mere gathering of knowledge to an integrated approach, by focusing on the application of knowledge, skills and values. Rote learning has to make way for self-led engagement. This will improve competency as well as values, attitudes or orientations that would benefit learners after leaving school or completing training.

This approach will ensure life-long learning and will benefit the young and the old who left school without acquiring certificates or who have never studied formally but have some sort of experience and skill.

OBE, however, frustrated many teachers due to the difficult terminology used. Teachers were told to forget about differences between aims and objectives of a lesson and concentrate on the term *specific outcomes* without first being taught the differences between the three terms. Teachers who had never attended courses on the functional approach, encountered more difficulties as they only knew the traditional method. Those who passed through the second phase had an easier transition.

Marambana's research, conducted in 1987, indicated that there were 20 000 teachers without certificates. Pressure was put on teachers, some of whom had only completed Matriculation, some the Junior Certificate (J.C.) and some the Primary Teacher's Course (P.T.C.), to improve their qualifications through distance education — hence the birth of the Vista University Distance Education Campus (VUDEC). Some of these teachers taught (especially African languages) because they could not find other work or present the subjects they favoured. Others taught African languages because they were mother-tongue speakers.

However, even certificated teachers who came to the in-service training, had problems with teaching approaches. Snyman et al. (1996: 152), indicated that shortcomings in the curriculum for teacher training were due to the fact that

—   90% of the time was spent on theories of grammar or literature, and the history of languages, and

—   only 10% of the time was spent on different skills or methods of teaching grammar and literature.

2.1   Problems Caused by Rote Learning

Examples of these can be given from the writing of essays, the testing of comprehension and the teaching of poetry.

(a)   Writing essays

Learners are often given a full-fledged essay to write without having been taught pre-essay-writing exercises. For example: the use of stylistic devices like long and short sentences for creating a certain atmosphere should be explained to learners.

Another example is that of encouraging learners to employ the five senses, e.g. touching, smelling, seeing, hearing and tasting in developing writing techniques that force learners to use certain parts of speech, idioms, proverbs, ideophones, etc. These aspects of grammar are still taught in isolation and not in context. When marking essays, teachers often just indicate incorrect words without substituting the correct word or encouraging learners to look up the right word in the dictionary (Malimabe 1990: 72). Teachers also used to give learners model essays written by themselves — a procedure which does not encourage creativity and initiative on the learner's part.

(b)   Testing comprehension

Some teachers are unable to identify the keywords in a sentence, the keysentences in a paragraph and the main idea of a passage. They are also unable to explain the meaning of words or use idiomatic expressions in context. This inability of some teachers to summarise, paraphrase or explain is a major problem in the teaching of African languages. For example, in an examination learners were asked to summarise the contents of M. Kgasi's poem "Sefikantswe" (tombstone or statue) in their own words. The learners performed poorly, for, even though they were familiar with the words 'tombstone' or 'statue', they did not know the meaning of the word *sefikantswe* in Tswana. The word was misinterpreted as follows: se- + -fika + -ntswe.

se-	=	class prefix
-fika	=	Zulu verbal stem meaning "to arrive"
-ntswe	=	Tswana word meaning "voice"

As a result the content of the poem as given in their own words did not make sense at all.

This same problem of misinterpretation is also found in Afrikaans. Mother-tongue speakers, who teach second language speakers, tend to neglect the teaching of simple everyday terminology. For example, learners misinterpreted the following topics in essay writing:

'n Nare droom (An Unpleasant Dream)

In the Sotho languages, *nare* is a buffalo and is spelt the same way as the Afrikaans word *nare* (unpleasant). So learners confused these two words and wrote: *'n Groot nare het my gejaag, en ek het gehardloop en gehardloop.* (A big buffalo chased me, and I ran, and ran.)

Poetsbakkery (Practical Joking)

Learners only recognised part of the word *bakkery* (baking) and wrote: *My ma het 'n koek gebak. Dit was baie lekker.* (My mother baked a cake. It was very tasty.)

In a recent examination paper students in the "Culture in Literature" section were asked to compare courtship in Western and African cultures. One student interpreted the word "court" as having the legal meaning.

From these examples it can be seen that problems are caused by a lack of vocabulary and dictionary usage.

(c)   Teaching poetry

A number of teachers lack the knowledge of suitable procedures of teaching poetry correctly. If a poem, for example, refers to an animal such as a cow, they assume the poem deals with the advantages that can be derived from a cow, i.e. use it for obtaining milk, for ploughing, for paying lobola, etc. In Tswana, the poem "Malenkhu a Marumo" by P. Leseyane compares the problems brought about by owning or not owning cattle. If you do own some, you do not sleep because other groups attack you or wild animals hunt the cattle. If you do not own any, you will never be in a position to get married, because you will be unable to pay lobola.

This reflects on the teaching approach of reciting poems using actions without understanding the meaning of the poems. This point is emphasised by Wästberg (1968: 10) who states: "A poem's overall success is equally dependent on the writing and translation of the author as well as the interpretation of the reader through his or her knowledge of the subject matter." Many traditional poems have a high percentage of "difficult" vocabulary. Teachers neglect to use a dictionary to look up these words, and if they do, the words often do not appear in the dictionaries available in the African languages. Most second language speakers who are given literature in the African languages to read are frustrated that they do not find the meaning of many of the words that appear in the text. Lecturers end up by advising them to ask other mother-tongue speakers for the meanings, or otherwise lecturers must waste time explaining the words.

From the above examples it can be deduced that the traditional or rote learning method is to a great extent responsible for the problems experienced. This method does not encourage the use of dictionaries. Rumboll et al. (1988) in the foreword to *English in Context Grade 5* states:

> The reading, understanding and comprehending Text's sections sets out to move beyond the traditional practice of simply close reading the material for its supposed ultimate meaning, to goal-orientated creative interaction with written, spoken and visual texts. The exercises are designed to improve comprehension (approaching the texts at an informal, concrete level and responding at a more abstract, formal level).

> While creative self-expression is encouraged in the sections on writing, talking, and listening, the primary intention there is to develop and refine the skills of clear, effective communication. The material on the use of dictionaries and other reference works deals with a further essential yet often neglected aspect of language skills acquisition.

The *English in Context* series is the only series of prescribed text books in which the use of dictionaries are encouraged. Most text books in the African languages do not.

2.2   Problems Related to Outcomes Based Education (OBE)

Although Outcomes Based Education (OBE) offers a greater scope for the use of dictionaries, this is not always realized as teachers are often unprepared for implementing the syllabi to the full. Teachers encourage learning from the known to the unknown, but they neglect or ignore what learners already know, i.e. using the functional approach as a basis or link as is the case with the outcomes approach. For example, after presenting a lesson on adjectives according to the functional way of the deductive and inductive approach, the facilitator can then consider what learners could do should they leave school immediately after the lesson on adjectives. Keeping in mind that they can differentiate things/people through colour, size/shape and numbers, the facilitator can then inform them about the different careers they can follow based on a lesson on adjectives. For example: interior decorating, dressmaking, art, make-up art, etc. All these careers need knowledge of and experience in colour, size/shape and numbers.

The five Compulsory Phase Organisers that have been introduced are:

— Personal development and empowerment,
— Communication,
— Environment,
— Culture and society, and
— Economy and development.

All eight Learning Areas therefore have the same themes. This means that each Phase Organiser could have about fifteen programmes which will be impossible to cover in one year. It is important that all facilitators of different Learning Areas should plan together to avoid overlapping of topics and confusing learners. For example, in Economy and Development, one of the Programmes is learning about a bank. A Language, Literacy and Communication (LLC) teacher can teach a text about being mugged or robbed while at the bank. Some terminology about the bank will be introduced, but the focus will be on language teaching. The Economic Management Sciences (EMS) teacher on the other hand, will teach all aspects of the banking system and management, bonds,

savings accounts, current accounts, loans etc. When assignments of different teachers overlap, it shows that proper lesson planning has been neglected. There prevails much uncertainty among teachers which, for instance, results in the following:

—   Some teachers copy the Specific Outcomes and use all of them for writing any programme. Not all Specific Outcomes will be relevant for every topic.

—   Being unable to do the application, facilitators sometimes rote learn the theory on outcomes. This could easily encourage teachers to resort to what they are familiar with — using a second language approach in the teaching of a first language.

3.   Problems of Dictionary Use

3.1   Lack of the Availability of Dictionaries

The government does not supply dictionaries, encyclopaedias (which do not exist in the African languages) and informative materials to schools, nor does it even subsidise the acquiring of these materials to foster proper learning.

Some schools do not even have libraries where reading is encouraged, let alone dictionaries with the help of which user skills can be taught. If there is a library, the question to be asked is: How many dictionaries and types of dictionaries plus encyclopaedias should be kept in the library?

Some teachers/facilitators do not have dictionaries in the classroom. The dictionaries found in class, if any, are for the teacher's/facilitator's personal use when studying — not to encourage learners to use them.

Some parents concentrate on buying study materials other than dictionaries, especially dictionaries for the African languages. English dictionaries in private schools, especially previously White schools, are part of the study materials, but not dictionaries in the African languages.

3.2.1   Lack of a Dictionary Culture

Through all eleven years of schooling, no dictionaries are used in class. Learners come into contact with dictionaries only at Matriculation level and then usually only as a result of reading Shakespeare where they encounter difficult words. In African languages, even the issues of the series *Terminology and Orthographies* are not made available or compulsory to learners. Some teachers are even unaware of these publications.

In most of the model lessons from Grades 3–6 the use of dictionaries is not encouraged. Teachers normally write out the difficult words from a text and explain them.

Some teachers are not skilled to do critical analysis in literature, which

requires a broad knowledge of words, their meanings and effect in order to interpret a text. These teachers are less likely to encourage learners to use dictionaries.

Many words learners encounter in their reading are not included in dictionaries available in the African languages, and these also do not give other essential information such as tonal stress marks for pronunciation. Deficiencies like these may easily cause learners to lose interest in using dictionaries.

At the beginning of every dictionary there are specific notes informing the user how to employ the dictionary. Before trying to find the meanings of words, the learner or teacher should first read and master this section thoroughly. So far, no lesson or course showing learners and teachers how to decode this section has been introduced. Most activities encourage arranging words in alphabetical order only.

In an attempt to encourage the use of dictionaries, one of the OBE publications encourages learners to build databanks with words from the dictionary. This is not a suitable method of teaching dictionary usage. Instead of building a databank of words from a text such as a story they are studying in class, learners randomly select words they do not know. However, if the text they are studying has words, which do not appear in the specific dictionary used, teachers will still have to explain the words with the result that learners will still not often use a dictionary.

Another popular feature in OBE publications is puzzles. Learners are given instructions to fill in the puzzles, which makes it possible for them to work with words. However, in the instructions or in the teacher's guide, no mention is made of asking or encouraging learners to use a dictionary while trying to do the puzzle.

3.2.2   Suggestions for Establishing a Dictionary Culture

PanSALB has established dictionary units for all the official South African languages. The units face the challenge of producing dictionaries urgently needed, especially picture dictionaries for Grades R–3 in the African languages. If learners are taught how to use dictionaries in these classes, an ongoing use of dictionaries will be effected.

Each grade should have a dictionary suitable for its level of education. This will not be a problem because the syllabus is structured in such a way that the same themes/programmes can be taught in the Five Compulsory Phase Organisers.

Workshops for teachers should be encouraged in which they are taught strategies of incorporating the use of dictionaries into learners' activities. These may include the following:

— In every grade, a period should be set aside for teaching learners how to use a dictionary and how to use the front matter, interpret the signs, abbreviations, explanations, etc.

— Before the reading period in school, learners should read a chapter of a prescribed book at home to give them an opportunity to use a dictionary. Reading a prescribed book in class with learners in the higher grades, like Grades 9–12, should be avoided.

— There is a tendency to play word games without emphasising the usage of words in context. If there are dictionaries in the classroom, learners should be encouraged to look up the meanings of words.

Many of today's children are computer-literate and could use on-line dictionaries effectively, provided the teacher is also competent to give the necessary guidance.

4.   Conclusion

There is no perfect method of teaching, which is a very complex activity. In most cases, the method used in a lesson will be determined by different situations, the type of subject being taught, the resources, time and the type of teacher and learner. The culture of using dictionaries will also be determined by these situations. Teachers can play an important role in enhancing dictionary use in class if they know how to incorporate it in their teaching. However, without good African language dictionaries that are revised on a regular basis, learners will continue compiling lists of words in the classroom without knowing the importance of having a good dictionary and knowing how to use it. Teachers need to be the ones who do the lesson presentations in the form of micro-teaching lessons where different methods of teaching may be explored and evaluated. However, if the workshop presenter does all the talking, giving his/her examples only, rote learning will still be encouraged.

Bibliography

Afolayan, A. 1999. *The Alienated Role of the Mother Tongue Literacy Education in Sustainable National Development: The Western Nigerian Yoruba Example*. Paper presented at a Reading for All: Pan African Strategies Conference organised by the Department of Education, Pretoria, Republic of South Africa, 6–9 August 1999.

Ausebel, D.P. 1978². *Educational Psychology: A Cognitive View*. New York: Holt, Rinehart and Winston.

Goodman, K. 1986. *What's Whole in Whole Language?* Portsmouth: Heinemann.

Gravett, S. 2001. *Adult Learning: Designing and Implementing Learning Events — Dialogic Approach*. Pretoria: Van Schaik.

Malimabe, R.M. 1990. *The Influence of Non-Standard Varieties on the Standard Setswana of High School Pupils*. Unpublished M.A. Dissertation. Johannesburg: Rand Afrikaans University.

Malimabe, R.M. 1997. *Subject Didactics: Northern Sotho, Southern Sotho, Tswana*. Pretoria: University of South Africa Press.

Marambana, N.R. 1987. *A Decade of Crisis in Black Education, 1976–1986*. Unpublished M.Ed. Dissertation. Sovenga: University of the North.

Rumboll, F.C.H. et al. 1988. *English in Context Grade 5*. Cape Town: Maskew Miller Longman.

Sand, S. and R. Forthering. 1999. *Teaching in Adult Basic Education (ABE)*. Pietermaritzburg: University of Natal Press.

Snyman, J.W. et al. 1996. The Current Position of Setswana in the Republic of South Africa. Legère, K. (Ed.). 1996. *Cross-Border Languages*. Okahandja: Gamsberg Macmillan.

South African Government. 1997. *White Paper on Outcomes Based Education.* Pretoria: Government Printer.

Wästberg, Per (Ed.). 1968. *The Writer in Modern Africa*. Uppsala: The Scandinavian Institute of African Studies.

Monolingual and Bilingual Learners' Dictionaries[*]

Rufus H. Gouws, *Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa (rhg@sun.ac.za)*

Abstract: When deciding on the best learners' dictionary for a specific user and a specific situation of usage one often has to make a choice between a monolingual and a bilingual learners' dictionary. This article discusses some aspects of the user-driven approach so prevalent in modern-day lexicographic thought, focuses broadly on dictionary typology and takes a closer look at monolingual and bilingual learners' dictionaries. Some problems users experience when learning a new language, e.g. language distortion and problems related to the phenomenon of false friends, especially in closely related languages, are mentioned. It is indicated that a typological hybrid dictionary could assist certain users. The importance of an unambiguous identification of the relevant lexicographic functions is emphasised and the notions of *function condensation* and *function merging* are introduced. It is shown that the typological choice should be determined by a function-based approach to dictionary usage.

Keywords: BILINGUAL DICTIONARY, FALSE FRIENDS, FUNCTION CONDENSATION, FUNCTION MERGING, GENUINE PURPOSE, LEARNERS' DICTIONARY, LEXICOGRAPHIC FUNCTIONS, MONOLINGUAL DICTIONARY, TEXT PRODUCTION, TEXT RECEPTION, TYPOLOGICAL HYBRID, TYPOLOGY.

Opsomming: Eentalige en tweetalige aanleerderwoordeboeke. Wanneer besluit moet word oor die beste aanleerderwoordeboek vir 'n spesifieke gebruiker en 'n spesifieke gebruiksituasie moet daar dikwels gekies word tussen 'n eentalige en 'n tweetalige aanleerderwoordeboek. Hierdie artikel bespreek bepaalde aspekte van die gebruikersgedrewe benadering wat kenmerkend is van die moderne leksikografiese denke, fokus breedweg op woordeboektipologie en gee in meer besonderhede aandag aan sekere aspekte van eentalige en tweetalige aanleerderwoordeboeke. Bepaalde probleme wat gebruikers ervaar by die aanleer van 'n vreemde taal, bv. taalversteuring en probleme verwant aan die verskynsel van valse vriende, veral in nou verwante tale, kry aandag. Daar word aangedui dat 'n tipologiese hibriede woordeboek van waarde kan wees vir sekere gebruikers. Die belang van 'n ondubbelsinnige vasstelling van die tersaaklike woordeboekfunksies word beklemtoon en die begrippe *funksieverdigting* en *funksiesamesmelting* word aan die orde gestel. Daar word aangetoon dat die tipologiese keuse bepaal moet word deur 'n funksie-gebaseerde benadering tot woordeboekgebruik.

Sleutelwoorde: AANLEERDERWOORDEBOEK, EENTALIGE WOORDEBOEK, FUNKSIE-

---

SAMESMELTING, FUNKSIEVERDIGTING, LEKSIKOGRAFIESE FUNKSIES, TEKSPRODUKSIE, TEKSRESEPSIE, TIPOLOGIE, TIPOLOGIESE HIBRIED, TWEETALIGE WOORDEBOEK, VALSE VRIENDE, WERKLIKE DOEL

Introduction

There is nothing as practical as a good theory. Theoretical lexicography may never be regarded as theory merely for the sake of theory. In metalexicography theoretical deliberations should always endeavour to enhance the quality of the lexicographic practice. The lexicographic practice is performed by means of dictionaries. Consequently dictionaries have to be regarded as practical tools in the hand of the language user. As practical tools dictionaries should be devised in such a way that a well-identified user can retrieve the kind of information needed for a specific purpose. This implies that dictionaries should be seen as utility instruments, to be used by the ordinary member of a speech community. A dictionary has to reflect the real language as it is encountered in daily communication and not the language which purists would necessarily regard as the preferred form.

In order to be functional instruments, dictionaries have to fulfil a specific purpose. The identification and formulation of this purpose has to precede the compilation process because the compilation process has to be steered by the purpose of the specific dictionary project. In metalexicographical terms this purpose of a dictionary is known as the *genuine purpose* (cf. Wiegand 1998). According to Wiegand, the genuine purpose of a dictionary is that it should be used for the retrieval of information from lexicographic data regarding those linguistic expressions included for treatment in the specific dictionary.

Tarp (2000: 198) directs his description of the genuine purpose of a dictionary at a concrete product which has to be used in specific circumstances to solve specific problems:

> The dictionary covers this or that area and is conceived to assist users with these or those characteristics in this or that situation in order to solve problems of this or that sort.

Taking the cue from Wiegand (1998), Gouws (2001) defines the genuine purpose of a dictionary as follows:

> The genuine purpose of a dictionary implies that a dictionary is produced so that the target user who uses the dictionary in a typical usage context will have an instrument to assist him in achieving a successful dictionary consultation procedure by reaching the goals that motivated the search. The genuine purpose of a dictionary should therefore be to ensure successful dictionary consultation procedures. A successful dictionary consultation procedure depends on the way in which the needed linguistic information can be retrieved.

The last number of years have witnessed the emergence of a user-driven approach in lexicography. This is made clear by Hartmann (1989: 103, 104) when he says:

"An analysis of users' needs should precede dictionary design" and "The design of any dictionary cannot be considered realistic unless it takes into account the likely needs of various users in various situations." Besides the emergence of a focus on the user an awareness of dictionary functions has also increased. Wiegand (2001: 217) welcomes the fact that the latest dictionary research pays so much attention to dictionary functions. It can rightfully be argued that the compilation of any dictionary should not only be preceded by a clear identification of the intended target user, the needs and the reference skills of the target user and the user and usage situation of the target user but also by an unambiguous decision regarding the functions of the planned dictionary.

When determining the functions of a dictionary and the way in which the lexicographic treatment will ensure that the needed functions prevail the lexicographer needs to be familiar with the intended user of the dictionary. Tarp (2000: 194) indicates the importance of this knowledge regarding the characteristics of the target users of a dictionary. These characteristics include their language competence, i.e. which language is their mother tongue, at what level they have mastered their mother tongue and at what level they have mastered the foreign language. When speaking about the user needs Tarp (2000: 195-196) says that a user might need among others

— information about the native language,

— information about a foreign language,

— a comparison between the native and a foreign language, and

— information about culture and the world in general.

This is not only true of general monolingual and/or bilingual dictionaries but also of learners' dictionaries and should lead to a choice by the lexicographer regarding the functions of the intended dictionary. A dictionary can be monofunctional, bifunctional or polyfunctional. Economic and pragmatic reasons have led to a situation where the publication of dictionaries is often not determined by the needs of the users but by the financial situation of the publishing house. Consequently lexicographers often have to embark on procedures of function merging or even function condensation because more than one function has to prevail in a given dictionary. In many cases this can only be possible if the dictionary does not give a full realisation of a given function but allows it in a condensed way.

Dictionary typology

The needs of the user do not only influence the contents and the structure of a dictionary but the typological classification of a dictionary should also be a direct result of the needs in a specific speech community. When dealing with dictionary typology one should be careful that the nature and extent of traditional typological categories do not impede the successful implementation of

user-determined functions. To illustrate this: working only with a broad and general typological category like *bilingual dictionary* says very little about the needs of e.g. primary school learners or professional translators. Neither does it say anything about text reception or text production or about native or foreign language users. A mere typological classification like *monolingual dictionary* or *bilingual dictionary* is actually a very blunt instrument which has a limited value when it comes to the functions of the dictionary.

Working with typological classifications one should negotiate the fact that many typological categories need a thorough subclassification to ensure a proper and valid data presentation and distribution in terms of the real needs of real users. However, parallel to the notions of function merging and function condensation run the notions of typological merging and typological condensation. One of the implications of these notions is the obliteration of subtypological categories in favour of broad and general typological categories, the so-called "superheroes" of dictionary typology. In the field of dictionary typology the term *typological hybrid* is used to indicate a dictionary which displays features from more than one typological category, e.g. a monolingual descriptive dictionary with a translation equivalent included in the comment on semantics of each article. The term *typological hybrid* could actually also be applied to the typological superheroes because they display features from various typological categories. In a speech community with a less-developed dictionary collection, it is more likely than not that one would find superheroes representing what should ideally be represented by a range of smaller dictionaries belonging to a variety of subtypological categories.

Learners' dictionaries represent one typological category which tends towards superhero status on account of the diverse nature of the data included in learners' dictionaries and on account of the wide spectrum of the user groups at which learners' dictionaries are directed. When one looks at dictionaries displaying the typical functions of learners' dictionaries it is evident that two distinct user groups can be determined, i.e. learners using the dictionary in a school environment and learners not involved with formal school activities. In his classification of dictionaries, Gouws (2001: 74) uses the category *pedagogical dictionaries* to include both learners' and school dictionaries. There it is argued that school dictionaries are primarily for mother-tongue speakers learning their own language whereas learners' dictionaries are for studying a foreign language. *Pedagogical dictionaries* is also a superhero term and not a very apt term because it does not really present a distinctive and specific lexicographic category due to the fact that the majority of dictionaries do have a pedagogic assignment in terms of e.g. their knowledge- and their communication-directed functions.

The term *school dictionary* is also not a very apt designation — it has the nature of a mini-superhero. So many different dictionaries could be classified under this term, including picture dictionaries for pre-primary use, monolingual and bilingual dictionaries for the different school phases as well as lan-

guage for special purpose dictionaries dealing with the terminology of school subjects. All of these dictionaries display certain features typically relevant to learners' dictionaries.

The term *learners' dictionary* should rather be used as an umbrella term for both school- and non-school-directed dictionaries which display typical learners' features.

It is important to look at the primary use of a learners' dictionary. In a school situation, it could be used by the teacher as an teaching instrument or by the learner as a learning instrument. In German, a distinction is made between *Lehrwörterbücher* (teaching dictionaries) and *Lernwörterbücher* (learning dictionaries). When discussing learners' dictionaries one should also make provision for these two features which are relevant to this dictionary type.

For the present discussion *learners' dictionary* will be used as a superhero term but the emphasis will be on a subclassification within this broad category. Although the traditional terms *bilingual* and *monolingual* will be used in this subclassification it will also be indicated that these terms are often misused and frequently fail to convey the real nature of a specific dictionary.

Monolingual or bilingual

Before referring to learners' dictionaries, one must have a clear and unambiguous understanding of the terms *bilingual dictionary* and *monolingual dictionary.* One of the first Afrikaans dictionaries, the *Proeve van een Kaapsch Hollandsch Idioticon* (Mansvelt 1884) was published to give an account of the differences between early Afrikaans and Dutch. The dictionary endeavoured to convince a Dutch user group that Afrikaans had developed as a fully-fledged language. The outer texts of this dictionary are in Dutch and the central list displays an Afrikaans macrostructure but the treatment is in Dutch, e.g.

> **1. penwortel,** ... de hoofdwortel van sommige boomen, die loodrecht naar bene-
> den gaat.

The treatment allocated to this lemma is the typical treatment to be found in a monolingual descriptive dictionary. One language is presented as the source language and the treatment is directed at that language although the treatment is done in another language. Is this a monolingual or a bilingual dictionary? It is a monolingual dictionary.

The occurrence of translation equivalents in the comment on semantics of a dictionary article does not necessarily elevate that dictionary to the level of being classified as bilingual; cf. the following examples from *Nuwe woordeboek sonder grense:*

> 2. **aanval** ■ werkwoord **(het aangeval)**
>   Dit beteken om iemand of iets seer te maak of skade aan te doen.
>   ➔ *Ons gaan die vyand met knopkieries en skilde aanval [attack].*

➜ *My vriendin is deur 'n hond aangeval [attack]; gevolglik moes sy hospitaal toe gaan.* ➜ *Diere soos ape <u>val</u> mense in motors <u>aan</u> [attack] omdat hulle kos wil hê.*
   ■ naamwoord **(aanvalle, aanvalletjie)**
'n Aanval is die aksie om iemand of iets seer te maak of skade aan te doen.
➜ *Tydens die aanval [the attack] is baie geboue en huise verwoes.*
➜ *In 'n aanval [an attack] op die bank is vyf wagte gewond.*

3. **aar** naamwoord **(are, aartjie)**
   ❶ 'n Aar is die pypie waardeur bloed na die hart vervoer word.
   ➜ *Die dokter het die inspuitnaald in 'n aar [vein] in my arm gedruk.*
   ❷ Dit is die deel van byvoorbeeld die koringplant waaraan die korrels vas is.
   ➜ *Tydens die oesseisoen word die are van die koring [ears of corn] gepluk.*

This is a dictionary compiled for a specified group of primary school learners with Afrikaans as third or fourth language, to be used in their Afrikaans curriculum. The dictionary should give them an indication of the meaning of a limited number of Afrikaans words relevant to their learning endeavour of Afrikaans as a foreign language. Although the target users of this dictionary primarily have to use this dictionary in their learning of Afrikaans their command of English is most probably better than their command of Afrikaans. Important from a lexicographic perspective is that they get access to the lemma sign representing the Afrikaans word they have encountered in their texts. This can be done by means of the alphabetically ordered macrostructure of the dictionary. However, it is just as important that they get access to an Afrikaans equivalent needed to produce a text in Afrikaans. If they do not know the Afrikaans word the explanation of meaning is of little relevance to them and there is no way they will find the needed word if they cannot do it via a language with which they are more familiar. Consequently the dictionary plan makes provision for access via an alphabetically arranged English equivalent list presented as a back matter text. These equivalents are also given as entries in the comment on semantics of each article. Notwithstanding these English equivalents and the alphabetical equivalent register presented as a back matter text, this dictionary remains a monolingual dictionary — although in the words of Tarp a monolingual dictionary "with a bilingual dimension."

There are different interpretations of the terms *monolingual dictionary* and *bilingual dictionary.* When a distinction is made between monolingual and bilingual learners' dictionaries the reference to *monolingual dictionaries* typically bears on dictionaries where the source language items are treated in the source language. The source language treatment primarily focuses on a paraphrase of meaning and the presentation of cotext entries to illustrate the typical use as well as the syntactic features of the treatment unit. Bilingual learners' dictionaries are typically seen as dictionaries in which lexical items from two given partner languages are co-ordinated. The source language provides the lemmata or primary treatment units whereas the target language provides the translation equivalents.

What choices should the user have?

The availability of learners' dictionaries

It is a well-known fact that the planning of a learners' dictionary has to be preceded by a clear identification of the intended target user of that dictionary. This identification of the target user does not only imply a knowledge of the level of language proficiency or the data types needed by the user but also a knowledge on the side of the lexicographer with regard to the mother-tongue of the target user. Bilingual learners' dictionaries are earmarked for a specific language pair whereas monolingual dictionaries, specifically those without a bilingual dimension, are directed at a more general target user group interested in learning a given foreign language.

Some problems in this regard

Learning a new language confronts learners not only with the lexicon and the grammar of that language but learners constantly have to negotiate their knowledge of their mother tongue and that of other languages familiar to them. In the process of learning a new language, language distortion is a real issue. A speaker of Zulu who wants to learn Chinese will experience little influence of language distortion. Because these two languages are vastly different the Zulu learner of Chinese will not be tempted to convey the structures of Zulu to Chinese. But when a Zulu speaker wants to learn another Nguni language, e.g. Xhosa, the target language will easily be influenced by his/her knowledge of the mother tongue. Closely related languages display a much stronger occurrence of language distortion.

   German and Afrikaans are related languages. German has a word *durchbringen* and its Afrikaans equivalent is *deurbring*. As is the case with the German word *durchbringen* the Afrikaans word *deurbring* is also a polysemous lexical item. These words share senses like "to get something through" and "to waste money". The Afrikaans word *deurbring* is also used where German would use *verbringen*: *die vakansie deurbring* "to spend the holiday". The German–Afrikaans word list of a bilingual dictionary treating these two languages will present the Afrikaans word *deurbring* as a translation equivalent of the German lemma sign *durchbringen*. This would not be an incorrect equivalent but although they share some senses the Afrikaans form has a wider polysemous range and the lexicographer may not rely on the intuition of the dictionary user to know the differences. For all the senses of the lexical item represented by the lemma sign *durchbringen* the Afrikaans form *deurbring* could probably be used as an equivalent. A mere co-ordination of the forms *durchbringen* and *deurbring* will be of little assistance for the user in terms of needing to be familiar not only with the similarities between the source and target language forms but also with the differences. A listing of illustrative examples may help to indicate

the mutual senses and usage situations but merely omitting an example illustrating a specific sense of the target language form represented by the translation equivalent which does not form part of the polysemous paradigm of the source language item represented by the lemma sign can be regarded as too implicit a way of conveying the more restricted polysemous range of the source language form. Users do not consult the entries in the search zone allocated to illustrative examples to ascertain the polysemous spectrum of a given source language form. Consequently it is very important that the comment on semantics of the article of a lemma sign like the German *durchbringen* should be divided into a subcomment on semantics for each of the senses of this lexical item — even if it implies the repetition of the same target language form as translation equivalent in each subcomment on semantics.

A monolingual German learners' dictionary not directed at target users from a specific speech community will not focus on the problems regarding *durchbringen* and *deurbring*. A dictionary aimed at speakers of Afrikaans who need to learn German will have to focus on this issue. When planning and using a monolingual learners' dictionary one has to ascertain whether the lexicographer has compiled the dictionary with a specific language group in mind because it is bound to influence the presentation and treatment of the dictionary.

Closely related languages also confront the lexicographer with problems with regard to partial false friends.

From a lexicographic perspective false friends do not really pose much of a problem for the lexicographer. A bigger problem is posed by partial false friends, i.e. the occurrence of formal identity or close identity and on a semantic level there is a relation although not one of identity but rather of relatedness. German and Afrikaans have the words *die See/see* and *das Meer/meer*. The German *die See* has the Afrikaans *meer* as translation equivalent and the German *das Meer* has the Afrikaans form *see* as translation equivalent. Both Dutch and Afrikaans have the adverb *amper* but they have direct opposite meanings: the Dutch *amper* means "almost not" whereas the Afrikaans *amper* means "almost". The fact that these lexical pairs can often occur in similar usage environments adds to the confusion potential they have for the user. However, within the comment on semantics the translation equivalents should give enough guidance to make the user aware of these pitfalls.

The relevant issues to be accounted for by the data exposure structure in a bilingual dictionary dealing with closely related languages are not always on the level of the entries participating in the lemmatic addressing procedures but often resides in the combinatory possibilities of the word represented by the lemma sign or the translation equivalent. In the treatment of the Dutch lemma sign *postzegel* (postage stamp) the Afrikaans translation equivalent *posseël* should occur as one of the microstructural entries. This lemma and translation equivalent display systematic similarities and this pattern continues in combinations like *postzegels verzamelen* x *posseëls versamel* (to collect postage stamps).

A typical usage of the Dutch word is in a combination like the following: *een postzegel op een brief doen* (to put a stamp on a letter). This is not translated into Afrikaans as *'n posseël op 'n brief doen* but as *'n posseël op 'n brief plak.* For the user of a Dutch–Afrikaans dictionary this difference is relevant and the dictionary should give account of it. The lemma and the translation equivalent(s) occur in a position of salience in an article. In a dictionary dealing with two closely related languages, the entries which could assist the user with real usage problems are often not the lemma and the translation equivalent(s) and are therefore not presented in that position of salience. Data exposure procedures are needed to draw the focus to these entries (cf. Gouws Forthcoming).

A bilingual dictionary should concentrate on those aspects that are contrastively relevant to the intended target user.

In the treatment of closely related languages, a lexicographer should be aware of the varied functions of the dictionary and the fact that the typical user could have different needs compared to the user of a bilingual dictionary dealing with unrelated languages. The emphasis in articles should not only be on procedures of lemmatic addressing but rather on the treatment of source language usage. In the article of a lemma sign like the Dutch *postzegel*, the lemma and the translation equivalent do have a role to play — in many usage situations the translation equivalent will primarily be looked up to confirm the proper form, and confirmation of a given form is a real part of text reception (cf. Gouws and Tarp 2004). However, the typical user will benefit more from the inclusion of the Dutch phrase *een postzegel op een brief doen (*to put a stamp on a letter) and its Afrikaans translation *'n posseël op 'n brief plak.* The presentation of illustrative material should not only be carefully selected but the lexicographer should employ the data exposure structure of the dictionary to make the user aware of relevant contrasting forms or, where applicable, of the relevant similarities. Similarities could be treated as unmarked entries but contrasting forms should be allocated an additional entry to expose the nature of this data.

Typological variations

In a language with a spectrum of dictionary types and a variety of learners' dictionaries, the user in need of a learners' dictionary typically has to decide whether he/she needs a monolingual or a bilingual dictionary. It has been indicated (cf. Ilson 1985), that many language learners opt for a bilingual dictionary in which their mother-tongue is co-ordinated with the target language. The dictionary consultation process is then either to find a word in the foreign language for a given item in the mother tongue or to find the proper word in the mother tongue for a word encountered in the foreign language.

From a language didactic perspective it has been argued that it is better for the user to consult a monolingual dictionary because having to go through the paraphrase of meaning and the example sentences in the foreign language will

assist the user to become more familiar with the foreign language. Once again it is important to take cognisance of the level of language proficiency of the intended target user. If the user is someone reasonably familiar with the foreign language the use of a monolingual dictionary could be an ideal solution. If a learner of say Sepedi is confronted with the Sepedi word *reka* he should consult a monolingual Sepedi learners' dictionary to find the lemma sign *reka* and the paraphrase of meaning in which the word is defined in Sepedi. If that learner already has the ability to understand the definition given in the monolingual dictionary then he/she can benefit from the monolingual dictionary by finding not only the meaning of the word but also some additional assistance. From a didactic perspective it may be true that the user will benefit more from a foreign language monolingual dictionary but from a user's perspective it is not always necessarily the case. Quite often one could argue in favour of the contrary. Once again the typical target user of the dictionary plays an important role.

If an adult English-speaking learner of Sepedi is confronted with the word *reka* the main purpose of his/her dictionary consultation process will not be to retrieve information regarding the activity to which the word *reka* refers. Although a definition of the word will be beneficial in terms of being confronted with a well-formed Sepedi sentence this may not be what the user really needs. The user may be familiar with the concept "to buy" but unfamiliar with the fact that *reka* is the Sepedi equivalent of the English verb *to buy*. Using the Sepedi–English component of a bilingual dictionary with English and Sepedi as partner languages may result in a much more rapid comprehension of the Sepedi word by having it co-ordinated with the English equivalent *to buy*.

The examples from *Nuwe woordeboek sonder grense* illustrate a similar aspect. The non-Afrikaans user of this dictionary may be familiar with the English word *vein* as well as with the meaning and use of this word. However, the person is not familiar with its Afrikaans equivalent *aar*. This user needs a dictionary giving him/her access to the Afrikaans form but it has to go via the English form. A monoaccessible Afrikaans descriptive learners' dictionary will be to little avail in such an endeavour.

Yet another possibility exists for users with a certain level of language proficiency. The lexicographer may decide to compile an explanatory dictionary with the lemma list in the foreign language but the explanations of meaning given in the mother tongue of the intended target user.

Typology versus functions

From the above-mentioned issues it should be clear that both monolingual and bilingual dictionaries could be used for different functions. Looking at the type of dictionary needed for a specific consultation procedure the focus should not be on bilingual or monolingual but rather on text production or text reception. We need a much stronger function-based dictionary classification.

In conclusion

Existing methods of dictionary classification do not always reflect the role of that dictionary or the ways in which it can assist users. General and blunt typological classifications should be avoided in favour of a stronger function-based approach.

Bibliography

Gouws, R.H. 2001. Lexicographic Training: Approaches and Topics. Emejulu, James Duplessis. (Ed.). 2001. *Éléments de lexicographie gabonaise. Tome I*: 58-94. New York: Jimnacs-Hillman.

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

Gouws, R.H. Forthcoming. Issues Regarding the Comment on Semantics in Bilingual Dictionaries Dealing with Closely Related Languages. Petkov, P. (Ed.). Forthcoming *Proceedings: Internationales Kolloquium zur Wörterbuchforschung.*

Gouws, R.H., M. Stark and L. Gouws. 2004. *Nuwe woordeboek sonder grense.* Cape Town: Maskew Miller Longman.

Gouws, R.H. and S. Tarp. 2004. Wie leer wat uit Afrikaanse (aan)leerderwoordeboeke. *Tydskrif vir Geesteswetenskappe.*

Hartmann, R.R.K. 1989. Sociology of the Dictionary User: Hypothesis 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.

Ilson, Robert (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*. Oxford: Pergamon Press.

Mansvelt, N. 1884. *Proeve van een Kaapsch-Hollandsch Idioticon.* Cape Town: Van de Sandt de Villiers.

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

Wiegand, H.E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband.* Berlin/New York: De Gruyter.

Wiegand, H.E. 2001. Was eigentlich sind Wörterbuchfunktionen? Kritische Anmerkungen zur neueren und neusten Wörterbuchforschung. *Lexicographica* 17: 217-248.

The Access Structure in Learner's Dictionaries*

Mariza Steyn, *Unit for Afrikaans, Language Centre, University of Stellenbosch, Stellenbosch, Republic of South Africa (ms@sun.ac.za)*

Abstract: Dictionaries have to be compiled in accordance with specific target users and dictionary goals. User orientation is, however, not the only criterion for measuring the success of a dictionary. Present-day lexicography requires the evaluation of the accessibility of dictionaries. The ease with which macro- and microstructural information is located and retrieved determines the quality and accessibility of the presentation. User-orientated, accessible learner's dictionaries can contribute substantially to the improvement of communication in a multilingual South Africa with different groups of language learners.

Metalexicography makes provision for guide structures that facilitate access to information in the dictionary. The outer, inner and meta-access structure are functional structure elements that increase the retrievability of information and help the user to reach, as swiftly as possible, the information he/she is looking for. During the planning of a learner's dictionary, the lexicographer must draw up, among others, a dictionary plan in which a detailed explanation of the outer and inner access structure is given. Such a dictionary plan contributes to a consistent application of lexicographical decisions. Decisions regarding the access structure must also be explained and illustrated in the front text in order to enable learners to acquire dictionary skills and become effective dictionary users.

Keywords: METALEXICOGRAPHY, OUTER ACCESS STRUCTURE, INNER ACCESS STRUCTURE, LEARNER'S DICTIONARY, FRAME STRUCTURE, UNINTEGRATED, INTEGRATED, MONO- AND POLY-ACCESSIBLE, OUTER TEXTS, OUTER RAPID ACCESS, DATA DISTRIBUTION STRUCTURE, MICRO-ARCHITECTURE, DATA-ARCHITECTURE, DATA EXPOSURE STRUCTURE, META-ACCESS

Opsomming: Die toegangstruktuur in aanleerderwoordeboeke. Woordeboeke moet saamgestel word met inagneming van 'n spesifieke teikengroep en woordeboekdoel. Gebruikersgerigtheid is egter nie die enigste kriterium waaraan die sukses van 'n woordeboek gemeet word nie. Die hedendaagse leksikografie vereis die beoordeling van die toeganklikheid van 'n woordeboek. Die gemak waarmee makro- en mikrostrukturele inligting opgespoor en ontsluit word, bepaal die kwaliteit en die toeganklikheid van die aanbieding. Gebruikersgerigte toeganklike aanleerderwoordeboeke kan 'n wesenlike bydrae lewer tot die verbetering van kommunikasie in 'n veeltalige Suid-Afrika met verskillende groepe taalaanleerders.

Die metaleksikografie maak voorsiening vir gidsstrukture wat toegang tot inligting in die woordeboek vergemaklik. Die eksterne, interne en metatoegangstrukture is funksionele struktuur-elemente wat die herwinbaarheid van inligting verhoog en die gebruiker help om so vinnig moont-lik uit te kom by die verlangde inligting wat hy/sy soek. Die leksikograaf moet tydens die beplan-ning van 'n aanleerderwoordeboek 'n woordeboekplan opstel waarin onder andere die eksterne en interne toegangstruktuur uiteengesit word. So 'n woordeboekplan dra by tot 'n konsekwente toe-passing van leksikografiese besluite. Besluite oor die toegangstruktuur moet ook in die toeligting beskryf en geïllustreer word om aanleerders in staat te stel om woordeboekontsluitingsvaardig-hede te verwerf en effektiewe woordeboekgebruikers te word.

Sleutelwoorde: METALEKSIKOGRAFIE, EKSTERNE TOEGANGSTRUKTUUR, INTERNE TOEGANGSTRUKTUUR, AANLEERDERWOORDEBOEK, RAAMSTRUKTUUR, GEÏNTEGREER, NIE-GEÏNTEGREER, MONO- EN POLI-TOEGANKLIK, BUITETEKSTE, EKSTERNE KITS-TOEGANG, DATAVERSPREIDINGSTRUKTUUR, MIKRO-ARGITEKTUUR, DATA-ARGITEK-TUUR, DATABLOOTLEGGINGSTRUKTUUR, METATOEGANG

1.   Introduction

Recent research in the field of metalexicography has focused on the structure of dictionaries. The following six main structures have been identified by Haus-mann and Wiegand (1989): macro-, micro-, distribution, frame, reference and access structure. The access structure can be regarded as the search route the dictionary user follows during a dictionary consultation procedure. Bergen-holtz and Tarp (1995: 16) define it as the structure of lexicographical indicators directing the user to the information required. A distinction is made between the outer and inner access structure. The outer access structure leads a user up to the lemma sign, introducing the article from which the needed information is to be retrieved, and the inner access structure guides the user within a dic-tionary article to the search zone in which the relevant data is presented.

The objective of the first part of this article is to give an overview of the elements which comprise the access structure of a dictionary. These elements of the outer and inner access structure will be illustrated and discussed briefly. Only some aspects of the outer access structure will be dealt with in more depth. Although the focus will be on monolingual learner's dictionaries, exam-ples from bilingual learner's dictionaries will also be included.

2.   What is the correlation between the access structure and the dictionary consultation procedure?

The access structure is the search route the user follows during a consultation procedure and it correlates directly with the actions taking place during the consultation process. The consultation process therefore affects the access structure of a dictionary. According to Béjoint (1994: 155), the following actions take place when the user consults a dictionary:

1.   Choose the correct dictionary according to the problem.

2.   Identify the problem area, word or phrase.

3.   Determine the most likely place in the dictionary where the word, phrase is treated.

4.   Guess which form of the word is listed as lemma.

5.   Guess where the word is placed in the alphabetical order of the dictionary. Here capital letters, geographical words, abbreviations etc. play a role.

6.   Find the word in its place as entry or subentry.

7.   Interpret the information given in the entry.

Tono (1992: 238) gives a schematic illustration of the consultation procedure:
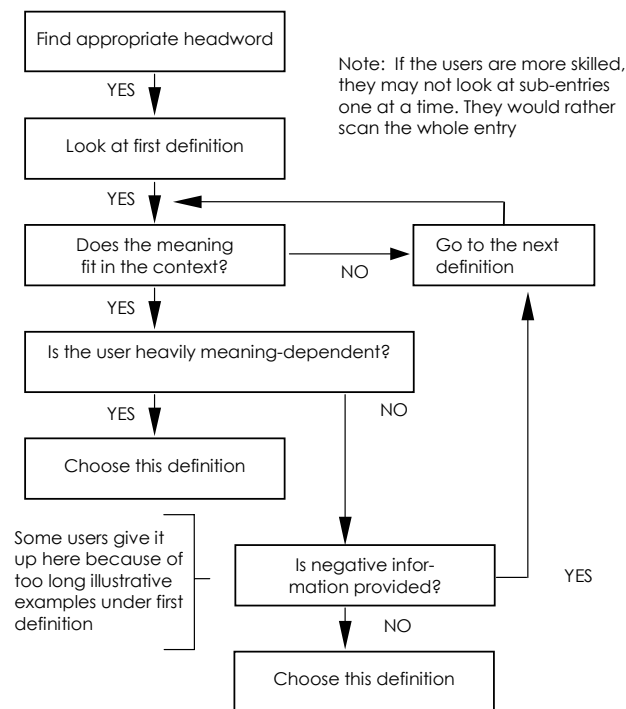
It is evident from both examples that the consultation activities consist of two stages, which correlate directly with the two main structures of the dictionary, namely the macro- and microstructure. By implementing elements of the outer

access structure, the user can find information in the macrostructure, while the inner access structure leads the user to specific information in the microstructure. The first part of the dictionary consultation procedure where the user follows the outer search path to the desired lemma will now be considered.

3.   The outer access structure of a dictionary

The outer access structure starts on the cover of the dictionary. Almind and Bergenholtz (2000) have made a detailed study of the aesthetic dimension of the dictionary where the outer access structure includes, for instance, the type and amount of information, as well as the presentation on the back and front pages of the dictionary.

3.1   The frame structure

Once the relevant dictionary has been found, the next step is to find the right component in the dictionary. This brings us to the frame structure. Accessibility of data is not only determined by the use of a well-defined access structure, leading the user to a venue within the word list structure of the dictionary, but it is also determined by a functional positioning of the data within the word book structure of the dictionary, by employing different texts in the dictionary to explain and explicate the contents of the dictionary. Utilising a word book approach, a dictionary can display a so-called frame structure (Kammerer and Wiegand 1998: 230, 233). Any dictionary is regarded as a carrier of different texts and text types. The frame structure makes provision for three textual focal points, i.e. the front matter, the central list and the back matter.

The central list is the textual component which is the most typical venue for dictionary consultation procedures. The structure of the central list should therefore facilitate the quick and unimpeded access to the required data. The outer access structure has to guide the user to a specific article stretch within the central list and within the article stretch to a specific lemma sign.

The front matter section contains all the texts preceding the central list and the back matter section contains all the texts following the central list. The texts in the front and back matter of a specific dictionary constitute the outer texts of that dictionary. According to Gouws (2001: 103), neither the front nor the back matter are functional parts of a dictionary, although both may contain individual texts, for instance the users' guidelines, which have specific functions in the dictionary. Traditional dictionaries have often used the front and back matter to present outer texts. However, in many dictionaries these were not functional texts and contributed very little to the genuine purpose of the dictionary. Outer texts included in the front and back matter should help the user to gain access to the required data as swiftly as possible. The user should furthermore be assisted in accessing information included in the outer texts.

3.2   Unintegrated and integrated outer texts

According to Wiegand (1998: 299), every dictionary has a genuine purpose. Outer texts can be classified according to their contribution to the achievement of the genuine purpose of the dictionary. An unintegrated outer text is not integrated into the genuine purpose of the dictionary. However, these unintegrated outer texts can "also serve a purpose as functional texts and they have their own genuine purpose, not identical to that of the dictionary but also important to ensure the eventual successful use of the dictionary" (Gouws 2002: 474). Examples of unintegrated outer texts are the title page, introduction, foreword, etc.

Another example of an unintegrated outer text with its own genuine purpose is the *Language Awareness Notes* in *Macmillan English Dictionary for Advanced Learners*. This additional text consists of eleven subtexts and is inserted into the central word list. It functions as a twice partially extended complex outer text. The *Language Awareness Notes* is preceded by two subtexts: an explanatory note and a table of contents. These secondary outer texts give immediate access to this unintegrated text and encourage the user and teacher to use the language tools included in the main text.

**LANGUAGE
AWARENESS**

These pages are designed to give up-to-date information on topics that are relevant to everyone who has any interest in the way the English language is being used at the beginning of the 21st century. They have all been written by people who are experts in the field they are writing about here; indeed, most of the authors have published more detailed work, which you might like to read if you find these pages interesting. They are intended to give you a deeper understanding of such areas as Spoken Discourse, Metaphor, and Pragmatics, all of which are crucial for users of the English language. Our aim is to encourage you to think about the language, and to realize what an exciting tool it is when you can use it with confidence. There is a lot of food for thought on these pages, and teachers will find there is plenty of scope in these topics for classroom activities and discussion.

Contents

Numbers LA2
Phrasal verbs LA3
Academic writing LA4
Metaphor LA8
Computer words LA10
Pragmatics LA12
Spoken discourse LA14
Sensitivity: avoiding offence LA16
British and American English LA18
Business English LA20
Word formation LA22

Example 1: Two subtexts preceding an unintegrated text

Integrated outer texts are directed at achieving the genuine purpose of the dictionary. The inclusion of these texts in learner's dictionaries should be expanded because of the added value and contribution to the successful dictionary consultation process. According to Gouws (2002: 473), there are two types

of integrated texts: "These texts belong either to the class of so-called 'help texts' with the function to assist the user with a more successful dictionary consultation procedure or, in as far as they contain a part of the lexicographic data presented in the dictionary, they are integrated texts from which the user can retrieve information regarding the subject matter of the dictionary."

The users' guidelines text is an example of a help text that assists the user in order to have a successful dictionary consultation procedure. By consulting the users' guidelines, complemented by a sample article, the learner gains insight into the structure and contents of the dictionary.

In order to complement the data presented in the central list a part of the lexicographic data can be accommodated in lists, traditionally presented in the back matter. Grammar codes presented in *Longman Dictionary of Contemporary English* (LDOCE) are integrated into the distribution of lexicographic data and assist the user in retrieving information regarding the subject matter of the dictionary. The same dictionary also includes a list of geographical names with adjectives forming part of the macrostructure of the dictionary. By adding this extra list with grouped and related words, the lexicographers support information retrieval and enhance language learning.

3.3   Mono- and poly-accessible dictionaries

In most general monolingual dictionaries, the central list is the only outer access structure; therefore the macrostructure of these dictionaries coincides with the outer access structure. These dictionaries, classified according to their outer access profile, are mono-accessible dictionaries with only one defined search path to solve the search problem. A learner who consults the mono-accessible *Tweetalige Aanleerderswoordeboek/Bilingual Learner's Dictionary* (TAW), for the translation of "mountain" uses the following outer search path: he/she finds the letter *M* and the relevant article stretch, uses the guide words and scans through the partial article stretch for the lemma sign.

Some dictionaries have additional wordlists and/or indexes which form part of the comprehensive macrostructure. In these poly-accessible dictionaries the user may arrive at lexicographical data by using more than one defined search path. Poly-accessible dictionaries only have one macrostructure but more than one access structure. The following example illustrates the different search paths available to the user:

A learner using the *Oxford Advanced Learner's Dictionary* (OALD) has two options in solving the search problem in reaching "mountain". Firstly, the learner can scan through the central list until after "mount²" and before "mountainside". Another option is to use *Appendix 10, Defining Vocabulary* to identify the lemma ("mountain") and then move along the outer search path to the lemma in the central list. The learner can therefore first confirm whether a specific word is listed and used without having to scan through multiple article

stretches. This word list also gives an overview of the macrostructural scope of the dictionary.

Another example of an outer text creating a poly-accessible dictionary is the extra wordlist provided in *Collins COBUILD Learner's Dictionary* (CO-BUILD) as a text in the back matter. In this outer text, the usage frequency of the lemmata is indicated and as an extra access structure adds value for the beginner who uses/learns the most frequent words in English. The list of geographical names in the same dictionary does not, however, make the dictionary poly-accessible, since the names are not listed in the central list. It is an unintegrated text aimed at the systematizing of information to enhance knowledge transfer.

The German learner's dictionary *Wörterbuch Deutsch als Fremdsprache* (WDAF) uses outer texts to rectify the unnatural alphabetical placing of, in Bolinger's words (1985: 69) "carrots, onions, beetroot and salsify next to one another". The lexicographers provide lists with words belonging to the same semantic fields. The learner can, for instance, find a grouping of words relating to physical processes. These functional texts presented in the outer text not only place words in their specific context, but also create a poly-accessible dictionary, giving the learner accessible search path options. The poly-accessibility of this dictionary also functions in another direction. Many articles contain reference entries with sections in the outer text (*Wortfelder*) as reference addresses. The alphabetical ordering of the central list is complemented by the semasiological ordering in the outer text (Gouws 2002: 480-481).

Wortfelder I.: Der Mensch

1. Körperteile/Organe (↗ *auch* 4, 5, 6)
   1.1. Substantive
   Kopf, Haupt, Schädel, Gehirn, Hirn, **Haar,** Locke, Pupille, **Auge,** Augenbraue, Wimper, Lid, Schläfe, Stirn, **Gesicht,** Fresse, Antlitz, Backe, Wange, Nase, Ohr, Ohrmuschel, **Mund,** Maul, Schnabel, Zunge, Lippe, Rachen, Gaumen, Schlund, Gurgel, **Zahn,** Backenzahn, Kiefer, Kinn,
   Hals, Kehlkopf, Speiseröhre, Luftröhre, Zäpfchen, Nacken, Genick, Adamsapfel, Mandel,
   Rumpf, Schulter, Rücken, Buckel, Brust, Brustwarze, Rippe, Wirbelsäule, Leib, Lende, Hüfte, Becken, Kreuz, Achsel, **Bauch,** Magen, Darm, Herz, Lunge, Leber, Niere, Milz, **Unterleib,** Gebärmutter, Scheide; Hoden, Penis, Glied, Sperma; **Gesäß,** Hintern, After, Arsch,
   Arm, Oberarm, Unterarm, Ellenbogen,
   Hand, Handgelenk, Flosse, Pfote, **Finger,** Fingernagel, Fingerspitze, Daumen, Zeigefinger, Mittelfinger, Ringfinger,
   Bein, Oberschenkel, Unterschenkel, Knöchel, Knie,
   Fuß, Fußsohle, Sohle, Spann, Hacke(n), Ferse, Zehe,
   Ferner: Blut, (Blut)kreislauf, Knochen, Nerv, Gelenk, Sehne, Ader, Aorta, Gefäß, Blutgefäß, Arterie, Schlagader, Haut, Schwiele

1.2. Verben
bluten, verrenken, ausrenken, einrenken, verstau-
chen, verschlucken, schlucken, rülpsen, knien
1.3. Adjektive/Adverbien
blutig, knochig, sehnig, schwielig

Example 2: Semantic fields in the back matter of WDAF

In order to help the learner to organize and memorize words, the same diction-
ary gives tables in which grammatical information is grouped together. Words
which belong to the same semantic, morphological and syntactic groups are
presented together to give the learner an image of the language. These texts
create other access possibilities for the user and contribute to a user-friendly
dictionary.

Tafel V

Liste der im Wörterbuch dargestellten unregelmäßigen Verben

backen	**bäckt**	**buk**	**gebacken**
befehlen	**befiehlt**	**befahl**	**befohlen**
beginnen		**begann**	**begonnen**
beißen		**biss**	**gebissen**
bergen	**birgt**	**barg**	**geborgen**
bersten	**birst**	**barst**	**geborsten**
bewegen		**bewog**	**bewogen**
biegen		**bog**	**gebogen**
bieten		**bot**	**geboten**
binden		**band**	**gebunden**
bitten		**bat**	**gebeten**
blasen	**bläst**	**blies**	**geblasen**
bleiben		**blieb**	**geblieben**
bleichen		**blich**	**geblichen**
braten	**brät**	**briet**	**gebraten**
brechen	**bricht**	**brach**	**gebrochen**
brennen		**brannte**	**gebrannt**
bringen		**brachte**	**gebracht**
denken		**dachte**	**gedacht**
dreschen	**drischt**	**drosch**	**gedroschen**
dringen		**drang**	**gedrungen**
dürfen	**darf**	**durfte**	**gedurft**
empfangen	**empfängt**	**empfing**	**empfangen**
empfehlen	**empfiehlt**	**empfahl**	**empfohlen**

Example 3: List in the back matter of WDAF

All these examples illustrate the poly-accessibility of dictionaries in terms of
the word list structure, thus an enhanced access to the central list.

3.4   Primary and secondary outer texts

According to Hausmann and Wiegand (1989: 331), each dictionary should con-
tain two compulsory texts, i.e. the central list, which presents the lemmata as
macrostructural entries and primary treatment units, and an outer text, usually
presented in the front matter, with the users' guidelines. This text already has
an access function because it not only informs users of the data collection in the

dictionary, but also explicates the positioning and arrangement of data types. This allows the user a more rapid dictionary consultation procedure. A typical component of such a users' guidelines text is a sample article taken from the central list to illustrate the default article structure, micro-architecture and search zones.

The users' guidelines function as an immediate constituent of the dictionary as a whole. The front matter text presenting the table of contents of the dictionary also falls in this category. Both can be regarded as first order primary outer texts. The table of contents gives access to most of the elements of the word book structure and occasionally to parts of these elements. Where the typical outer access structure guides a user to a lemma sign, the table of contents includes entries which guide the user to the different components of the frame structure and to the individual texts contained in these components of the dictionary as a big text. The functional value of a table of contents, which increases the accessibility of the dictionary, may never be underestimated.

OALD and COBUILD both include tables of contents which give access to other lists and information in the front and back matter. Once again the learner can follow two search paths, i.e. the direct route to the required information or rapid access by means of the table of contents. The lexicographer must also ensure accessibility to the outer texts.

The users' guidelines of some learner's dictionaries focus only on explicating the data presented in the central word list. It therefore functions not as an immediate constituent of the dictionary as a whole, but as an immediate constituent of the extended central list. Such an outer text functions as a second order primary outer text. Another example of an outer text which is a constituent only of the central word list is the list of defining vocabulary. It increases the access possibilities and focuses on data presented in the central list text.

Many dictionaries contain primary outer texts with their own outer texts, e.g. an explanatory text, a table of contents or a users' guidelines text directed not at the dictionary as such but only at a specific outer text. These texts are complex outer texts that are extended by adding subtexts. The subtexts function as secondary outer texts and are not immediate constituents of the dictionary as a big text but rather of the specific outer texts. A major function of secondary outer texts is to provide access to their primary outer texts and to increase accessibility to data accommodated outside the central list of a dictionary.

In WDAF a secondary text refines access to the texts in the back matter of the dictionary. The lexicographers provide a classification of all the semantic fields listed in the back matter. This secondary text places all the semantic fields in context and facilitates the search and learning processes. A table of contents with page references gives quick access to texts in the back matter. This functional secondary text increases the accessibility and retrievability of information and is a good example of special extended access to an outer text.

Tafeln
(Übersicht)

Example 4: Access to outer texts

LDOCE bridges the traditional bias directed at the central word list and employs an outer text in the front matter to familiarise users with the structure of the users' guidelines. This secondary text is an attempt to enhance the access to the compulsory and important text that explains the system of the dictionary. It forms part of a more sophisticated access structure which is the key to a successful lexicographic product.

●
●
●
● **Guide to the Dictionary**

xiv

…………………………………………………………………………………………………

Contents

……………………………………………………

1   How to find what you are looking for

  1.1   Compound words

  1.2   Phrasal verbs

  1.3   Derived words without definition

  1.4   Homographs

  1.5   Other types of headwords

  1.6   Phrases and idioms

2   Understanding meaning

  2,1   Words with more than one meaning

  2.2   Definitions

  2.3   Examples

  2.4   Collocations

  2.5   Finding the meaning you want – Signposts

  2.6   Long entries with menus

1   How to find the word you are looking for

………………………………………………………

Words are listed in this dictionary in alphabetical order

1.1   Compound words

………………………………………………………

Compound words are groups of two or more words with a fixed form and a special meaning, such as **front man** and **front line**. Most of these are shown as full headwords (but see section 1.6 Phrases and idioms). They are treated like ordinary words in the alphabetical order; the space or hyphen between the two parts is ignored.

front·al
frontal sys·tem
front-and-center
front bench
front·bench·er

2.7   Showing words with similar and opposite meanings

3   Frequency

4   Grammar

4.1   Word classes

4.2   Inflections

4.3   Syntax – verbs

4.4   Phrasal verbs

4.5   Syntax – nouns

4.6   Syntax – adjectives and adverbs

4.7   Very infrequent words

5   Information on register and usage

5.1   Indicating register

5.2   Spoken words and phrases

5.3   Usage notes

6   Pronunciation

6.1   Compound words

7   British and American English

7.1   Pronunciation and spelling differences

7.2   Words and meanings – British and American differences

7.3   Differences in grammar

7.4   Differences in phrases and collocations

front door
fron·tier

1.2   Phrasal verbs
....

Multi-word verbs, like **give up** or **put off**, are listed in alphabetical order directly after the entry for their main verb. For example:

face[2] *v* [T]
 face sb · · **down**
 face up to
 face sb **with**
face card

1.3   Derived words without definition
....

Some words do not need a definiton, because they are derived from a headword by adding a suffix. For example **gracefully** and **gracefulness** are derived from **graceful**, and their meaning is simply that of the main word plus the meaning of the suffix. These words are shown at the end of the entry for the word that they are derived from.

grace·ful /ˈɡreɪsfəl/ *adj* 1 moving in a smooth and attractive way, or having an attractive shape: *a slim graceful figure* 2 behaving in a polite and pleasant way: *a graceful apology* – **gracefully** *adv*: *When I am no longer needed, I shall retire gracefully*. – **gracefulness** *n* [U]
In this case, **gracefully** just means 'in a graceful way,' and **gracefulness** just means ' the quality of being graceful.'

Example 5: A secondary text in LDOCE

3.5   Alphabetical ordering, multiword units and collocations

The following three elements also form part of the outer access structure and have to be included in the dictionary planning phase. Firstly, the alphabetical ordering of the macrostructure facilitates access to lemmata since all users may be reasonably expected to master the alphabet. Opitz (1983: 172) states that the first and foremost of the required skills of the dictionary user is the ability to find and identify lexical items in a total word list. In practically all cultures and at almost all levels of activity, the alphabet as established by the language community in question, serves to forge a reliable link of communication between individuals and groups of different intellectual or occupational levels. The ordering of the macrostructure is, however, not so trivial because of a typical distinction between straight alphabetical ordering and an ordering presenting a sinuous lemma file.

In straight alphabetical ordering main lemmata are typically displayed whereas a sinuous lemma file presents sublemmata together with the main lemmata. The placing of sublemmata can create problems for users in terms of retrievability and accessibility. The system of the dictionary concerning sublemmata, like compounds and derivatives, has to be explained and illustrated in the front matter of the dictionary. Where learners with a restricted knowledge of a language are concerned, it would be best to restrict the inclusion of sublemmata to a minimum. If the lexicographer decides to include sublemmata

286   Mariza Steyn

because of a more advanced language knowledge of the users or because of space considerations, they must be marked by non-typographical markers and treated consistently.

The following examples illustrate possible external access problems that users might experience because of an inaccessible placement of sublemmata.

Luft [lʊft]**, die**; ~, ⟨o.Pl.⟩ geh. Lüfte [lʏftə] **1.** ⟨o.Pl.⟩ **1.1.** 'Gemisch aus bestimmten Gasen, bes. Sauerstoff, Stickstoff, das die Erde als Hülle umgibt': *warme, kalte, trockene, feuchte ~; die ~ ist mild; frische ~ ins Zimmer lassen; ~ (auf einen Reifen) (auf)pumpen* **1.2.** 'Luft (1.1), die zum Atmen dient': *~ ↗ holen; die ~ ('den Atem') anhalten; keine ~ bekommen* ('vorübergehend nicht od. nur sehr schwer atmen können') **1.3.** *an die frische ~* ('ins Freie') *gehen; ich gehe mal an die ~* **2.** 'der Raum der Luft (1.1) über der Oberfläche der Erde': *die Eingeschlossenen aus der ~* ('durch Flugzeuge') *mit Lebensmitteln versorgen; Fotos aus der ~ma-chen; die Arme, den Ball in die ~* ('in die Höhe') *werfen; der Vogel, der Ballon erhob sich in die ~/* geh. *Lüfte* **3.** 'Spielraum Zwischenraum': *lass noch etwas ~ zur Wand hin; nach der Konferenz habe ich wieder etwas ~* ('Zeit') *für andere Arbeiten* ❖ **lüften, luftig, Luftikus, Lüftung – Kaltluft, Zugluft;** vgl. **Luft/luft-**
• /jmd., etw./ **sich in ~ auflösen 1.** /etw., bes. Plan/ 'völlig fallen gelassen werden': *das Projekt hat sich in ~ aufgelöst* **2.** /etw., jmd./ 'plötzlich nicht mehr da sein, verschwunden sein': *das Haus, der Mann kann sich doch nicht in ~ aufgelöst haben!; irgendwo ist/herrscht dicke ~* 'irgendwo herrscht eine gespannte, gereizte Atmosphäre': *zu Hause ist heute dicke ~;* /etw. bes. Information/ **aus der ~ gegriffen sein** 'frei erfunden sein, nicht der Wahrheit entsprechen': *die Geschichte ist völlig aus der ~ gegriffen!;* /jmd./ **in die ~ gehen** 'die Nerven verlieren, seinen Zorn, seine Wut durch Äußerungen, Verhalten deutlich machen': *wenn er das erfährt, geht er in die ~; etw. liegt in der ~* ('etw. Drohendes, Erwartetes steht bevor'); /jmd./ **sich ~ machen** 'frei und im Affekt aussprechen, was einen bedrückt': *ich musste mir erst mal ~ machen;* umg. /jmd./ **etw. in die ~ jagen** ('etw., bes. ein Gebäude, durch Sprengen zerstören'); /jmd./ **für jmdn. ~ sein** 'von jmdm. demonstrativ nicht beachtet werden': *der is für mich ~!;* **die ~ is rein** ('für das Gelingen eines Vorhabens besteht keine Gefahr mehr'); /jmd./ **(frische) ~ schnappen** ('ins Freie gehen'); /jmd., Tier/ **nach ~ schnappen** ('rasch und mühsam Atem holen'); /jmd./ **jmdn. an die ~ setzen 1.** 'jmdn. entlassen': *nach diesem Vorfall wird er ihn sicher an die ~ setzen* **2.** 'jmdn. hinauswerfen (2)': *wenn der Streit nicht aufhört, setze ich euch beide an die ~!;* **jmdm. bleibt die ~ weg** ('jmd. ist fassungslos')

Luft/luft [ˈ..]**.]|-abwehr, die** 'Gesamtheit der Maßnahmen, Kräfte und Mittel zur Abwehr (1) gegnerischer Angriffe aus der Luft (2.1)': *die ~ alarmieren* ❖ ↗ wehren; **-ballon, der** 'meist aus Gummi bestehender Hohlkörper, der aufgeblasen wird und Kindern als Spielzeug dient'; SYN Ballon (2) (↗ BILD): *den ~ fliegen lassen* ❖ ↗ Ballon; **-dicht** ⟨Adj.; o. Steig.⟩ 'so fest verschlossen, dass keine Luft (1.1) eindringen od. entweichen kann' /auf Behältnisse bez./: *eine ~e Verpackung; etw. ~ verschließen; die Packung ist ~* ❖ ↗ dicht; **-druck, der** ⟨o.Pl.⟩ 'Druck (1) der atmosphärischen Luft (1.1)': *den ~ messen; der ~ schwankt, steigt, fällt* ❖ ↗ drücken

lüften [ˈlʏftn̩], lüftete, hat gelüftet **1.** /jmd./ *etw. ~* 'etw., bes. Kleidung, einen Raum, von frischer Luft (1) durchdringen lassen': *die Betten, das Zimmer, die Wohnung ~; du musst, solltest mal ~* ('frische Luft ins Zimmer lassen') **2.** /jmd./ *etw. ~* 'etw., das etw. bedeckt, verdeckt, kurz hochheben': *den Vorhang, Schleier ~; den Hut zum Gruß ~; den Deckel ~* ❖ ↗ **Luft**

Luft [lʊft..]**.]|-fahrzeug, das** 'Fahrzeug, das zur Fortbewegung in der Luft (2) dient' ❖ ↗ fahren; **-gewehr das** 'Gewehr, bei dem das Geschoss durch komprimierte Luft unter Druck aus dem Lauf (4) getrieben wird'; ↗ FELD V.6.1: *mit dem ~ auf Spatzen schießen* ❖ ↗ wehren; **-hoheit, die** 'Souveränität eines Staates über den Luftraum seines Hoheitsgebiets': *die ~ respektieren; die ~ wurde verletzt* ❖ ↗ hoch

luftig [lʊftɪç] ⟨Adj.⟩ **1.1.** ⟨Steig. reg.; nicht bei Vb.⟩ 'hell, geräumig und mit ständig durchziehender frischer Luft' /auf einen Raum bezogen/: *ein heller, ~er Raum; die Laube ist ~* **1.2.** ⟨o. Steig.; nur attr.⟩ *in ~er* ('von frischer, freier Luft umgebener') *Höhe: sie schwebten mit dem Ballon in ~er Höhe* **2.** ⟨Steig. reg.⟩ 'leicht (1.2) und so zugeschnitten, dass große Teile des Körpers nicht bedeckt sind' /auf Kleidung bez./: *ein ~es Sommerkleid; sie ist ~ gekleidet* ❖ ↗ **Luft**

Luftikus [ˈlʊftikʊs]**, der;** ~/~ses, ~se scherzh. 'leichtsinniger, oberflächlicher Mensch': *so ein ~!; er is ein ~* ❖ ↗ **Luft**

Luft/luft [lʊft..]**.]|-kissenfahrzeug** [kɪsn..]**, das** 'Fahrzeug, dessen unterer Teil, ähnlich einem Kissen, mit komprimierter Luft gefüllt ist und sich dicht über dem Erdboden oder der Oberfläche des Wassers schnell fortbewegt' ❖ ↗ Kissen, ↗ fahren; **-lande-truppe** [landə..]**, die** 'Truppe, die mit Flugzeugen (hinter der gegnerischen Front) abgesetzt wird, um am Boden zu kämpfen' ❖ ↗ landen, ↗ Truppe; **-leer** ⟨Adj.; o. Steig.; vorw. attr.⟩ 'keine Luft enthaltend, ein Vakuum bildend': *ein ~er Raum; ein ~es Gefäß; etw. ~ pumpen;* METAPH *ein Politiker kann nicht im ~en Raum schweben* ('kann nicht ohne enge Verbindung mit der Realität leben') ❖ ↗ leer; **-linie, die** ⟨o.Pl.⟩ 'kürzeste Entfernung zwischen zwei Punkten auf der Oberfläche der Erde': *die ~ zwischen Berlin und Paris beträgt nur etwa 900 km* ❖ ↗ Linie; **-matratze, die** 'großes Kissen von etwa zwei Meter Länge, das aufgeblasen wird und zum Liegen, Schlafen verwendet wird': *die ~ aufblasen; sich auf die ~ legen* ❖ ↗ Matratze; **-post die** 'Beförderung der Post, Fracht durch Flugzeuge': *einen Brief, ein Paket per ~ absenden, befördern* ❖ ↗ Post; **-pumpe, die** 'Gerät, mit dem Luft od. andere Gase in einen od. aus einem Hohlraum gepumpt wird' (↗ TABL Fahrzeuge): *mit der ~ das Fahrrad aufpumpen;* ❖ ↗ Pumpe; **-raum, der** 'Raum über der Oberfläche der Erde, bes. als Gebiet der Lufthoheit': *den ~ verletzen, verlassen* ❖ ↗ Raum; **-röhre, die** 'röhrenförmige Verbindung zwischen Kehlkopf und Lunge, durch die die Luft beim Atmen in die Lunge gelangt'; ↗ FELD I.1.1: *von der Nahrung ist etwas in die ~ gelangt* ❖ ↗ Rohr; **-schiff, das** 'durch Motoren angetriebenes lenkbares, langsam fliegendes Luftfahrzeug, das aus einem großen, mit Gas gefüllten länglichen Behälter und daran befestigten Gondeln (2) zum Transport (1) von Passagieren und Fracht besteht'; ↗ FELD VIII.4.2 ❖ ↗ Schiff; **-schloss, das** 'etw. Wunderbares, etw., das man sich wünscht, das aber nur in der Phantasie existiert': *das ist nicht zu verwirklichen, sondern nur ein ~* • /jmd./ **Luftschlösser bauen** ('Pläne machen, die unrealistisch sind und sich nicht verwirklichen lassen'); **-streitkräfte, die** ⟨Pl.⟩ 'Teil der Streitkräfte, der militärische Operationen, Kampfhandlungen unter Einsatz von Luftfahrzeugen durchführt'; SYN Luftwaffe: *die ~ kommen zum Einsatz* ❖ ↗ streiten, ↗ Kraft

Lüftung [ˈlʏft..]**, die;** ~, ~en **1.1.** ⟨vorw. Sg.⟩ 'das Lüften (1)': *für gute ~ sorgen; die ~ im Zimmer ist ausreichend* **1.2.** 'Vorrichtung zum Lüften (1)': *eine moderne ~ einbauen; die ~ des Hauses funktioniert nicht richtig, ist defekt* ❖ ↗ **Luft**

Luft [ˈlʊft]**.]|-waffe, die** SYN 'Luftstreitkräfte' ❖ ↗ Waffe; **-zug, der** ⟨o.Pl.⟩ 'Bewegung der Luft, leichter Wind': *ein leichter, kühler ~; durch den ~ bewegten sich die Zweige, Blätter, Gardinen* ❖ ↗ ziehen

Example 6: *Wörterbuch Deutsch als Fremdsprache*

If a learner uses WDAF in order to find the meaning of the compound "Luftpumpe" he/she will experience problems with respect to the outer access structure. The consultation process starts at the lemma "Luft", continues to "Luft/luft" without success, while a third "Luft" and a fourth "Luft/luft" also do not provide the solution. Furthermore, the search route of the learner is interrupted by the alphabetical placing of "lüften", "lüftig" and "Luftikus". At long last the required sublemma is found in the fifth article. For a detailed discussion about the problems of nesting and multiple niching consult Gouws (2001: 106-107).

In the users' guidelines of the *Target English Learner's Dictionary*, it is stated that only derivates are treated as sublemmata. However, two derivatives, namely "ending" and "endless", are treated as separate lemmata instead of being treated as sublemmata in the article of "end". Many other examples illustrating similar inconsistencies can be found. The lexicographer needs to place and treat sublemmata consistently so that users can find and retrieve the information.

The outer access structure can be enhanced by the use of typographical devices, for instance, the bold printing of lemmata helps the user in his/her search for a particular lemma in the word list. Lemmata can also be indented or protruded thereby being more conspicuous and identifiable. Typographical devices acting as structural indicators of homonyms also form part of the outer access structure.

Decisions about the placement of multiword units and collocations also have implications for the outer access structure and consequently the accessibility of the dictionary. Multiword units and collocations will not be discussed further, but the following comment can suffice: depending on the level of the learner and the function of the dictionary, these data categories should be organised according to a consistent, transparent system which will enhance access to the data.

3.6   Outer rapid access structure

The last part of the outer access structure consists of the outer rapid access structure. To speed up access in the word list, structural indicators are usually used to separate the different word list sections. According to Hausmann and Wiegand (1989: 335), these bold face capitals A, B … Z represent the access alphabet of the dictionary. These non-typographical structural indicators can be supported by a thumb index on the outside of the dictionary page. It forms part of the outer rapid access structure giving quick access to the pages dealing with the letter or letters in question. The outer access structure is continued in the running heads of each page. They are guiding elements which consist of one or more words or parts of words at the top of each page.

Different dictionaries employ different systems of running heads, for instance OALD gives the first lemma of the left page and the last lemma of the right page. A clearer indication of the scope of a page forms part of a user-friendly approach. In the *Cambridge Learner's Dictionary* the space reserved for running heads is used to explicate an important non-typographical structural marker. The user thereby has rapid access to specific information without having to page to the users' guidelines for clarification.

A further element of the outer access structure which contributes to a rapid detection of different dictionary parts is coloured pages or edges of pages. In a bilingual learner's dictionary these pages can divide the two main parts of the dictionary, whereas in a monolingual learner's dictionary they can indicate study pages, minigrammars or illustration leaflets/plates.

3.7   Conclusion

Lexicographers have to provide an innovative and creative outer access structure to facilitate the outer search route of the learner. Decisions regarding the outer access structure must be explained and illustrated in the front matter. Lexicographers have various options to improve the outer access structure of the dictionary. Gouws (1996: 21) sums it up as follows: "On the one hand, the lexicographer could add special entries or alter the presentation of macrostructural elements and on the other hand, the lexicographer could include additional access structures to create a poly-accessible dictionary." An example of a dictionary employing an innovative outer access structure is the *Kenkyusha's Lighthouse English–Japanese Dictionary* where structural markers indicate the usage frequency of the lemmas. Furthermore, the core vocabulary is presented in colour and a bigger type face. It is especially useful for the beginner because core vocabulary can be identified at a glance.

4.   The inner access structure

A second main component of the access structure of a dictionary is the access structure of the microstructure or the internal access structure. The user's internal search route starts at the lemma and continues into the dictionary article. Svensén (1993: 16) describes the internal search route of the user as follows: "When the right entry word has been found, the user also has to find the right section within the entry, and within that section the material is grouped in a certain order and presented according to certain conventions, which can be fairly difficult to learn."

4.1   Data distribution structure

The strategic placement of data in the dictionary article is determined firstly by the data distribution structure. The design of a data distribution structure and data distribution program must form part of the dictionary plan. Bergenholtz and Tarp (1995: 188) describe the distribution structure as the structure of linguistic (and encyclopaedic) information distributed across or occurring in different places in the dictionary. Basically, this information may occur in three different places: in the individual articles, in different types of framed articles or in separate components like a dictionary grammar. The inner access structure also refers to the form and ordering of data in the dictionary article. In this regard Opitz (1983: 174) demands the following: "In order to prove successful, user convenience must receive high priority. This includes first of all the choice of a type of type style that can be deciphered even under such adverse conditioning as poor lightning, haste and fatigue. Next come matters of organisation: how to allow for fast scanning of an entry in order to find a particular item of information while avoiding mistakes and confusion, or even irritation, on the user's part."

4.2   Micro-architecture

A dictionary article should also display a micro-architecture. According to Wiegand (1996: 143), the micro-architecture "aids the reception of this mode of presentation so that dictionary articles which are arranged in this manner are user-friendlier than articles which lack a micro-architecture". The following articles illustrate the lack and occurrence of micro-architecture.

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

Example 7: Article without micro-architecture

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

Example 8: Article with micro-architecture

The conventional linear presentation of data types in example 7 makes it difficult for the user to retrieve the information and the process also takes longer. In example 8 the data types are clearly separated, also by the use of space, and the polysemes are placed under one another. The internal search route of the second article is facilitated by the finely planned micro-architecture. Every microstructural element must have a function and place in the dictionary article and it must be presented in an unambiguous way so that the information can be retrieved easily and rapidly.

The micro-architecture of dictionary articles also implies that all the different microstructural elements have to be employed in an interactive relation to one another. The different microstructural categories should complement each other without increasing the degree of textual condensation in the article. It is therefore of great importance that the lexicographer already decides on the design and well-structured micro-architecture of the article during the planning stages in order to assist the user when he/she is negotiating the article internal search route. The lexicographer can use a lay-out sample with specimen articles to make sure that the micro-architecture leads to user-friendliness and readability.

4.3   Order of data types in the microstructure

Another aspect of the inner access structure is the order in which data types are presented. According to Gouws (1993: 37), the hierarchical character of the microstructure gives prominence to the first entries and gives them a position of major focus, whereas the last entries receive far less focus. This should influence the organisation of information in the dictionary. Being familiar with the specific typological criteria, a lexicographer knows what the priority of his/her dictionary is and which entry should receive the most emphasis. The lexicographers of TAW deem the translation equivalents, followed by the example material, as the most important data categories in the dictionary article. Consequently these entries are placed at the beginning of the article in the focus position. On the other hand, parts of speech are indicated at the end of the article. Because the ordering of the data categories in this dictionary deviates from the traditional placing, it is marked with a structural marker, namely a square.

4.4   Data-architecture

One of the problems users experience on their internal search is the unclear division between data categories. This problem can be solved, firstly by using a well-developed micro-architecture, thereby incorporating white spaces. A second solution is to distinguish between different search zones by using typographical structure indicators, for example semibold, italic, normal, capital letters, etc. Although they are not functional text segments but attributes of the data form, they indicate boundaries between different search zones. Decisions about the typographical structure indicators or markers form part of the data-architecture of the article. It requires that the lexicographer actively participate in the layout of articles in order to create unambiguously defined data units which enhance the access to information. The data-architecture should also help to create a coherent and integrated article. Examples 9.1 and 9.2 illustrate the effect of typographic structural markers or the lack thereof on the easy retrieval of information.

An important fact to keep in mind when planning the data-architecture, as with any other aspect of a learner's dictionary, is to guard against the excessive varying of the structural indicators which can hinder readability and complicate the internal search route.

gek, (s) **(-ke),** fool, madman, booby; *~ke en DWASE skryf hulle name op deure en glase,* fools' names and donkeys' faces are to be seen in public places; a white wall is a fool's paper; *elke ~ het sy GEBREK,* every rose has its thorn; every bean has its black; *iem. vir die ~ HOU,* make a fool of someone; *die ~ke kry die KAARTE,* fortune favours fools; *staan en KYK soos 'n ~,* stare like an idiot; *dit is nog LANK nie so ~ nie,* it is not a bad idea; *iem. vir die*

~ laat LOOP, send someone on a fool's errand; *die ~ SKEER met iem.,* make a fool of someone; *soveel WERD as 'n ~ daarvoor wil gee,* it is worth as much as a fool would give for it; *een ~ kan meer vrae vra as wat 'n honderd WYSES kan beantwoord,* one fool can ask more questions than ten wise men can answer; (w) **(ge-),** play the fool; (b, bw) **([-ke]; -ker, -ste),** foolish, mad, queer, crazy, fond (of); *~ wees van ANGS,* be frantic with anxiety; *~ GENOEG, sy ...,* oddly enough she ...; *~ na sy KIND wees,* dote on his child; *dis te ~ om LOS te loop,* that is most absurd; it does not bear thinking of; *nie so ~ as wat hy LYK nie,* he is not as green as he is cabbage-looking; *iem. ~ MAAK,* drive a person crazy; *~ NA,* mad about (on); *jou ~ SOEK,* hunt high and low; *die ~ste van die TOESTAND,* the oddest feature of the situation; *~ na VIS wees,* be very partial to (fond of) fish; *~ van WOEDE,* mad with rage; *~ WORD,* go mad.

Example 9.1: *Pharos*: punctuation

gek[1] fool *Only a* **fool** *would leave his car with the keys still in it.* Net 'n **gek** sou sy motor verlaat met die sleutels nog daarin.

☐ **gek** *selfstandige naamwoord (meervoud* **gekke***)*

gek[2] ❶ foolish, silly, stupid *It is* **foolish**/**silly**/**stupid** *to go too deep into the sea if you can't swim.* Dit is **gek** om te diep in die see in te gaan as jy nie kan swem nie. ❷ mad "*Let's go for a ride on our bicycles.*" – "*Are you* **mad***? It's raining!*" "Kom ons gaan ry 'n ent op ons fietse." – "Is jy **gek**? Dit reën!"

♦ **gek maak** drive mad *That dog will* **drive** *me* **mad** *with its barking!* Daardie hond sal my nog **gek maak** met sy geblaf!

☐ **gek** *byvoeglike naamwoord (attributief* **gek***)* **gekker, gekste**

Example 9.2: TAW: punctuation + structural markers

4.5   Inner texts

Usage notes are an important inner access structure, which is employed in most learner's dictionaries. These inserted inner texts are especially appropriate for learner's dictionaries. They can easily be distinguished from the rest of the article, because they are presented in a different manner, e.g. in boxes or tables. If the lexicographer wants to draw the attention of his intended target user to a specific aspect of a lemma, the inclusion of an inserted inner text will be a clear milestone on the search path of the inner access structure. The lexicographer has to decide on a system for the placing of usage notes, either after a specific data type, or at the end of the article.

4.6   Inner rapid access structure: data-exposure structure

The use of non-typographical structural markers is a last method of indicating and marking borders between data categories in order to enhance the inner rapid access structure. They are used firstly to mark different polysemous senses of a lemma. The effect of a grouped polysemy in example (7) compared with an ungrouped polysemy in example (8) is evident.

Example 7 shows senses which are presented in succession and lead to a horizontal, user-*un*friendly search route. In example 8 all the senses are grouped under one another, in vertical text blocks, which are easy to distinguish and form part of the inner access structure.

Explicit non-typographical indicators marking specific data categories form part of the article internal data-exposure structure. These inner rapid access structures expose the presented data and allow the user a faster access in

the data look-on. The data-exposure program of dictionary articles, especially in a learner's dictionary, should be clear and transparent. If not, the user has to be familiar with the inner structure of the articles in the dictionary and with the order of the article positions.

4.7   Conclusion

Renovation in present-day learner's dictionaries focuses particularly on the inner access structures and rapid inner access structures. In many cases, these changes increased the quality of the dictionary, but it can also lead to an increased text density. This special balance that must be maintained is expressed as follows by Svensen (1993: 16): "It is not merely the information needs of different categories of users that are of importance. An equally important factor is their ability to find and make use of the information given. Here the lexicographer is forced into a difficult balancing act. On the one hand, the dictionary must contain as much information as possible without becoming unnecessarily bulky. On the other hand, this information must not be so compressed that the intended user will not be able to utilise it."

5.   Meta-access structure

A third and last aspect of the access structure which will not be addressed in this article is the meta-access structure. This refers to the language in which data is presented. Hausmann and Wiegand (1989: 341) argue that from one and the same lexicographic data unit or treatment unit two different users may reconstruct different information. In order for the user to extract this information the language usage should match the target user group of the dictionary. Therefore the language used in the front and back texts of the dictionary as well as definitions, examples etc. should be accessible. The metalanguage should therefore also help the user to retrieve information as easily and rapidly as possible.

6.   Three important questions

When the access structure in learner's dictionaries are considered, as is the case in this article, certain questions immediately arise.

1.   What is the use of the access structure in general?

2.   Is there a difference between the access structure of general dictionaries and learner's dictionaries?

3.   Do different levels of learners affect the access structure?

6.1   Question one

As already explained, the access structure can be viewed as the form in which content is presented. Without an accessible form, data presented and arranged in the macro-, micro-, frame and distribution structures cannot be accessed and employed by the user. The failure of dictionary consultation procedures is often attributed to the inadequate reference skills of the users, in this case learners. But it is more often than not the structure of the dictionary that impedes the user from sufficient access to data. According to Gouws (2001: 102), the rapid and unimpeded access of the users to the relevant data presented in the dictionary has to be regarded as a prerequisite for a successful lexicographic product in a user-driven approach. The ease with which macro- and microstructural information is located and retrieved determines the quality and accessibility of the presentation. The lexicographer can therefore use the access structure as primary instrument to ensure the user-friendliness of a dictionary.

6.2   Question two

Different outer texts in different dictionaries suggest that the outer access structure of a dictionary is designed with a specific user in mind. In most comprehensive and standard dictionaries, the outer search route will start at the cover, possibly continue to the table of contents, the running heads and the thumb index and end at the desired lemma. Recently learner's dictionaries have employed certain bridge texts to facilitate and enhance the outer search route of the dictionary. These integrated outer texts are additional wordlists or equivalent registers, found mostly in learner's dictionaries for beginners. They help the user with a restricted knowledge of the language and specifically the spelling of a foreign language word to reach the lemma via a mother tongue equivalent. An equivalent register enables the learner to reach the unknown word with the aid of a translation and thus enhances text reception and production in the foreign language. In the same manner as the users' guidelines, an outer text presenting an equivalent register should be regarded as a compulsory text in monolingual and monoscopal bilingual learner's dictionaries for beginners and intermediate learners. If the learner reaches the advanced stage, where thinking in the foreign language becomes a possibility, the need for this special outer text decreases.

In the *Passport English Hebrew Learner's Dictionary*, the equivalent register is given the status of a Hebrew–English dictionary or "reverse dictionary" that also "forms an integral part of the dictionary in order to enable the learner to locate the English headword" (p. viii).

It is also evident that the inner access structure varies according to the dictionary type. If one examines articles from different dictionaries it would not be difficult to decide what dictionary type they represent. It is not only the

294   Mariza Steyn

amount and density of data that lead one, but the internal structure of the arti-
cles also suggests that different dictionary types have different inner access
structures to correlate with specific users.

| English	Hebrew		Hebrew	English
image	תַּדְמִית נ.		**ת**	
printout	תַּדְפִּיס ז.			
frequency (of radio)	תֶּדֶר ז. (תְּדָרִים)			
briefing	תַּדְרִיךְ ז.			
tea	תֵּה ז.			
to wonder, ask oneself	תָּהָה פ״ע (לִתְהוֹת)			
abyss	תְּהוֹם ז״נ			
fame, glory	תְּהִילָה נ.	תּ׳	*tav* (22nd letter of Hebrew alphabet); 400	
Book of Psalms	תְּהִילִים ז״ר			
procession	תַּהֲלוּכָה נ.	תָּא ז.	cell; locker, compartment; cabin	
process	תַּהֲלִיךְ ז.	– תָּא הַטַּיָיס	cockpit	
mark; note (music); character (computers)	תָּו ז.	ת״א ר״ת תֵּל־אָבִיב	TA, Tel-Aviv	
(facial) features	תָּווֵי־פָּנִים	תָּאֲבוֹן ר׳ תֵּיאָבוֹן		
tav (22nd letter of Hebrew alphabet)	תָּו נ.	תַּאֲווָה נ.	lust, passion, desire	
matching	תּוֹאֵם ת.	תֵּאוּם ז.	coordination; adjustment	
title; degree (academic)	תּוֹאַר ז.	תְּאוֹם ז. תְּאוֹמָה נ.	twin	
transport	תּוֹבָלָה נ.	– תְּאוֹמִים	twins	
prosecutor	תּוֹבֵעַ ז. תּוֹבַעַת נ.	– מַזַּל תְּאוֹמִים	Gemini	
thanks, gratitude	תּוֹדָה נ.	תְּאוּנָה נ.	accident	
with thanks	– בְּתוֹדָה	– תְּאוּנַת־דְּרָכִים	road accident	
Many thanks!	– רַב־תּוֹדוֹת	תְּאוּצָה נ.	acceleration	
Thank you very much!	– תּוֹדָה (רַבָּה)!	– לְקַבֵּל תְּאוּצָה	accelerate	
thank God!	– תּוֹדָה לָאֵל	תֵּאוּר ז.	description	
thanks to	– תּוֹדוֹת לְ־	תְּאוּרָה נ.	lighting	
consciousness	תּוֹדָעָה נ.	תָּאַם פ״י (לְתְאוֹם)	correspond	
chaos	תּוֹהוּ וָבוֹהוּ	תְּאֵנָה נ. (תְּאֵנִים)	fig	
feature (of landscape), line	תְּוָואי ז.	תַּאֲרִיךְ ז.	date	
label, tag; stigma	תָּווִית נ.	– תַּאֲרִיךְ יַעַד	deadline	
hope	תּוֹחֶלֶת נ.	תְּבוּאָה נ.	corn, grain crops	
life expectancy	תּוֹחֶלֶת־חַיִּים	תְּבוּנָה נ.	intelligence, sense	
inside	תּוֹךְ ז. (תּוֹכוֹ)	– בִּתְבוּנָה	intelligently, wisely	
inside; within, in; among	– בְּתוֹךְ	תְּבוּסָה נ.	defeat	
into	– לְתוֹךְ	תְּבִיעָה נ.	demand, claim	
out of, from	– מִתּוֹךְ	– תְּבִיעָה מִשְׁפָּטִית	lawsuit	
while, in the course of	– תּוֹךְ כְּדֵי־	תֵּבֵל נ״ו	the world	
parrot	תּוּכִּי ז.	תַּבְלִין ז.	spice	
contents; content	תּוֹכֶן ז. (תְּכָנִים)	תַּבְנִית נ. (תַּבְנִיּוֹת)	mould; (baking) tin	
(table of) contents	– תּוֹכֶן הָעִנְיָינִים	תָּבַע פ״י (לְתְבּוֹעַ)	to claim, demand; sue	
software	תּוֹכְנָה נ.	– תָּבַע לְדִין	sue	
programme (of shows)	תּוֹכְנִייָה נ.	תִּגְעֹרֶת נ.	fire	
plan, programme, project	תּוֹכְנִית נ.	– בַּקְבּוּק תַּבְעֵרָה	petrol bomb	
curriculum; syllabus	– תּוֹכְנִית־לִימוּדִים	תַּבְרוּאָה נ.	sanitation	
be planned, be designed	תּוּכְנַן פ״ע	תַּבְשִׁיל ז.	dish (cooked food)	
		תָּג ז.	badge, label, tag	

Example 10: An equivalent register in *Passport*

A comprehensive dictionary, such as the WAT, has a sophisticated user in
mind and consequently only a few extra indicators facilitate the search route of
the user within complex dictionary articles. The structure of articles from a
standard Afrikaans dictionary, HAT, also suggests a user with developed/
adequate reference skills. Some typographical and non-typographical markers
indicate the borders between different data categories. The *Cambridge Learner's
Dictionary* uses a sophisticated access structure, designed to help the learner to
overcome possible obstacles, whether it be because of language level or lack of
reference skills in the search for specific data. Boundaries between all data
categories are marked, polysemes are distinguished through guide words and
illustrations help the user with immediate knowledge transfer.

6.3   Question three

The third important question related to the topic is whether different levels of learners affect the access structure of learner's dictionaries. The following graphs will illustrate the possible relationship between the levels of learners and the outer and inner access structure. If one groups together the different outer access structures, as presented in existing learner's dictionaries, they fall into three categories. Firstly there are those structures that are compulsory, such as the alphabetical order, the access alphabet, running heads and the users' guidelines in the front text. Secondly some dictionaries also have a few optional structures in order to enhance access to the macrostructure. These optional structures are thumb indexes, coloured pages and some outer texts. Lastly some learner's dictionaries for beginners also add bridge texts (as illustrated previously) to assist the beginner in finding the required lemma via the mother tongue.
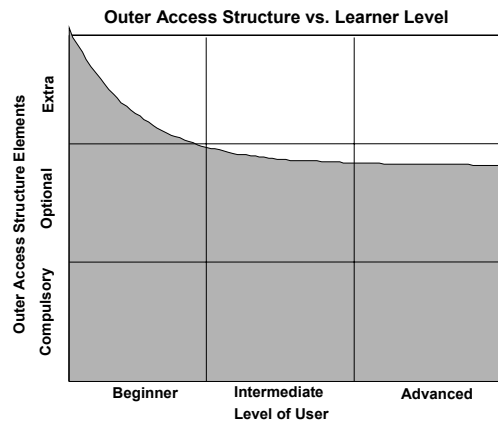
The horizontal axis of the graph represents the learners' levels and the vertical axis the three groups of outer access structures. The graph shows a decline in the number and type of outer structures in relation to the level of user. At beginner's level, the user needs as much help as possible to facilitate the outer search route.

There is no distinction between learner's dictionaries with a function of production in the foreign language and reception of the foreign language, because it would not influence the outer access structure directly. A learner's dictionary designed for the purpose of reception might have a larger macrostructure with more lemmata than one for production. On this graph the lines indicating production and reception therefore fall together.

In the case of the inner access structure of learner's dictionaries, the distinction between dictionaries for production and reception does have an influ-

ence on the number of inner access structures. The three categories of inner access structures are the following. The first contains implicit inner access structures being the data distribution structure and the order of the data types. The second category comprises explicit compulsory access structures, including data-architecture and micro-architecture. The third contains explicit optional rapid access structures, inner texts and data exposure structures.

Inner Access Structure vs. Learner Level
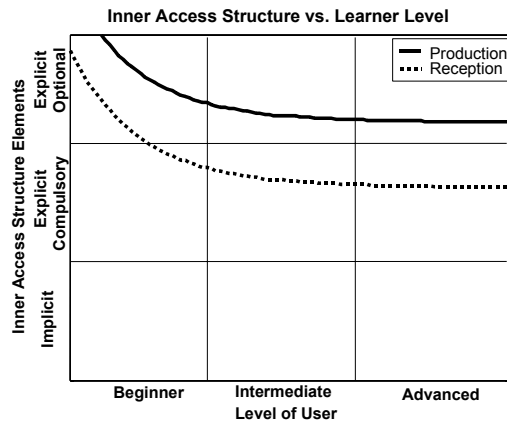
The second graph also indicates a decline in inner access structural elements in relation to the specific category of learner. In this case, however, a distinction between the productive and receptive modes appears. Lexicographers of dictionaries with production in the foreign language as dictionary function need to add extra data in order to enable the learner to use it actively. These data categories, like the indication of frequent collocations and proverbs, and the inclusion of syntactical combination principles etc. demand more explicit structural elements than is the case with dictionaries designed for reception.

7.   Conclusion

The access structure functions as primary structure in guiding the user during the consultation process. The user's search problem can be solved swiftly and easily if the access structure of a dictionary, specifically a learner's dictionary, is developed and applied accurately. Existing learner's dictionaries, as shown in this article, include new outer texts and subtexts to explicate specific data. They also employ innovative strategies of presenting data in dictionary articles. However, most learner's dictionaries lack a well-developed access program that guides the user along every step of the search route. In this article, some suggestions have been made and illustrated in order to improve the access profile of learner's dictionaries.

References

Dictionaries

Crowther, J. et al. (Eds.). 1995[5]. *Oxford Advanced Learner's Dictionary*. Oxford: Oxford University Press.

Du Plessis, M. 1998. *Tweetalige Aanleerderswoordeboek/Bilingual Learner's Dictionary*. Cape Town: Pharos Woordeboeke.

Eksteen, L.C. 1997[14]. *Pharos Groot Woordeboek/Major Dictionary*. Cape Town: Pharos Woordeboeke.

Gefen, R. et al. (Eds.). 1996. *Passport English Hebrew Learner's Dictionary*. Jerusalem: Kernerman Publishing.

Gillard, P. (Ed.). 2001. *Cambridge Learner's Dictionary*. Cambridge: Cambridge University Press.

Kempcke, G. et al. (Eds.). 2000. *Wörterbuch Deutsch als Fremdsprache*. Berlin: Walter de Gruyter.

Odendal, F.F. et al. (Eds.). 1994. *Verklarende Handwoordeboek van die Afrikaanse Taal*. Johannesburg: Perskor.

Rundell, M. et al. (Eds.). 2002. *Macmillan English Dictionary for Advanced Learners*. London: Bloomsbury Publishing.

Sinclair, J.M. (Ed.). 1996[2]. *Collins COBUILD Learner's Dictionary*. London: HarperCollins.

Summers, D. (Ed.). 1984. *Target English Learner's Dictionary*. Cape Town: Maskew Miller Longman.

Summers, D. (Ed.). 1995[3]. *Longman Dictionary of Contemporary English*. Harlow: Longman.

Takebayashi, S. and Y. Kojima (Eds.). 1990[2]. *Kenkyusha's Lighthouse English–Japanese Dictionary*. Tokio: Kenkyusha.

Van Schalkwyk, D.J. (Ed.). 1996. *Woordeboek van die Afrikaanse Taal*, *Part X*. Stellenbosch: Bureau of the WAT.

Other literature

Almind, R. and H. Bergenholtz. 2000. Die ästhetische Dimension der Lexikographie. Fix, U. and H. Wellmann (Eds.). *Bild im Text — Text und Bild*: 259-288. Heidelberg: C. Winter.

Béjoint, H. 1994. *Tradition and Innovation in Modern English Dictionaries*. Oxford: Clarendon Press.

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

Bolinger, D. 1985. Defining the Indefinable. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 69-73. Oxford: Pergamon Press.

Gouws, R.H. 1993. Afrikaans Learner's Dictionaries for a Multilingual South Africa. *Lexikos* 3: 29-47.

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

Gouws, R.H. 2001. The Use of an Improved Access Structure in Dictionaries. *Lexikos* 11*:* 101-111*.*

Gouws, R.H. 2002. The Outer Texts in the *De Gruyter Wörterbuch Deutsch als Fremdsprache*. Wiegand, Herbert E. (Ed.). 2002. *Perspektiven der pädagogischen Lexikographie des Deutschen II*: 471-484. Tübingen: Max Niemeyer.

Hausmann, F.J. and H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J., O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*: 328-360. Berlin: Walter de Gruyter.

Kammerer, M. and H.E. Wiegand. 1998. Über die textuelle Rahmenstruktur von Printwörter-
büchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-238.

Opitz, K. 1983. The Terminological/Standardised Dictionary. Hartmann, R.R.K. (Ed.). 1983. *Lexico-
graphy: Principles and Practice*: 163-180. London: Academic Press.

Stark, M. 2001. *Die toegangstruktuur in verklarende aanleerderwoordeboeke*. Unpublished M.A. Thesis.
Stellenbosch: University of Stellenbosch.

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

Tono, Y. 1992. The Effect of Menus on EFL Learners' Look-up Processes. *Lexikos* 2: 230-253.

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

Wiegand, H.E. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie,
Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband.* Berlin/New York: Walter
de Gruyter.

Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language[*]

Sven Tarp, *Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark (st@asb.dk)*

Abstract: After a discussion of a selected part of the existing theoretical literature, the concept of foreign-language text production is analysed within the framework of the broader concept of the foreign-language learning process. Two main types of foreign-language text production are discussed, i.e. text production with and without an outline in the learner's mother tongue, followed by an analysis of the kind of assistance dictionaries can provide to the user. The main conclusion drawn from this discussion is that a dictionary designed to assist a user with both types of text production should include both an L_1–L_2 word list and an L_2 (or L_2–L_1) word list. On this basis, a proposal for the lexicographic data that should be included in the respective word lists is presented, and this data is then compared with the data needed to assist the user in foreign-language text reception. The conclusion is that it is quite possible to conceive bifunctional dictionaries that can provide assistance for both foreign-language text production and foreign-language text reception.

Keywords: LEXICOGRAPHY, LEARNER'S LEXICOGRAPHY, LEARNER'S DICTIONARIES, LEXICOGRAPHIC FUNCTIONS, COMMUNICATION-ORIENTATED FUNCTIONS, FOREIGN-LANGUAGE TEXT RECEPTION, FOREIGN-LANGUAGE TEXT PRODUCTION, USER CHARACTERISTICS, USER NEEDS, USER SITUATIONS, LEXICOGRAPHIC DATA

Opsomming: Gedagtes oor woordeboeke ontwerp om gebruikers te help met teksprodukise in 'n vreemde taal. Na 'n bespreking van 'n uitgesoekte deel van die bestaande teoretiese literatuur, word die konsep van vreemdetaalteksproduksie ontleed binne die raamwerk van die breër konsep van die aanleerposes van 'n vreemde taal. Twee hooftipes vreemdetaalteksproduksie word bespreek, nl. teksproduksie met en sonder 'n oorsig van die aanleerder se moedertaal, gevolg deur 'n ontleding van die soort hulp wat woordeboeke aan die gebruiker kan verleen. Die hoofgevolgtrekking gemaak uit hierdie bespreking is dat 'n woordeboek wat ontwerp is om 'n gebruiker met altwee tipes teksproduksie te help, sowel 'n L1–L2-woordelys as 'n L2- (of L2–L1) woordelys behoort in te sluit. Op hierdie basis word 'n voorstel van die leksikografiese data wat in die onderskeie woordelyste ingesluit behoort te word, aangebied, en hierdie data word dan vergelyk met die data benodig om die gebruiker by vreemdetaalresepsie te help. Die gevolgtrekking is dat dit heeltemal moontlik is om bifunksionele woordeboeke te bedink wat hulp kan verleen met sowel vreemdetaalteksproduksie as vreemdetaalteksresepsie.

---

[*] This article is a slightly revised version of the closing paper read at the Seminar on Learners' Dictionaries presented by the African Association for Lexicography at the University of Pretoria, Pretoria, Republic of South Africa, 11–12 November 2002.

Sleutelwoorde: LEKSIKOGRAFIE, AANLEERDERSLEKSIKOGRAFIE, AANLEERDERS-
WOORDEBOEKE, LEKSIKOGRAFIESE FUNKSIES, KOMMUNIKASIEGEORIËNTEERDE FUNK-
SIES, VREEMDETAALTEKSRESEPSIE, VREEMDETAALTEKSPRODUKSIE, GEBRUIKERSEIEN-
SKAPPE, GEBRUIKERSBEHOEFTES, GEBRUIKERSITUASIES, LEKSIKOGRAFIESE DATA

1.   Introduction

In modern theory of lexicographic functions, one of the basic communication-
orientated functions of dictionaries is to assist users in solving problems related
to text production in a foreign language (Tarp 1992, 1994, 1995, 1998, 2000,
2003, Bergenholtz and Tarp 2002, 2003, 2004a, 2004b). But what exactly is meant
by text production in a foreign language? How does it take place? What is the
typology of the users who would be expected to consult a dictionary in order to
solve problems related to this kind of text production? What are these prob-
lems? And how can dictionaries assist users in solving them? These are funda-
mental questions of both theoretical and practical importance that must be
answered in order to further develop the functional theory of lexicography.
This article will try to advance in this direction. It will argue that, in terms of
dictionaries, foreign-language text production can only be understood within
the framework of learner's lexicography in the broad sense of the word. It will
put forward a proposal for a general model of a dictionary designed to assist
with the production of foreign-language texts. And finally it will show how
such a dictionary can easily be developed also to serve the reception of foreign-
language texts, i.e. to combine at least two different lexicographic functions.

2.   Theoretical contributions

Some ten years ago, the German lexicographer Joachim Mugdan (1992a, 1992b)
shocked quite a number of lexicographers when he argued that future L_2–L_1
dictionaries, in order to avoid the production of too many dictionaries for each
pair of languages, should be designed in such a way that they could also assist
L_1 users in solving problems related to text production in L_2. At that time the
so-called "active–passive" theory reigned in the world of lexicography (Kro-
mann et al. 1984, 1991). This theory, among others, included L_2 text production
in the concept of L_1–L_2 "translation". Hence, according to this theory, and the
prejudices of many lexicographers, assistance with L_2 text production should
be provided in an L_1–L_2 dictionary, i.e. the opposite view of the one proposed
by Mugdan. He, however, was not the first lexicographer to put forward such
an unorthodox idea. Fifteen years earlier, his fellow countryman F.J. Haus-
mann had suggested something similar in his famous book on modern French
dictionaries (Hausmann 1977). Hausmann recommended that free text produc-
tion in a foreign language should be supported by a monolingual L_2 dictionary
and that bilingual dictionaries generally should be used as late as possible.

Hausmann made a clear distinction between free foreign-language production and translation-related production (Hausmann 1977: 145. Translated by ST):

> Production of a foreign-language text can, as free production, be based on one's own thoughts or can be transferred from a mother-tongue text into the foreign language (L_1–L_2 translation).

Both Mugdan's and Hausmann's observations and recommendations are of great value for the development of a theory of the lexicographic function of "assistance to foreign-language production". Even more important and valuable are, however, the reflections made by the Russian lexicographer L.V. Shcherba in 1940. Although Shcherba apparently considered lexicography a subdiscipline of linguistics, he has without doubt the merit of being the founder of a general theory of lexicography. Shcherba discussed the production of foreign-language texts within the framework of the language-learning process. According to the Danish scholar Hans Kristian Mikkelsen who has studied Shcherba's original Russian works, Shcherba was generally strongly opposed to the contrastive methodology in foreign-language didactics because it could contribute to "a mixed bilingualism due to numerous transfers from L_1" and therefore "only goes for L_2 learning at the beginner's level" (Mikkelsen 1992: 34). In his famous essay on a general theory of lexicography, Shcherba (1940: 341) once more stressed this point of view:

> Any true pedagogue advises students to discard translating dictionaries as soon as possible and switch to the defining dictionary of the foreign language. A translating dictionary, then, is only useful for beginning foreign language students.

Shcherba was extremely critical of the existing bilingual dictionaries, above all because they did not help the user to grasp the meaning of the foreign word. Although he pointed out that the faults of these dictionaries could be partly eliminated by adding various notes and examples, he himself preferred what he called a radical solution to the problem, namely the compilation of "foreign defining dictionaries in the students' native language". These dictionaries should mainly contain definitions — or explanations — of the foreign words, written in the user's own language. Shcherba (1940: 341), however, also accepted mother-tongue equivalents in those cases "when this would simplify definition and would not be detrimental to a full understanding of the foreign word's true nature".

From the very way this whole idea is presented, it becomes clear that the defining dictionary proposed by Shcherba is not just a traditional dictionary for the purpose of foreign-language reception, but mainly a dictionary designed to assist the assimilation of the foreign-language system (lexis and grammar) and the production of texts in this language. Until this new type of dictionary was developed, Shcherba considered the translation dictionary a "malum necessarium". Although he himself was a polyglot who easily assimilated a new language, Shcherba on the other hand was very much aware of the fact that for-

eign-language students at a beginner's level would have big difficulties in using such a dictionary due to their limited knowledge of the foreign language. In this way he introduced a very important distinction, in terms of different lexicographic needs, between users at a beginner's level and more experienced and advanced users (Shcherba 1940: 341-342):

> However, a special type of translating dictionary must still be retained for the people who do not know the foreign language very well, but nevertheless need to translate something into that language from time to time. Generally speaking, the basic rule of a competent foreign-language teaching methodology is that one should not — even mentally — translate from one's own native language, but try to think in the foreign language within the limits of one's knowledge, referring when necessary to ideological or synonym dictionaries as well as to good foreign defining dictionaries, but by no means to translating dictionaries. However, when applied to practical life, this presupposes a rather high level of skill in the foreign language.

Here it must be noted that when Shcherba speaks about translation in the above quotation, he apparently does not mean translation in the narrow sense of the word. According to Mikkelsen (1992) and Mugdan (1992a), for Shcherba real translation seems to be a creative activity that generally should be carried out without the consultation of dictionaries. Shcherba's idea was that when you have understood a text, then you would also be able to translate it and find the corresponding equivalents. So the use of the word *translation* in the above quotation rather seems to refer to a situation where the student first thinks in his/her own language and then "translates" it and produces a text in the foreign language. This theoretical distinction between two different types of foreign-language text production, apart from the one related to real translation, has wide-ranging consequences for practical lexicography. Shcherba (1940: 342) himself suggests what he calls a "special type of dictionary" in order to cover the needs of the foreign-language students at a beginner's level. This dictionary should enable these users "to translate non-fictional texts into the foreign language without making gross mistakes" and, as such, it should provide exact indications on how to translate mother-tongue words "in various contexts in order to be understood and to avoid appearing ridiculous". Shcherba (1940: 342) summarises his ideas as follows:

> To summarise this section, I repeat what I said in the preface to my dictionary: for every pair of languages, four dictionaries are needed. Two foreign defining dictionaries with explanations in the native language of the user are absolutely necessary, and depending on actual needs, two translating dictionaries of the special type indicated above, from the native language into the foreign language.

Although he speaks of four different types of dictionaries, what Shcherba really does in his important contribution to lexicography is to discuss different lexicographic functions, i.e. giving assistance to users with different mother tongues and levels of L_2 mastery in solving problems related to different user situations. This was a big step forward for lexicography.

3.   Who needs assistance with foreign-language text production?

As could be seen in the previous section, Shcherba only discussed the production of foreign-language texts within the framework of the language-learning process. This approach, however, can be extended to all kinds of foreign-language text production. Who else than foreign-language learners, in the broad sense of the word, would be expected to produce texts in a foreign language? Ever since the publication of the first edition of the *Oxford Advanced Learners' Dictionary* in 1948, the term "advanced learner" has been commonly accepted and used both within lexicography and linguistics. This term gives a broad sense to the meaning of the word "learner". The learning of a foreign language is normally a life-long process. Very few people succeed in reaching the same degree of mastery of their second language as is the case with their mother tongue. And if they eventually do, they can no longer be considered to be speaking in a *foreign* language. Hence, although it is not common in theoretical literature, it seems quite acceptable to analyse and discuss lexicographic assistance to foreign-language text production within the framework of learner's lexicography. This, however, at the same time simplifies and complicates the problem. It simplifies it because it is no longer necessary to distinguish between "normal" L_2 production and learner-related L_2 production. And it complicates it because it requires an analysis of the concept of a learner.

Tarp (2003) has put forward the following elements aimed at a determination of a learner profile and the general circumstances, in which the foreign-language learning takes place:

(a)   Proficiency level
(b)   Adult or child
(c)   Emigrant or other type of learner
(d)   "Joint" or separate culture
(e)   Level of general culture
(f)   Conscious or spontaneous learning
(g)   Learning inside or outside the language area
(h)   Learning being exposed or unexposed to the mother tongue
(i)   Learning with or without contrastive methodology
(j)   Learning related or unrelated to a specific textbook or didactic system
(k)   Learning related or unrelated to a particular topic (economy, history, culture etc.)

All these elements are important in order to draw up a complete characteristic of the user profile for a specific dictionary. If the user is a child, for instance, the vocabulary should be more reduced, the definitions and the metalanguage easier, the structures and layout simpler, etc. If the learner's culture is different from that of the foreign language community, notes on cultures should be added. If the learning process is related to a specific textbook, the vocabulary of the dictionary should be adapted to that of the textbook. However, for the pur-

pose of this article points (a) and (i) are the most interesting. As has been seen in the previous section, both Shcherba and Hausmann give recommendations as to how to study a foreign language or produce an L_2 text. However, this sort of recommendations is not the task of lexicography, but of linguistics and language didactics. However sympathetic Shcherba's resistance to the contrastive methodology in language learning might seem, the task of lexicography is or should always be to study the real language-learning process followed by the intended users and to support this process by means of dictionaries, whether or not it is based on the contrastive method of language learning. This means that if the intended user group use this method, a bilingual dictionary of the sort that Shcherba dislikes is indispensable. The task of dictionaries — and lexicography — is to satisfy the needs of a specific group of users in terms of helping them to solve problems related to a specific user situation. And if this situation is due to an undesirable methodology, lexicography nevertheless fulfils its function as long as it does not conflict with more general ethical principles.

4.   What is foreign-language text production?

The foreign-language learning process can be illustrated in the following schematic way:
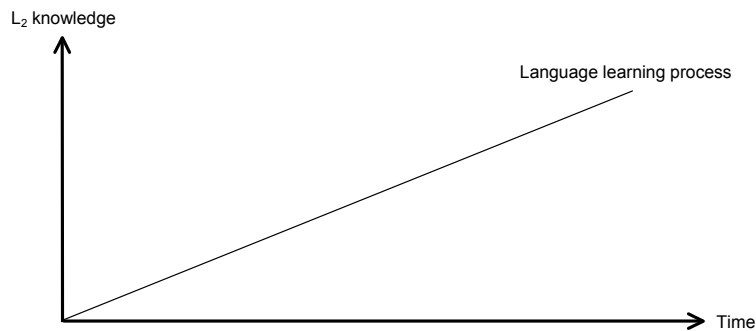
Illustration 1: Foreign-language learning process

Of course, this process can be slower or quicker depending on the concrete learner's talent and efforts. Furthermore, the real line will not be a straight one, but will include periods with slow growth followed by sudden dialectical jumps when quantity transforms itself into quality, i.e. when the passive vocabulary or grammar becomes active. And in the beginning of the process, the foreign-language knowledge will most probably grow quicker and then slower and slower, thus forming a curve. Moreover, every learner will have his/her own individual curve. Nevertheless, at an abstract level, the above schema is fully acceptable for the purpose of this article, i.e. to detect some lexicographically relevant problems in the language-learning process.

Theoretically, it is possible to separate various phases in the L_2 learning process and make dictionaries especially designed to cover learners' needs in each of these phases. A production dictionary conceived for the very first beginner's phase would, for example, require a more reduced vocabulary, fewer senses of the words' meaning, more basic and less specific grammatical data, fewer collocations etc. than a dictionary conceived for more experienced users, and so on. From an economic and logistic point of view, however, it would often be undesirable to produce a big number of highly specialised dictionaries that would only be helpful to the user in a limited period of the L_2 learning process. Dictionaries will, therefore, often be designed to cover several, specific phases, i.e. a much longer period of the learning process, thus combining the data needed to assist the user group in various phases. But in order to achieve a successful combination of data it is first of all necessary to have a thorough understanding of what is meant by foreign-language text production. In this respect, Tarp (1992) separated three different types of L_2 text production, each with its own characteristics:

(a)   Translation-related L_2 text production
(b) L_2 text production based on an outline in L_1
(c)   Free L_2 text production without an outline in L_1

Translation-related L_2 text production is a special type of L_2 production, which makes up the third phase of the L_1–L_2 translation process composed of L_1 reception, L_1–L_2 transfer and L_2 production. Having been thoroughly discussed and analysed by Tarp (2002), it is not the topic of this article.

L_2 text production based on an outline in L_1 is equivalent to what Shcherba called "translation" in his essay discussed in section 2. This user situation has a variety of specific expressions and is highly individualised. First of all, there is the situation where the learner, inside his/her brain, first formulates a phrase in the mother tongue and then "translates" it into the foreign language. Then there is the situation where the learner first formulates a general idea in the mother tongue and then reconstructs it in the foreign language. And, finally, there is the situation where the learner tries to think in the foreign language but due to the limited L_2 vocabulary frequently encounters concepts that he/she only knows how to express in the mother tongue and therefore has to "translate". Furthermore, one and the same learner would not necessarily use only one of these methods, but will most probably switch between them depending on the more or less complicated character of the phrases to be expressed in the foreign language.

Free L_2 text production without an outline in L_1 is the user situation where the learner tries to express him-/herself only on the basis of thinking in the foreign language. It will normally require a relatively high level of mastery of L_2.

The above distinction between foreign-language text production *with* an outline in the mother tongue and free foreign-language production *without* such an outline is of wide-ranging theoretical and practical importance. As has

been seen, Shcherba, Hausmann and Mugdan lay great emphasis on free L_2 text production without an outline in L_1. But for the majority of L_2 learners the real user situation when producing a text in a foreign language is most probably another one. In this regard, Tarp (1992: 39) concluded:

> It rather seems that "production in a foreign language" should be considered a process that in the beginning (i.e. with the learner at a beginner's level) is based on the "producer's" mother tongue (i.e. *with an outline* in the mother tongue) and then gradually is getting based on the foreign language (i.e. *without an outline* in the mother tongue) in correspondence with the "producer's" still greater mastery of this language. In other words: that "production in a foreign language" for the big majority takes place as a combination of production *with* and *without an outline* in the mother tongue.

This way of understanding text production related to the foreign-language learning process can be illustrated in the following schematic way:
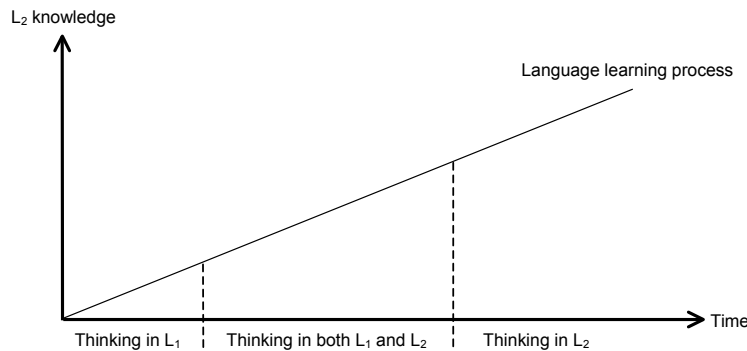

Illustration 2: Text production related to foreign-language learning

Once more, it is evident that the duration of the various phases is unique for each individual learner and might depend on a number of factors such as talent, efforts and methodology. If the contrastive method is used, the phase of L_1 thinking would, for instance, be supposed to last much longer. On the other hand, if the user is following Shcherba's recommendations, he/she might start thinking in the foreign language from the very first day of the learning process in order to formulate simple phrases in L_2. But, of course, if more complex phrases have to be expressed for one reason or the other, this type of learner will also have to switch to the bilingual method, at least for the time being.

5.   Lexicographic assistance to foreign-language text production

The above observations lend a fresh perspective to the question of lexicographic assistance to text production related to the foreign-language learning process.

In order to cover the user's real needs, production-related learner's dictionaries should allow him/her to make consultations through the language(s) relevant in each phase of the learning process. In the first phase when thinking primarily takes place in the mother tongue, the dictionary must provide for consultations through L_1. In the advanced phase when thinking is mainly done in the foreign language, the dictionary must in the same way provide for consultations through L_2. However, in the long phase in between where thinking takes place in both L_1 and L_2, the dictionary has to provide for consultations in both languages if it wants to cover the user's real needs. This can be summarised in the following way:
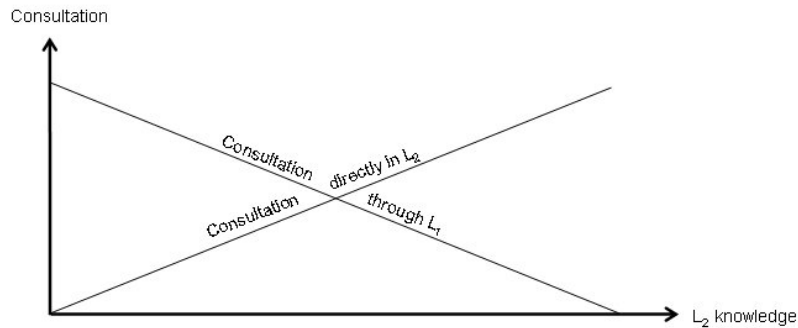
Illustration 3: Consultation of dictionaries to assist L_2 text production

It goes without saying that this complex character of the foreign-language text production has huge consequences for learner's lexicography. The important conclusion to be drawn from this is that — with the exception of dictionaries designed only to serve users who exclusively base themselves in the mother tongue or the foreign language — dictionaries conceived to assist L_2 text production should incorporate both a bilingual L_1–L_2 section and a section with L_2 lemmata, either in one or two volumes. The big challenge for lexicography is to provide a theory that can lead to the making of such dictionaries.

5.1   The lexicographic journey from L_1 to L_2

It has already been mentioned that if users follow the contrastive method in language learning, they will need a bilingual L_1–L_2 dictionary of the type Shcherba detested. However, Shcherba himself admitted that other learners at beginners' level would also have big difficulties in proceeding directly to the foreign words in an L_2 dictionary and he therefore made provision for improving the traditional L_1–L_2 dictionary and for creating a "special type of dictionary". Shcherba died in 1944 before he had the opportunity to put this idea into practice. So it is actually unknown how it would have been conceived in

detail. But it is, nevertheless, a question whether Shcherba did not betray his own opposition to mixed bilingualism with this proposal. There are a number of reasons why bilingual L_1–L_2 dictionaries might never be the most suitable to provide information on L_2.

What information do the users need in order to produce a correct text in L_2? First of all, they need the foreign word, its orthography, pronunciation, part of speech and gender. Then they need confirmation that the word in question has the right meaning or sense for the concrete purpose. Furthermore, they need synonyms, antonyms, hyponyms etc. as well of pragmatic and even cultural information in order to vary the style and assure that the word is used in the correct context. And finally, of course, they need information on inflectional morphology and syntactical properties which also includes word combinations such as collocations, idioms and proverbs.

This is a very big amount of lexicographic data that would have to be addressed to the "passive side" of the word list. Of course, if the users proceed directly from a specific mother-tongue word to a foreign-language word with the necessary meaning differentiation, they can be more or less sure that the meaning or sense of the latter is the right one in the specific context and can be used to produce the L_2 text in question. Although this does not necessarily mean that the complete meaning of the L_2 word is expounded and, therefore, may contribute to a "mixed bilingualism", the partial meaning or sense provided through the L_1–L_2 word list is sufficient for the purpose of foreign-language text production. Furthermore, it is also relatively easy to include data providing information on orthography, pronunciation, gender and part of speech in an L_1–L_2 word list, and the same applies to synonyms, antonyms, hyponyms, collocations, idiom, proverbs etc. Foreign-language collocations, idioms and proverbs that are not direct translation equivalents of mother-tongue collocations etc. would, of course, be missing, thus contributing to a poorer language and style. Pragmatic data on L_2 could also easily be given in the L_1–L_2 word list, at least if they are presented in a codified way, but it would be much more complicated to include items providing cultural information related to L_2 in such a word list. However, when it comes to inflectional morphology and, especially, syntactic properties, the knowledge of which, together with the word, is the starting point for any text production, the problems become really big.

Even if the lexicographer — like Shcherba — who just wants to help users to produce "non-fictional" texts in L_2 in order "to be understood and to avoid appearing ridiculous" can, therefore, exclude some of the above-mentioned lexicographic data, it will still be very difficult to incorporate the rest of the relevant data in an L_1–L_2 word list. And not only that: as full equivalence between the words of two languages is rather the exception than the rule, the lexicographer would necessarily have to address two or more L_2 equivalents to quite a number of L_1 lemmata. If the above-mentioned data — even in a reduced version — were to be attached to each and every of these equivalents, it would lead to a heavily loaded dictionary article and a very complex and surely

not very user-friendly article structure. Furthermore, many L_2 translation equivalents of L_1 collocations would not contain the L_2 equivalents addressed to the lemma. This would require that much additional data should also be attached to all these collocations. And the same applies even more to idioms and proverbs as well as to synonyms, antonyms, hyponyms etc. One such example is the Danish *spis brød til* (literally: "eat bread at the same time"), which means something like "take it easy" or "calm down" in English. It would not be easy to include any further information on "take" and "calm" under the lemma *spise* (eat) in a Danish–English dictionary or to make references to another Danish lemma where users could find the needed information about the English word.

Hence, it can be concluded that an L_1–L_2 word list is not the most suitable aid to provide assistance with foreign-language text production. Of course, if one, as Shcherba, accepts the production of L_2 texts "without gross mistakes", i.e. with some mistakes, a word list of the above scope would be capable of supporting the production of texts in a foreign language. But apart from leading to some mistakes, it would also allow some degree of mixed bilingualism in terms of language assimilation. It is, therefore, evident that lexicography should search for another solution to the problem. And this solution can actually be found in lexicographic practice itself. Many dictionaries for a great number languages include a special list of verbs or even a second bilingual word list and provide, by means of an advanced distribution structure, data addressed to a specific lemma or an equivalent in other component parts of the dictionary. The following examples are from one such dictionary:

> **comer** [E1] *vi* **(a)** (en general) to eat; **no tengo ganas de**
>   ~ I'm not hungry…

Example 1: Dictionary article from Spanish–English word list in Oxford (2000)

> **eat** /iːt/ (*past* **ate**; *past p* **eaten**) *vt*/*i* comer…

Example 2: Dictionary article from English–Spanish word list in Oxford (2000)

infinitive/	past tense/	past participle/
infinitive	*pretérito*	*participio pasado*
…	…	…
dwell	dwelt, dwelled	dwelt, dwelled
eat	ate	eaten
fall	fell	fallen
…	…	…

Example 3: Excerpts from verb list in Oxford (2000)

If Spanish-speaking users, in the context of English text production, need information on the inflection of an English verb, what then is the role of the L_1–L_2 word list in the above-mentioned dictionary? It clearly serves as an index or

a "bridge" to the place (verb or word list) where the relevant data can be found. The L_1–L_2 word list barely functions as part of the outer access structure guiding the user first to the word list and then to the article where these data are placed according to the distribution structure chosen for the dictionary. In this specific case, it has no other function. Other dictionaries might, apart from the L_1–L_2 word list, only incorporate an L_2 or L_2–L_1 word list where data on inflectional morphology could be found, whereas others might only have a specific verb list of the above type. This is, for instance, the case with many dictionaries of Spanish where there are up to 58 relevant forms of a verb to be treated, something which is rather complicated to incorporate in a normal dictionary article. But what exactly is a list of verbs or irregular verbs? In the lexicographic literature, it is often referred to as back matter. But if it is analysed according to lexicographic principles, it is, in fact, a word list just as any other such list. It is the result of a selection based on a set of principles (being verbs or irregular verbs); it has an alphabetic macrostructure as most lexicographic word lists and a well-defined microstructure although this may vary from that of other types of word lists. If a reference work only consisted of one such list of verbs, it would therefore necessarily have to be typologised as a dictionary, i.e. a dictionary made for a special purpose.

This experience can be generalised: When users need a specific type of information about L_2 in order to solve a specific type of problem related to a specific type of user situation, and when they might not know the L_2 lemma where they can find this information, the data providing this information can be addressed to the lemmata of an L_2 word list and an L_1–L_2 index can then serve as a bridge giving access to these data. But how should such an L_1–L_2 index or bridge be designed?

The answer to this question once more depends on the methodology applied by users in the foreign-language learning process. If the contrastive method is used, the L_1–L_2 index should actually be a traditional L_1–L_2 word list where additional data have been placed in an L_2 list. However, if this method is not being used and the compiler of the dictionary tries to avoid any contribution to a mixed bilingualism, then there are two options. One is to try to combine Shcherba's vision of a "special type of dictionary" with the above considerations and the other is to conceive a completely new set of principles for such an index.

The problem here is evidently that Shcherba himself never developed his proposal in the necessary detail. However, in one of his writings on lexicography, Shcherba formulated five general principles for a Russian–French "special type of dictionary" designed for Russian users. Mikkelsen (1992: 27) summarises these principles as follows:

> 1. Provide a translation, not an explanation, that will, in the appropriate grammatical form, fit into a correct French sentence which has been translated from a Russian sentence. Of all candidates choose the one which fits into most of the

> Russian contexts. If no general equivalent is found, make sure that the intended user has enough information to judge which one will serve him best.
>
> 2. Reject the translations which are "too French" and metaphorical, and take only the simplest ones in order not to let the user seem ridiculous.
>
> 3. Throw away all approximate translations and synonyms — again in order not to make anecdotal translations possible.
>
> 4. If no precise equivalent is available, give the approximate ones together with the corresponding explanations. If not even an approximate equivalent can be found, leave the lemma as untranslatable, followed by an explanation in brackets. Under certain circumstances bring a translation of whole contexts.
>
> 5. Provide the necessary grammatical information, so that the user can produce the correct morphological and syntactic forms.

The first four points are interesting although it is evident that they would lead to a very simple foreign-language text without much stylistic variation and, as such, suitable only for a beginner's level (what exactly was Shcherba's idea). It is, however, the fifth point that presents the biggest problems. Firstly, it is not clear what exactly is meant by the "necessary grammatical information"? What kind of morphological and, especially, syntactic information does it include? Are collocations, for instance, part of this information? Secondly, as it has been argued above, data providing this kind of information are exactly of the type that is difficult to incorporate in an L_1–L_2 word list. Hence, although they include innovative ideas, there is no clear-cut and definitive solution to be found in Shcherba's lexicographic reflections.

The conclusion is therefore that a completely new set of principles for an L_1–L_2 index to an L_2 word list has to be developed. Such a set of principles could be as follows:

(a)   A list of L_1 words is selected.

(b)   Data providing information on part of speech is addressed to these words (for some languages also information on gender). This is done in order to help the user to be assured that he/she has arrived at the right word.

(c) L_2 equivalents are addressed to all L_1 words. These L_2 equivalents, at the same time, make up implicit references to the corresponding lemma in the L_2 word list.

(d)   If several L_2 equivalents are addressed to the same L_1 word, then meaning differentiation is provided. This is done in order to help the user to select the right L_2 word.

(e)   No more data is addressed to the L_2 equivalents. This is done in order to force the user to proceed directly to the L_2 word list.

(f) L_1 idioms are selected and attached to the corresponding L_1 words.

(g)   An L_2 translation equivalent is addressed to each L_1 idiom together with the L_2 lemma where further information about this equivalent can be found. This is done because the L_2 translation of an L_1 idiom does not

necessarily contain one of the other equivalents addressed to the corresponding L_1 word and in order to permit the user to proceed directly to the L_2 word where further information is to be found.

As can be seen, an L_1–L_2 index based on these seven principles contains very few data and, thus, provides little information on L_2. Although it can be used as a "reminder" for those users who already know and manage the L_2 word, but have just forgotten it, in most cases it would force users to proceed to the L_2 word list in order to find the information needed. In this way, the above proposal reduces the danger of a "mixed bilingualism" to a minimum. Of course, no guarantee can be given, as even the best dictionary is never immune to improper use. But this is mainly a question of dictionary culture.

5.2   Proposals for an L_2 word list

How should the L_2 word list of an L_2 text production dictionary be conceived? What kind of lexicographic data should it contain? In order to answer this question, it must be stated that the basic principle of dictionary conception always is — or should be — that no data is included because of "tradition" or the practice of existing dictionaries. It is not a question of what users expect to find in the dictionary due to an improper lexicographic culture, but what they actually need. The user-friendliness of a dictionary can be analysed in two different and, frequently, contradicting ways. The compiler of the dictionary can, for instance, observe how a hundred persons are using his dictionary and what problems they may have in using it. He can then develop a new model that is easier to consult, i.e. more user-friendly in terms of its *use*. But he can also consider how the dictionary functions, how it helps users and contributes to their needs, and then develop a new type of dictionary which is more helpful and fulfils the needs of the users even better than the old one, i.e. is more user-friendly in terms of its *effect*. Of course, in this last case, it might be necessary to "re-educate" users and teach them how to use the new dictionary properly. Users' real needs should therefore always be put in the centre of dictionary conception and planning. Each and every type of data included in the L_2 word list should be argued on the basis of the function of the dictionary, i.e. to assist users in solving problems related to foreign-language text production.

Mugdan (1992a) argues that the optimal lexicographic tool to assist users in foreign-language text production is a monolingual L_2 dictionary, but that a bilingual L_2–L_1 dictionary, for the benefit of its users, could also be designed to provide for this function, a solution rather close to Shcherba's "explanatory dictionary". In this vein, Mugdan proposes that the L_2 "free production" dictionary — apart from orthographic, phonological and semantic data — should contain information on inflection and syntax and several examples of how to use the words. This proposal, however, has to be concretised and other types of data added. But first, it should be stated that any well-designed dictionary nec-

essarily contains two different types of lexicographic data with totally different functions: the data intended to cover the users' primary needs, i.e. to solve their needs in a specific user situation, and the data intended to satisfy the users' secondary needs, i.e. to solve the problems related to the use of the dictionary (Tarp 2000). On this basis, the following proposal is presented, including — apart from the lemma — data on part of speech, gender, orthography, pronunciation, meaning, pragmatic and cultural notes, inflection, combining forms, syntactic properties, collocations, idioms, proverbs, stereotype phrases, text examples, synonyms, antonyms, hyponyms etc.

Part of speech. This item has two functions. Firstly, it is a means of identification — at least for the languages where the same grapheme might represent words with different parts of speech — in order to assure users that they have arrived at the right word. In this sense it covers the secondary user needs. Secondly, it provides some basic or general information on semantics, the inflectional paradigm and syntactic properties, i.e. primary user needs necessary for the purpose of text production.

Gender. This item might also have two functions: to identify the right word, at least in languages with two or more genders, and to provide basic information on the inflectional paradigm and some general indications of the syntactic properties of the word.

Orthography. This type of data is included both to confirm the arrival at the right word and to provide indications for written text production in the foreign language. Data on orthography are normally furnished implicitly through the lemma form and the inflectional paradigm. But if there are two or more competing orthographic forms of a given word, these should all be indicated in order to avoid any confusion whether or not the user has found the right word.

Pronunciation. This type of data also has two functions, i.e. to confirm that the user has found the right word and to provide indications for oral text production in the foreign language.

Meaning. In a production dictionary, data on meaning are exclusively provided with the purpose of confirming to users that they are about to use the right word. Users are already supposed to have an idea of what they want to express, i.e. they already know the meaning and just need to put words to it. As such, it is evident that the dictionary should expose all the different meanings and senses of the word in order to satisfy the users' needs in any user situation. The meaning of the L_2 words can be given in at least four different ways: L_2 explanations, L_1 explanations, L_1 equivalents and illustrations. Which of these ways are the most convenient will be discussed in section 6.

Pragmatic and cultural notes. Pragmatic notes are necessary in order to avoid users applying certain words out of context. And cultural notes are likewise necessary in order to avoid possible misunderstandings. The equivalent in Fang (a Gabonese language) of the English word *wedding* is, for example, *àluk*, but whereas *wedding* for most members of the English-speaking community

refers to a contract between two persons, for most native Fang speakers *àluk* would refer to a contract between two families. This type of cultural information should, of course, be available in the dictionary.

Inflection. All data about inflection are exclusively provided in order to serve the primary user needs, i.e. to assist text production. The inflectional paradigm might be incorporated directly in the L_2 article which often is the case in languages with a relatively reduced number of inflectional forms such as English, Danish and Afrikaans, or it can be placed in a separate verb list, which is frequently the case with languages with a considerable number of inflectional forms, such as Italian verbs or Icelandic nouns. In order to save space, the regular inflectional paradigms might be omitted in the individual dictionary articles and implicit or explicit references made to the outside matter of the dictionary where the regular inflectional paradigm is exposed in a special section. But the explicit incorporation, in one way or the other, of all irregular and even infrequent inflectional paradigms should be compulsory in any dictionary for text production.

Combining forms. It is evident that knowledge about word formation, especially affixes and compound words, is crucial in order to create a text in a foreign language. For languages, such as Spanish and Afrikaans, with a tradition of adding diminutives, augmentatives and other affixes to the nucleus word, this should, of course, be reflected in a production dictionary. In a similar way, for languages with many compound words, such as Danish, Afrikaans and German, it is important to furnish data on the building properties of each word and illustrate it with some examples. A classical example is the Danish word *barn* (child) which, apart from a totally obsolete genitive form (*e*), is combined both with its singular and plural forms (*barn* and *børn*) in a quite unpredictable way: *barnepige* (babysitter), *barnevogn* (perambulator), *børnetøj* (children's cloth), *børnehave* (kindergarten) etc. No foreign-language learner would be able to construct these and similar compound words in Danish without the corresponding lexicographic data.

Syntactic properties. This aspect is often ignored or very badly treated in dictionaries (and in lexicographic literature in general). Information on the syntactic properties of a given word is, however, even more necessary than word formation in order to create phrases in the foreign language. The syntactic properties should, therefore, be indicated for all the relevant words, and this should be done in an explicit way and not, as is the case with most dictionaries that supply such data, implicitly through a number of text examples. Such text examples are, of course, important in order to illustrate or even complement the explicit indications, but they should never stand alone because this requires that users themselves must be capable of generalising from the specific example to the general rule, something which might be difficult and sometimes not even possible. For instance, in a dictionary designed to assist foreign-language speakers in text production in English, it should be explained explicitly that the verb *tell* might, at least, be combined syntactically in the following ways:

- — tell something
- — tell somebody something
- — tell somebody (that) + clause
- — tell somebody to + infinitive
- — tell somebody what/how to + infinitive

Each of these explicit indications could then be followed by a number of text examples and not, as in Cambridge (1996), the other way round. The syntactic indications should, however, never be presented in such an abstract and complex way, as is the case in De Gruyter (1999):

> **1.1.** *etw.* (bes. *das*, *es*) ~
> **1.2.** *jmdm. etw.* (bes. *etw.*, *das*, *es*, Nebens. mit *dass*) ~
> **1.3.** *sich* ⟨Dat.⟩ *etw.* ⟨vorw. *das*⟩ ~
> **1.4**. *etw. über jmdn.*, *etw.* ~, *etw. von jmdm.*, *etw.* ~
> **1.5.** ⟨oft im Fragesats.⟩ *etw.* ⟨vorw. *was*⟩ *zu etw.* ⟨Dat.⟩ ~
> **1.6.** *etw. mit etw.* ~
> **1.7**. *gegen jmdn.*, *etw. etw.* ~
> **1.8.** *etw. zu jmdn.*, *etw.* ~
> **…**

Example 4: Excerpts from dictionary article in De Gruyter (1999)

The above minirules might, indeed, only be understood by linguists or people with a certain degree of linguistic training. The "normal" learner would most certainly get lost in this highly condensed and codified language. This article is, however, not the right place to give a detailed treatment of the very complex question of syntax in production dictionaries.

    Collocations and idioms. Collocations constitute an important part of the specific syntactic properties of many words, but are often quite unpredictable. The way the Danish word *bil* (car) collocates with the word *køre* (drive) is, for example, not the same as in English and Spanish. Whereas the English and Spanish equivalents of the Danish collocation *køre bil* are *drive a car* and *conducir a coche*, the corresponding equivalents of the collocation *køre i bil* are *go by car* and *ir en coche*, i.e. with another verb. The English and Danish equivalents of the Spanish *lavar los dientes* (literally: *wash the teeth*) are *brush one's teeth* and *børste tænder* and so on. As even these frequent and daily used collocations are unpredictable, any good dictionary for text production should provide such collocations.

    According to linguistic theory, collocations are characterised by transparency, but practice nevertheless shows that foreign-language speakers frequently have difficulties in understanding them, which means that an explanation or equivalent in the users' native language should be furnished. More or less the same applies to idioms that also constitute a considerable part of a given language. Hence, the most common idioms should also be included in a production dictionary. Although many linguist theories have been developed in order

to distinguish between idioms and collocations, most users do not know how to establish this distinction and would waste unnecessary time if idioms and collocations were placed in different sections or fields of the dictionary article. Even though many linguists may protest, because of user-friendliness, collocations and idioms should therefore be treated in the same way and at the same place in a dictionary for L_2 text production.

Proverbs. Although it is quite possible to make oneself understood without the use of proverbs, it should nevertheless be recommended that such word combinations are incorporated in text-production dictionaries, at least if they are conceived for more advanced foreign-language speakers. And they should be clearly explained or a mother-tongue equivalent should be given.

Stereotype phrases. A number of stereotype phrases used in specific contexts such as *How are you?*, *Have a nice weekend*, *Give my love to Jane*, etc., should also be included in the dictionary and explained.

Text examples. Many lexicographers — like Mugdan (1992a) — argue that text examples should be provided in dictionaries in order to illustrate how the words are used in concrete contexts. This argument is both reasonable and important, but the text examples should not be isolated from the other data included in the dictionary article and placed in a separate section or field. Quite the opposite, they should — as is practice in a number of dictionaries — be integrated with these data and placed in direct connection with them in order to show how they could be used in a concrete context. They could, for example, be placed after each syntactic "minirule" in order to expose this rule in a more concrete way. And they could also — especially in electronic dictionaries where space is not a problem — be presented in such a way that they illustrate the use of collocations and idioms.

Synonyms, antonyms, hyponyms etc. The only purpose of including synonyms, antonyms, hyponyms and similar words in a dictionary for text production is to assist users in producing text with a more varied vocabulary. Synonymy is not only interesting at word level but also for collocations, idioms and phrases as suggested by Jónsson (1999). If there is not complete synonymy or antonymy, items of meaning differentiation should be added in order to help users to select the right word in any context. As all the relevant morphological and syntactic data cannot be provided in the dictionary article where they are indicated as synonyms, antonyms or hyponyms, all such words should also be selected as lemmata, to which users can be referred for more information.

The above proposal about information categories in dictionaries for foreign-language text production is, of course, not the last word to be said in this regard. Everything has to be worked out in a far more detailed way and adopted to each set of languages and different categories of users. In this regard, it is evident that more or less data of the above-mentioned types could be included in the dictionary depending on the L_2 level of the user group. Other types of data might also be included, especially for more advanced users. But

in general, these are the most important lines along which a new generation of dictionaries for L$_2$ text production should be conceived.

6.   Integration with foreign-language reception

It is rather expensive and inconvenient to make dictionaries designed for only one function. This might be possible for languages with a large number of speakers, like English, but for "smaller" languages it would be financially more feasible for publishing houses to produce multifunctional dictionaries. And for users, it would clearly also be more convenient if the same dictionary could be consulted in order to solve problems related to more than one user situation. As will be seen in the following, it is fairly easy to combine foreign-language text production with foreign-language text reception within the framework of one and the same dictionary. Actually, a well-conceived production dictionary already contains most of the data necessary to satisfy users' needs in terms of text reception. Only a few additional data need be included in the dictionary in order to serve this function and these should in no way disturb the general structure and user-friendliness of the dictionary. But what data are needed to assist users in foreign-language text reception?

In order to answer this question, it should be remembered that lexico-graphic data may have two completely different functions: Firstly, they may cover the primary user needs, in this case assistance with text reception, and, secondly, they may satisfy the secondary user needs, i.e. guiding the user to the lemma where this assistance is given and helping to confirm that the lemma is actually the right one.

In real texts, foreign-language readers or listeners may find irregular, rare or even "surprising" inflectional forms they neither understand nor are able to relate to the basic form lemmatised. In English, for instance, this would be "old" plural forms such as *teeth* (*tooth*), *geese* (*goose*) and *feet* (*foot*) and irregular past tenses of verbs like *ate* which is even to be found under quite another letter than the infinitive form *eat*. In Afrikaans, it would include irregular past tenses such as *was* (*wees*: be), *sou* (*sal*: shall), *wou* (*wil*: will), *kon* (*kan*: can) and *moes* (*moet*: must), "strange" plural forms of nouns like *skepe* (*skip*: ship) and *stede* (*stad*: city) which a learner, at least at the beginner's level, would not be ex-pected to know.

As foreign-language speakers are not supposed to know all these inflec-tional forms and might have problems understanding them, it is evident that, in a dictionary for text reception, these irregular or strange forms should be selected as lemmata from where a reference could be made to the lemmatised form, for example the infinitive or singular form, where all the relevant data are to be found. The same applies to orthographic variants. The Spanish noun *yerba* (grass), which can still be found in a large number of texts, is, for exam-ple, a variant of the more frequent *hierba*. Few foreign-language speakers that have not come across these forms before would be able to relate these two orthographic variants that are placed far from each other in any alphabetic

organised dictionary. Both variants should, therefore, be selected for a text-reception dictionary, and the less frequent one could then be given the status of a reference lemma referring users to the other variant where the relevant data are provided.

Once users have found the lemma they have been looking for, another type of secondary user needs occurs as they need confirmation that the lemma is actually the right one. This confirmation is, first of all, furnished through orthographic and/or pronunciation data (for written and/or oral text reception). But as users might have come across an inflectional form or orthographic variant, they also need confirmation that these are actually included in the lemma in question and this is, once more, done by incorporating the relevant data in the dictionary article. However, as have been suggested in section 5, the inclusion of these same data has already been done for the purpose of text production. Hence, as regards the secondary user needs, the only data to be added to the dictionary in order to enhance it from an L_2 production dictionary to an L_2 production and L_2 reception dictionary are a number of reference lemmata whose inclusion disturbs neither the general structure nor the content of the dictionary.

As regards the primary user needs, it is evident that a dictionary for L_2 text production cannot provide understanding of whole texts as such. It can only provide assistance to understanding at word, collocation, idiom and proverb level as well as to finding the meaning of some stereotype phrases. In section 5, it has already been argued that the meaning of collocations, idioms, proverbs and stereotype phrases should be provided in a dictionary for L_2 text production. Hence, this kind of data would already have been included in a production dictionary and might therefore also serve the purpose of text reception.

The remaining question now is how to explain the meaning of the lemmata. In section 5 it was mentioned that data on meaning could be provided in at least four different ways, i.e., by means of L_2 explanations, L_1 explanations, L_1 equivalents and illustrations, or as a combination of these. L_2 explanations are actually provided in many monolingual learner's dictionaries, for example the English Big Five (Oxford, Cambridge, Longman, Cobuild and Macmillan), the German De Gruyter, Langenscheidt and Grundwortschatz, the Spanish Vox, the Swedish Lexin, the Afrikaans Basiswoordeboek and so on. However, none of these dictionaries are conceived for users of a specific language community, so the decision whether or not to choose L_1 explanations or equivalents was probably never considered. Nevertheless, all the above-mentioned dictionaries pose one big problem: How to understand the explanations?

> **handicraft** […] (**a**) an activity such as sewing or weaving, done with one's hands and requiring artistic skills: *handicraft classes*. (**b**) items made in this way: *a sale of handicraft(s)*.

Example 5: Dictionary article from Oxford (1998)

> **Panzer** […] auf Raupen (2) laufendes, vollständig mit Platten aus Stahl versehenes, geschlossenes und bewaffnetes Fahrzeug für den militärischen Kampf

Example 6: Dictionary article from De Gruyter

> **hoek** […] (1) 'n Hoek van iets is 'n punt of 'n gebied waar twee of meer van sy kante, lyne of vlakke bymekaar kom. Dit kan ook die gebied of ruimte wees naby die plek waar die twee of drie kante of vlakke bymekaar kom […] (2) 'n Hoek kan die afstand of ruimte wees tussen twee lyne of kante by die punt waar hulle aan mekaar raak […] (3) 'n Hoek is 'n gebuigde stuk metaal wat gebruik word om iets, veral vis, mee te vang: vishoek.

Example 7: Dictionary article from Basiswoordeboek

It goes without saying that users of Oxford should not only have reached a certain level of mastery of the English language but should also have a relatively developed capacity for generalisation should they be able to understand the above explanation and grasp the full meaning of the word. De Gruyter uses several difficult words and an unnecessarily complex syntax in order to explain a word representing something that everybody knows and that could be far more easily explained by an illustration. Basiswoordeboek uses a simple and rather stereotype definition language but even so it is still evident that a certain level of Afrikaans is needed in order to understand the explanations. Some of the dictionaries mentioned above, for example the English Cambridge, Longman and Macmillan and the Spanish Vox, have selected a reduced vocabulary of 2 000 or 2 500 words for their explanations whereas Oxford has a slightly bigger defining vocabulary of 3 500 words.

> **outreach** […] relating to efforts to bring government and a range of social services to people where they live or spend time * *an outreach worker* * *AIDS outreach programs bring medical care, condoms and counselling to prostitutes on the streets*.

Example 8: Dictionary article from Cambridge

> **trepar** […] **1** *intr*. Subir a un lugar alto y difícil ayudándose de los pies y de las manos: *los niños treparon al árbol más alto; treparon por la pared y saltaron al otro lado; el gato trepó hasta la copa del manzano.* ⇒ **encaramar**. **2** Crecer ciertas plantas subiendo y sujetándose a los troncos y ramas de los árboles, a las varas y a otros objetos: *la yedra trepaba por la tapia; puso unas cañas para que trepara la vid.* **3** *fam.* Con-

> seguir un puesto importante o una posición social alta
> usando todos los medios posibles: *ha amasado su*
> *fortuna trepando y adulando a los demás.*

Example 9: Dictionary article from Vox

This reduced vocabulary, of course, makes it much easier, but it still does not solve the problem for users at the very beginner's levels. It does not solve the problems where a certain capacity for generalisation is needed in order to grasp the full extent of an explanation. And, at the same time, it is really a question if all the explanations are indeed unequivocal.

> **table** […]
> 1 A **table** is a piece of furniture with a flat top that
> you put things on or sit at. ► *She was sitting at the*
> *kitchen table eating a currant bun* […] *I placed his*
> *drink on the small table at his elbow.* […]

Example 10: Dictionary article from Cobuild

This explanation could be interpreted in various ways. Of course, the additional text examples also help to get the right understanding of the word. But why complicate matters for users who just need a quick answer? Why not add something as simple as an illustration to the lemma as has been done in Longman, Cambridge Macmillan, Oxford and the German Grundwortschatz. There are actually quite a number of words that could be better explained in this way, for example *chair*, *door*, *cat*, *dog*, *hand*, *foot* etc. An illustration referring to a reality already known by users would in most cases be the best way to explain the lemma. However, in other cases — such as *wonder*, *reconsider*, *feel*, *confirm*, *interpret*, etc. — an illustration would hardly be able to expose the true meaning of a word and, even less, its various senses. In such cases, which probably make up the big majority, language is necessary in order express the meaning. It is, nevertheless, a question if an explanation in the foreign language would always be enough to enable users to grasp the exact meaning of the word in question. Beginners would definitely have problems, but even more experienced users would, at least sometimes, have doubts whether or not they have understood it right. Advanced users would, of course, have fewer problems but if lexicographers want to make a dictionary designed for users with different levels of mastery of the foreign language, they have to look for other solutions.

As already mentioned in section 2, Shcherba (1940) proposes a model for a "foreign defining dictionary" where the explanations are provided in the users' native language and equivalents only accepted in those cases "when this would simplify definition and would not be detrimental to a full understanding of the foreign word's true nature" (Shcherba 1940: 341). This solution is a tempting one and deserves much more attention.

Shcherba's mission was to produce a dictionary assisting with the assimilation of the foreign-language system, i.e. lexis and grammar, but there is little

doubt that this function — in Shcherba's concept — up to a certain degree is interwoven with L_2 text reception, as the dictionary was conceived as a consultation dictionary and not a "reading dictionary" according to the classification made by Hausmann (1977). Some of the most valuable comments on Shcherba's proposal, at least outside Russia, have been put forward by Duda (1986) and Duda et al. (1986). They were preparing a so-called "active" Russian–German dictionary for German users and, in this connection, they resumed the criticism directed against another Russian–German dictionary, made by Bielefeldt, where the latter had used Shcherba's principles for a "defining dictionary". According to them, Shcherba's idea of providing explanations of the meanings of foreign words in the users' mother tongue had proved problematic in practice (Duda 1986: 13. Translated by ST):

> The user is, apparently, capable of fulfilling the analysis of the word's meaning as it is described in a monolingual dictionary. But it seems to be far more difficult for him to name the meaning on the basis of a given description, i.e. to proceed to the lexicalisation.

On this basis, Duda (1986) concludes that Shcherba's principle of giving priority to the explanation should be changed and that priority should be given to the equivalent whenever this is possible and sufficient. According to Duda, this would not change the foundations of Shcherba's ideas. The above argumentation does not seem to give Shcherba justice. If users understand the meaning of a word, nothing more is required from a dictionary whose purpose is to assist them in foreign-language assimilation and text reception. The lexicalisation into the users' native language is, in fact, only necessary when the function of the dictionary is to assist the translation of foreign-language texts into this language. Duda (1986) and Duda et al. (1986) also criticize the fact that explanations and equivalents are not separated and claim that this could create problems for users. But as the following example shows, there are cases where the simultaneous provision of both an equivalent and an explanation is the best solution:

> **peso** […] $\boxed{5}$ (Fin) peso (*unit of currency of many Latin American countries*); **no tiene un** ~ he doesn't have a cent *o* penny

Example 11: Dictionary article from Oxford (2000)

The comments by Duda (1986) and Duda et al. (1986) are, nevertheless, highly valuable and should be taken into consideration in the future development of lexicographic theory. There is, for instance, no doubt that a cumulative selection of equivalents in the users' mother tongue in many cases would be sufficient in a dictionary designed for foreign-language text reception. With such a cumulative selection users would be able to pick up exactly that equivalent which provides meaning to a word within a concrete context. In this regard, it does not constitute any problem that a word chosen as an equivalent also might

have an additional sphere of meanings and senses not included in the lemma. When the user situation to be assisted is the assimilation of the foreign-language lexis, this might perhaps create a wrong idea about the real meaning of the foreign word. But would it necessarily do this? It goes without saying that a dictionary conceived for L_2 text reception should avoid contributing to any form of "mixed bilingualism" and, therefore, harm the overall language-learning process. This is, however, not the topic of this article and should be left to future theoretical discussions.

For the purpose of this article it suffices to conclude that whatever method is used to explain the lemma for the sake of foreign-language text reception is also sufficient for the sake of foreign-language text production. Although it might frequently be necessary to provide more extensive data for the purpose of text reception, these data will in no way conflict with the data needed for text production as the latter constitute a genuine subset of the former.

To summarise this section: The above discussion has indicated that the additional data — in terms of reference lemmata and explanations — needed to adapt a foreign-language production dictionary to the purpose of foreign-language text reception, will in no way disturb its character as a production dictionary. On the other hand, whether or not the data included in the dictionary for the purpose of production and irrelevant for the purpose of reception will reduce its usefulness as a reception dictionary, will primarily depend on the internal distribution structure in the dictionary article and the corresponding search-field structure. Users who only need assistance with text reception should, of course, be able to proceed directly to the relevant data without passing through a lot of irrelevant, time-consuming data. This, however, is only a question of knowing the problems and choosing the most appropriate structures.

7.   Conclusions

The overall method used in this article was first to analyse and describe the dictionary functions and then to discuss the relevant categories of lexicographic data. This is the opposite method of the one used in most theoretical works about lexicography, for example Herbst (1999) and Zöfgen (1994), where the lexicographic data is discussed without any relation to specific functions. The former method points to a science-based lexicographic theory while the latter leads to a kind a theory that hardly can provide the necessary guidelines for the future high-quality dictionaries that modern society needs.

Through the combined methods of analysis-synthesis and induction-deduction, this article has taken some fundamental steps towards a general theory of dictionaries, whether printed or electronic, designed to assist the users in foreign-language text production and reception. It is evident that there are still important questions to be dealt with, such as the users' specific needs at different levels of foreign-language mastery, how to expose the syntactic prop-

erties, what are the benefits of the different ways of explaining the words and how to structure the lexicographic data in correspondence with the specific user needs. It is also evident that some of the categories treated in this article, up to a certain point, are language-specific and that others might be added for other languages, especially outside the Indo-European language family.

Another important question is to develop a method through which monolingual foreign-language production and reception dictionaries can be given a bilingual dimension. This is what has been done with the Swedish Lexin which, in spite of a highly codified and difficult metalanguage, is probably one of the most interesting learner's dictionaries in the world. This dictionary was originally conceived as a monolingual learner's dictionary for foreign immigrants in Sweden and then it was "translated" and specifically adopted to the needs of several of the most important emigrant languages in Sweden (see Gellerstam 1999). The generalisation of this idea would not only make it much easier and cheaper to produce dictionaries adapted to the specific language speakers' specific needs, but would, as such, also point to a much higher quality product than the actual monolingual dictionaries for foreign-language speakers in general.

Literature

Dictionaries

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

Cambridge 1996 = **Procter, Paul (Ed.).** 1996. *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press.

Cobuild 2001 = **Sinclair, John (Ed.).** 2001. *English Dictionary for Advanced Learners*. London: HarperCollins.

Dansk grundordbog 1994 = **Dollerup, Cay and Inge Padkær Nielsen.** 1994. *Dansk grundordbog. Genvej til det danske sprog*. Copenhagen: Høst & Søn.

De Gruyter 2000 = **Kempcke, Günter (Ed.).** 2000. *De Gruyter Wörterbuch Deutsch als Fremdsprache*. Berlin/New York: Walter de Gruyter.

Grundwortschatz 1983 = **István, Kosasras.** 1983. *Grundwortschatz der deutschen Sprache.* Berlin: Volk und Wissen Volkseigener Verlag.

Langenscheidt 1993 = **Götz, Dieter, Günther Haensch and Hans Wellmann (Eds.). In collaboration with Vincent J. Dochertz and Günther Jehle.** 1993. *Langenscheidts Großwörterbuch Deutsch als Fremdsprache. Das neue einsprachige Wörterbuch für Deutschlernende.* Berlin/Munich: Langenscheidt.

Lexin 1993 = **Gellerstam, Martin (Ed.).** 1993. *Lexin. Språklexikon för invandrare. Svenska ord — med uttal och förklaringar*. Uppsala: Norstedts Förlag.

Longman 1995 = **Summers, Della.** 1995. *Longman Dictionary of Contemporary English*. Harlow: Longman.

Macmillan 2002 = **Rundell, Michael (Ed.).** 2002. *English Dictionary for Advanced Learners*. London: Macmillan.

Oxford 2000 = **Carvajal, C.S. and J. Horwood (Eds.).** 2000. *The Pocket Oxford Spanish Dictionary. Spanish–English English–Spanish.* Oxford: Oxford University Press.

Oxford 1998 = **Crowther, Jonathan (Ed.).** 1998. *Oxford Advanced Learner's Dictionary.* Oxford: Oxford University Press.

Vox 1995 = 1995. *VOX Diccionario para la enseñanza de la lengua española.* Alcalá de Henares/Barcelona: Universidad de Alcalá de Henares/Biblograf.

Theoretical contributions

Bergenholtz, Henning and Sven Tarp. 2002. Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen. *Lexicographica* 18: 253-263.

Bergenholtz, Henning and Sven Tarp. 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 171-196.

Bergenholtz, Henning and Sven Tarp. 2004a. The Concept of "Dictionary Usage". Dollerup, Cai (Ed.). *Worlds of Words. A Tribute to Arne Zettersten. Nordic Journal of English Studies* 3(1): 23-36, 2004.

Bergenholtz, Henning and Sven Tarp. 2004b. Wörterbuchfunktionen. Barz, Irmhild, Henning Bergenholtz and Jarmo Korhonen (Eds.). 2004. *(Lerner-)Lexikografie: Ein- und zweisprachig für Deutsch und Finnisch.* Frankfurt am Main/Bern/New York/Paris: Peter Lang.

Duda, Walter. 1986. Ein «aktives» russisch–deutsches Wörterbuch für deutschsprachige Benutzer? Günther, Erika (Ed.). 1986. *Beiträge zur Lexikographie slawischer Sprachen*: 9-15. Linguistische Studien. Reihe A. Arbeitsberichte. Akademie der Wissenschaften der DDR. Zentralinstitut für Sprachwissenschaft 147. Berlin: Akademieverlag.

Duda, W., M. Frenzel, E. Wöller and T. Zimmermann. 1986. *Zu einer Theorie der zweisprachigen Lexikographie. Überlegungen zu einen neuen russisch–deutschen Wörterbuch.* Linguistische Studien. Reihe A. Arbeitsberichte. Akademie der Wissenschaften der DDR. Zentralinstitut für Sprachwissenschaft 142. Berlin: Akademieverlag.

Gellerstam, Martin. 1999. LEXIN — lexikon för invandrare. *LexicoNordica* 6: 3-18.

Hausmann, Franz Josef. 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher.* Tübingen: Max Niemeyer.

Herbst, Thomas and Kerstin Popp (Eds.). 1999. *The Perfect Learners' Dictionary (?).* Lexicographica. Series Maior 95. Tübingen: Max Niemeyer.

Jónsson, Jón Hilmar. 1999. Fraseologiens plass i ordbøker for innvandrere. *LexicoNordica* 6: 65-78.

Kromann, Hans-Peder, Theis Riiber and Poul Rosbach. 1984. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. Wiegand, Herbert Ernst (Ed.). 1984. *Studien zur neuhochdeutschen Lexikographie V. Germanistische Linguistik* 3-6: 159-238. Hildesheim/New York: Olms.

Kromann, Hans-Peder, Theis Riiber and Poul Rosbach. 1991. Principles of Bilingual Lexicography. Hausmann, Franz Josef, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta (Eds.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*: 2711-2728. Handbücher zur Sprach- und Kommunikationswissenschaft 5.1–5.3. Berlin/New York: Walter de Gruyter.

Mikkelsen, Hans Kristian. 1992. What Did Shcherba Actually Mean by "Active" and "Passive" Dictionaries? Hyldgaard-Jensen, Karl and Arne Zettersten (Eds.). 1992. *Symposium on Lexico-*

graphy V. Proceedings of the Fifth International Symposium on Lexicography May 3–5, 1990, at the University of Copenhagen: 25-40. Tübingen: Max Niemeyer.

Mugdan, Joachim. 1992a. Zur Typologie zweisprachiger Wörterbücher. Meder, Gregor and Andreas Därner (Eds.). 1992. *Worte, Wörter, Wörterbücher: Lexikographische Beiträge zum Essener Linguistischen Kolloquium*: 25-48. Tübingen: Max Niemeyer.

Mugdan, Joachim. 1992b. On the Typology of Bilingual Dictionaries. Hyldgaard-Jensen, Karl and Arne Zettersten (Eds.). 1992. *Symposium on Lexicography V. Proceedings of the Fifth International Symposium on Lexicography May 3–5, 1990, at the University of Copenhagen*: 17-24. Tübingen: Max Niemeyer.

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

Tarp, Sven. 1992. *Prolegomena til teknisk ordbog*. Ph.D. Dissertation. Århus: Spansk Institut. Handelshøjskolen i Århus. <http://www.lng.hha.dk/dml/spa/phd.pdf>.

Tarp, Sven. 1994. Funktionen in Fachwörterbüchern. Bergenholtz, Henning and Buchard Schaeder (Eds.). 1994. *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*: 229-246. Tübingen: Gunter Narr.

Tarp, Sven. 1995. Wörterbuchfunktionen: Utopische und realistische Vorschläge für die bilinguale Lexikographie. Wiegand, H.E. (Ed.). 1995. *Studien zur zweisprachigen Lexikographie mit Deutsch II*: 17-51. Lexicographica. Series Maior 110. Hildesheim/New York: Olms.

Tarp, Sven. 1998. Leksikografien på egne ben. Fordelingsstrukturer og byggedele i et brugerorienteret perspektiv. *Hermes, Journal of Linguistics* 21: 121-137.

Tarp, Sven. 1999. Lørnerordbøger for indvandrere og andet godtfolk. *LexicoNordica* 6: 107-132.

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

Tarp, Sven. 2002. Translation Dictionaries and Bilingual Dictionaries – Two Different Concepts. *Journal of Translation Studies* 7: 59-84.

Tarp, Sven. 2003. Basic Elements of Lexicographic Theory/Éléments de base de la théorie lexicographique. Emejulu, James Duplessis (Ed.). 2003. *Éléments de lexicographie gabonaise. Tome II*: 7-35. New York: Jimacs-Hillman.

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

Zöfgen, E. 1994. *Lernerwörterbuch in Theorie und Praxis: Ein Beitrag zur Metalexikographie mit besonderer Berücksichtung des Französischen*. Lexicographica. Series Maior 59. Tübingen: Max Niemeyer.

A New English–Arabic Parallel Text Corpus for Lexicographic Applications

Hashan Al-Ajmi, *Department of English, Kuwait University, Kuwait,*
(hashan98@yahoo.com)

Abstract: Bilingual lexicographers, translation specialists and English teachers in the Arab world do not have access to computerized corpora of parallel texts for the English–Arabic language pair. This project has been carried out to meet this requirement by establishing the first general parallel corpus of English texts and their Arabic translations. The first phase of the project involved the selection of general source texts having appropriate lexical and stylistic features. The chosen source texts deal with a variety of topics such as the environment, globalization, psychology, history, politics, drama, etc. Their Arabic translations were taken from *The World of Knowledge* series published by the National Council for Culture, Arts and Letters (NCCAL) in Kuwait.

Keywords: PARALLEL CORPUS, LEXICOGRAPHY, TRANSLATION, BILINGUAL DIC-TIONARY, COLLOCATIONS, ALIGNMENT, SYNONYMS, DERIVATIVES, ANTONYMS, GLOS-SARY, FREQUENCY

Opsomming: 'n Nuwe Engels–Arabiese parallelletekskorpus vir leksiko-grafiese toepassings Tweetalige leksikograwe, vertaalkundiges en Engelsonderwysers in die Arabiese wêreld het nie toegang tot gerekenariseerde korpusse van parallelle tekste vir die Engels–Arabiese taalpaar nie. Hierdie projek is onderneem om in dié behoefte te voorsien deur die eerste algemene parallelle korpus van Engelse tekste en hul Arabiese vertalings tot stand te bring. Die eerste fase van die projek het die keuse van algemene brontekste behels wat geskikte leksikale en stilistiese eienskappe besit. Die gekose brontekste handel oor 'n verskeidenheid onderwerpe soos die omgewing, globalisering, psigologie, geskiedenis, politiek, drama, ens. Hul Arabiese vertalings is geneem uit *The World of Knowledge*-reeks gepubliseer deur die National Council for Culture, Arts and Letters (NCCAL) in Koeweit.

Sleutelwoorde: PARALLELLE KORPUS, LEKSIKOGRAFIE, VERTALING, TWEETALIGE WOORDEBOEK, KOLLOKASIES, OOREENSTEMMING, SINONIEME, AFLEIDINGS, ANTO-NIEME, GLOSSARIUM, FREKWENSIE

1. Introduction

Parallel corpora ought to be treated as essential tools in bilingual lexicographic activities in the Arab world where translation plays a key role in the transfer of knowledge, especially from English sources. Yet, Arab translation specialists and training programs lack corpus-based bilingual dictionaries and lexical data

bases. Unfortunately, and despite the growing awareness of their importance, no parallel corpora for the English–Arabic language pair have yet been developed. This is partly due to the lack of the necessary programs to compile such resources and the funding authorities' doubts and uncertainty regarding the effectiveness of parallel corpora. There are also those who confuse this type of corpus with translation memories which are widely used by professional translators and by developers of machine translation systems and term banks. The available literature on parallel corpora outlines several applications of this type of corpus including:

— Development of bilingual dictionaries:

  English–Arabic dictionaries are considered to be inefficient lexical sources as they rely on the provision of non-contextual translation equivalents. They force the user to undertake most tasks such as the selection of the appropriate equivalent and ordering the words in the target language phrases or sentences. Here, the parallel corpus can play a vital role by providing several appropriate context-based equivalents that can be included in a bilingual dictionary (cf. Dickens and Salkie 1996).

— Helpful information for users:

  Users of the corpus can see both source and target words in their contexts and be their own judges when selecting the appropriate synonym — an advantage typically unavailable in the use of bilingual dictionaries.

— Language learning:

  Many language teachers around the world have already started to integrate and utilize parallel corpora in the foreign-language curriculum. Many of them have noticed that the parallel corpus contributed in making learners more independent, e.g. by making comparisons, understanding how different contexts lead to differences in meaning.
  Also, parallel corpora have been used in a variety of ways such as learning by discovery, enhancement of vocabulary, preparation of teaching materials, and understanding of translation problems and techniques (Baker 1993, Barlow 1996, St. John 2001, Zanettin 1994).

In the following sections, the English–Arabic parallel text corpus will be described in terms of its sources, the programs used in compiling and searching it, its components, and its uses in the improvement of bilingual dictionaries and the development of collocations dictionaries.

2.   Sources

The aim was to find original English texts characterized by both currency and

variety of topics to ensure that the corpus will provide new and modern vocabulary and usage. Therefore, the focus has been on those texts written quite recently, e.g. in the nineties. Publications of the Kuwait National Council for Culture, Arts and Letters were considered to be the appropriate resource. These publications enjoy great popularity among educated readers in the Arab world, especially the *World of Knowledge* series of translated books that explore a wide variety of topics and are written by the best translators in the region. It is believed that Arabic translations in the higher stylistic level of this diglossic language (cf. Ferguson 1972) are necessary in this corpus because most translations from English are of scientific, legal, political or economic texts with similar stylistic levels. Accordingly, these translations would make the corpus a practical tool in translator training.

3.   Means of access: Restricted web site

To ensure that a larger number of users can access it, the corpus has been allocated a URL on the Kuwait University server. But due to copyright restrictions online users of the corpus will need to enter their user names and passwords. Students taking lexicography and translation courses at the Department of English use this web site in various course activities such as the extraction of bilingual word lists and the study of translator equivalents versus dictionary equivalents.

4.   Corpus processing and search tools

The source and target texts were scanned, converted to text files through an OCR (optical character recognition) program and saved as XML (extensible markup language) files. Another software program aligned the converted texts on the sentence level and each textual unit was assigned an alignment type and identification. Header files stored in two separate corpora were used to link the English and Arabic texts. The search tool employed in this project is the al-Idrisi search program developed by Sakhr Software. This program was chosen for its capability of handling Arabic morphology and syntax. Its search abilities cover exact matches, root-based searches as well as derivatives, synonyms, antonyms and recognition of Arabic affixes. It also uses wild card searches, employs relevancy ranking and ignores common Arabic errors, e.g. *hamza*, *ha/ta* endings entered by the user or present in the original text. It searches for all or any of the words or phrases.

5.   Components

The corpus is composed of a number of HTML links for a glossary of English words, and a search page as well as a help page.

5.1   The English glossary

This glossary is actually an edited index of all words in the English part of the corpus. It includes word frequencies and is used as a reference tool for looking up words and knowing their frequencies directly. This is done by selecting the first letter of the needed word, then pressing the second letter to show all the words in that section with their frequencies. The user of the glossary is subsequently capable of accessing the results page by clicking on the chosen hypertext word which will highlight its English results.

5.2   Search page

This page contains several options allowing users to find results of a specific nature. For example, users can specify a source text or subject such as economics, arts, psychology, etc. The search options here include the search language, which makes it possible to search for identical words or phrases and for both Arabic and English synonyms as well as derivatives which can be the focus of morphological studies of English and Arabic. Users can utilize the corpus as either a monolingual or a bilingual corpus. They are also able to choose the number of needed results which can range between ten and fifty. They can opt to ignore Arabic common errors such as *hamza* and the *ya* ending. The corpus has search capabilities similar to those available in Google such as searching for all or any of the words or phrases, but it is also able to search for Arabic words with their affixes, e.g. the attached definite *al* or pronouns which characterize Arabic. Users can access more text of any given result by clicking on 'more' to see the whole sentence highlighted in its wider context on an HTML page.

6.   Corpus-based exercises

The corpus has already been used in a number of course activities in translation and lexicography at the Department of English. One of these involved comparisons between the collocations provided in the *Oxford Collocations Dictionary for Students of English* (2002) and the collocations in the corpus. It was found that the corpus contains about 25% of the OCD collocations but many corpus collocations have not been listed in the dictionary. Of course, the corpus was useful in providing Arabic equivalents for these collocations and, therefore, it can be a good source for a corpus-based bilingual dictionary of this type. In another exercise the corpus was compared with the *Al-Mawrid English–Arabic Dictionary* (2003) indicating that about 10 000 Arabic equivalents and senses are missing in the dictionary. And in a further exercise comparisons were made between frequency ordering in the corpus and sense ordering in some English–Arabic dictionaries. The comparison has revealed that many entries need to be reordered in order to match users' expectations when dealing with living texts in translation and reading.

The corpus has other capabilities especially in the field of bilingual lexicography, e.g. obtaining illustrative examples in one or both languages.

7.   Conclusion

It is hoped that this parallel corpus will meet the needs of researchers, translators and English teachers in the Arab world. However, given its small size (three million words) it should at this stage be regarded as a prototype for a larger corpus which is planned to be compiled in future. The envisaged version would be about ten times the current size. It will be bi-directional with more textual genres and styles, and it will be processed with POS (part of speech) tags to improve the search options in both directions.

Acknowledgement

References

Dictionaries

Ba'albaki, M. 2003. *Al-Mawrid: A Modern English–Arabic Dictionary*. Dar El-Ilm: Lil-Malayin.
Crowther, J. 2002. *Oxford Collocations Dictionary for Students of English*. Oxford: Oxford University Press.

Other references

Baker, M. 1993. Corpus Linguistics and Translation Studies — Implications and Applications. Baker, M., G. Francis and E. Tognini-Bonelli (Eds.). 1993. *Text and Technology*: 233-250. Philadelphia: John Benjamins.
Barlow, M. 1996. Parallel Texts in Language Teaching. Botley, S., J. Glass, T. McEnery and A. Wilson (Eds.). 1996. *Proceedings of Teaching and Language Corpora 1996*: 45-56. UCREL Technical Papers 9. Lancaster: University of Lancaster.
Dickens, A. and R. Salkie. 1996. Comparing Bilingual Dictionaries with a Parallel Corpus. Gellerstam, M., J. Järborg, S.G. Malmgren, K. Norén, L. Rogström and C. Röjder Papmehl (Eds.). *EURALEX '96 Proceedings I–II*: 551-559. Gothenburg: Department of Swedish, Göteborg University.
Ferguson, C. 1972. Diglossia. Cashdan, A. et al. (Eds.). *Language in Education: A Course Book*: 38-45. London: Routledge and Kegan Paul.
St. John, E. 2001. A Case for Using a Parallel Corpus and Concordancer for Beginners of a Foreign Language. *Language Learning and Technology* 5(3): 185-203.
Zanettin, F. 1994. Parallel Words: Designing a Bilingual Database for Translation Activities. Wilson, A. and T. McEnery (Eds.). 1994. *Corpora in Language Education and Research: A Selection of Papers from Talc94*: 99-111. UCREL Technical Papers 4. Lancaster: University of Lancaster.

Etymological Aspects of Idiomatic and Proverbial Expressions in the Lexicographic Development of Sesotho sa Leboa — A Semantic Analysis[*]

V.M. Mojela, *Sesotho sa Leboa National Lexicography Unit, University of the North, Polokwane, Republic of South Africa (mojelav@unorth.ac.za)*

Abstract: Idiomatic and proverbial expressions are important components of the oral tradition of Sesotho sa Leboa, and therefore a knowledge of the literal meaning of words as they appear in dictionaries without inclusion of their figurative meaning seems to be a shortcoming. An idiom or a proverb possesses one basic meaning, i.e. the meaning to which the idiom or proverb is basically meant to refer, but each idiom or proverb is made up of several lexical items. Each of these lexical items has its own meaning, which usually differs from the figurative sense of the idiom or proverb. Even though the meaning of the words in an idiomatic or proverbial expression seems to differ from the sense of the idiom or proverb, there is to a certain extent a relationship. It is this relationship which lexicographers can assist to explain in their definitions in order to clarify both the literal and the figurative meanings of words in Sesotho sa Leboa.

This article aims to stress the importance of having specialized dictionaries which will give users detailed etymological explanations of the meaning of idiomatic and proverbial expressions as used in Sesotho sa Leboa. The etymological analysis of the meaning of these lexical items (idioms and proverbs) will provide a better understanding of these expressions and enrich dictionaries with detailed definitions. This will create a better understanding of the relationship between the literal meanings of the expressions and their real (i.e. figurative) meanings.

Keywords: DIACHRONIC ANALYSIS, ETYMOLOGY, FIGURATIVE EXPRESSION, IDIOM, LEMMATIZATION, LEXICOLOGY, METAPHOR, METAPHORIC EXPRESSION, PROVERB, SYNCHRONIC ANALYSIS, TERMINOGRAPHY.

Opsomming: Etimologiese aspekte van idiomatiese en spreekwoordelike uitdrukkings in die leksikografiese ontwikkeling van Sesotho sa Leboa — 'n semantiese ontleding. Idiomatiese en spreekwoordelike uitdrukkings is belangrike komponente van die mondelinge tradisie van Sesotho sa Leboa en gevolglik skyn die letterlike betekenis van woorde soos hulle in woordeboeke voorkom sonder insluiting van hul figuurlike betekenis, 'n tekortkoming te wees. 'n Idioom of spreekwoord besit een basiese betekenis, d.i. die betekenis

---

waarna die idoom of spreekwoord bedoel is om basies te verwys, maar elke idioom of spreek-
woord bestaan uit verskeie leksikale items. Hierdie leksikale items het elk sy eie betekenis wat
gewoonlik verskil van die figuurlike betekenis van die idioom of spreekwoord. Selfs al skyn die
betekenis van die woorde in 'n idiomatiese of spreekwoordelike uitdrukking te verskil van die
idioom of spreekwoord, is daar in sekere mate 'n verwantskap. Dit is hierdie verwantskap wat
leksikograwe kan help om in hul definisies te verklaar ten einde sowel die letterlike as die figuur-
like betekenisse van woorde in Sesotho sa Leboa te belig.

Die doel van hierdie artikel is om die belangrikheid te beklemtoon van die besit van gespesi-
aliseerde woordeboeke wat gebruikers uitvoerige etimologiese verklarings sal gee van idiomatiese
en spreekwoordelike uitdrukkings soos in Sesotha sa Leboa gebruik. Die etimologiese ontleding
van hierdie leksikale items (idiome en spreekwoorde) sal 'n beter begrip van hierdie uitdrukkings
verskaf en woordeboeke met uitvoerige definisies verryk. Dit sal 'n beter begrip van die verhou-
ding tussen die letterlike betekenisse van die uitdrukkings en hul werklike (d.i. figuurlike) beteke-
nisse bewerkstellig.

Sleutelwoorde: DIACHRONIESE ONTLEDING, ETIMOLOGIE, FIGUURLIKE UITDRUK-
KING, IDIOOM, LEKSIKOLOGIE, LEMMATISERING, METAFOOR, METAFORIESE UIT-
DRUKKING, SINCHRONIESE ONTLEDING, SPREEKWOORD, TERMINOGRAFIE

1.   Introduction

Idioms and proverbs are figures of speech found in most, if not all languages
around the world. They all have one thing in common, i.e. they use figures of
speech to express figurative meanings in their culture, and in most cases these
figures of speech include idioms and proverbs. These idioms and proverbs are
important carriers of culture in the communities.

The main objective of this article is to highlight the importance of the ety-
mological definition of idioms and proverbs in lexicography, especially with
reference to specialized dictionaries for Sesotho sa Leboa and other African
languages, as a means of preserving the African cultures embedded in these
expressions. Like all other languages, Sesotho sa Leboa has idioms and prov-
erbs embedded within the culture of the Basotho community and these expres-
sions are used, *inter alia*:

—   to explain a wider meaning with a very short and concise expression
     designed in the form of metaphor, and

—   to compare the present with the past — thereby keeping alive the cul-
     ture, or the potential, almost forgotten culture of the community.

2.   Definitions: Idioms and Proverbs

The terms *idiom* and *proverb* are defined and described by various linguistic and
literary scholars such as Svensén (1993) and Gouws (1999) in several different
ways. These scholars include, *inter alia*, lexicographers and lexicologists, whose

main perception is directed towards the possible lemmatization and the inclusion of these figures of speech in dictionary corpora and linguistic and terminographic dictionaries. Gouws (1999: 72), for instance, says the following concerning idioms:

> As fully-fledged lexical items, idioms represent the lexicalization of specific semantic values. Because this lexicalization is a language specific process two languages will not necessarily have idioms to convey the same meaning.

Svensén (1993: 108) defines the term *idiom* as follows:

> A fixed group of words with a special meaning which is different from the meanings of the individual words.

Idioms are, according to Svensén (1993: 109) "fixed combinations", not only lexically but also to some extent grammatically, and "must therefore be shown in full in the dictionary".

3.   Etymological Analysis

The etymological definition of lexical items such as idioms and proverbs will bring a great improvement in the lexicographic development of Sesotho sa Leboa if figurative meaning is analysed more comprehensively. The available dictionaries are only concerned with the figurative meaning of idioms and proverbs without any regard for the significance of the literal meaning of these expressions. This means that dictionary users will only find the ultimate meaning of the lexical item without any knowledge of the background of the expression. In this regard Makkai (1969: 50) distinguishes between two "idiomaticity areas": "The area on the lexemic stratum will be called the *first idiomaticity area*, or the area of *polylexonic lexemes*; and that on the sememic stratum may be termed the *second idiomaticity area*, or the area of *polysemonic sememes*." This article therefore seeks to highlight the importance of taking both the literal and the figurative meanings of these expressions into consideration in order to supply the user with comprehensive definitions of these lexical items. This can be done through the compilation of a special encyclopedic dictionary which will include etymological definitions of these lexical items. This type of dictionary should have definitions concerning:

—   the basic meaning of the expression,

—   the relationship between the literal meaning of the expression and its real (i.e. figurative) meaning, and

—   the etymological background (meaning) of the individual words in an expression.

Etymology will play a major role in the determination of the relationships

between the literal and figurative meaning of the lexical items contained in this special type of dictionary. Etymological information will be interpreted according to the description of the term *etymology* as given by scholars such as Svensén (1993: 189):

> Information about the etymology of words tells us their history: how they were formed and evolved and finally took the shape and meaning they have in the language of today. Etymological facts lie along the time axis, and cut straight across the other information categories, combining elements from all of them: the formal, combinational, and semantic properties of words, and the things and events in the world outside language, all make their contribution.

This means that the dictionary should answer the following questions concerning an expression such as **Go ya ga maotwana hunyela** 'to die':

— What is the meaning of this expression?

— What is the literal meaning of this expression?

— Why did this expression get this meaning?

— What is the meaning of the words **maotwana** and **hunyela**?

— How are these words related to 'death'?

If we were therefore to answer these questions regarding the given expression, the following would be relevant answers:

— The meaning of this idiomatic expression is 'to die', which is completely different from its literal meaning derived from the literal meaning of the individual words which constitute it.

— The literal meaning would be 'to go to the place called **maotwana hunyela**'.

— As far as the etymological aspect underlying the origin of this idiomatic expression is concerned, it acquired its meaning from the traditional method used by the Northern Sotho people when burying their dead. The corpse is put in an upright position in the grave with the knees up against the chest (**go hunyela)** facing west with the whole body covered by the skin of a beast slaughtered specifically for this purpose. The beast's meat, called **mogoga**, is eaten without salt immediately after the burial.

— The word **hunyela** means 'to shrink', or in this case, 'to bend or to squeeze the body in a sitting position in the grave with the chin leaning on the knees'. **Maotwana** (literally meaning 'small feet') refers to the feet of a corpse which are 'shrunken up' in the grave before being covered by soil.

4.   Semantic (Metaphoric) Analysis

The example discussed above confirms the fact that idioms and proverbs use known cultural words in the form of metaphor to express a secondary meaning(s). The following metaphoric comparison is usually heard when a speaker of Sesotho sa Leboa uses a proverb to confirm the truth of what he/she says. In most cases the speaker will say: **Mogologolo o boletše a re: ya se rage e tla tlatša kgamelo**, or **Basotho ba boletše ba re: ya se rage e tla tlatša kgamelo.**

   This shows that these expressions did not originate in the recent but in the distant past. The semantic analysis of these figurative expressions shows that the meaning of the idioms and proverbs can be classified into three stages, i.e.:

—   Stage 1: A literal meaning which makes little or no sense, as becomes clear from the above-mentioned proverb, e.g.

   Ya se rage e tla tlatša kgamelo
   'If it doesn't kick, it will fill the milk container.'

—   Stage 2: A literal meaning which has sense even though it will not be the real meaning of the idiom or proverb, e.g.

   Ya se rage e tla tlatša kgamelo
   'If the cow which is being milked doesn't kick away the milk container and spill milk, the milking process will go on well and the container will ultimately be filled with milk.'

—   Stage 3: Which is the real (or figurative) meaning of the idiom or proverb, which is in most cases different from the literal meaning of these expressions, e.g.

   Ya se rage e tla tlatša kgamelo
   'If there are no disturbances, we will have positive results.'
   'If things go well, great things will happen.'

The first and second stages in the definition of these expressions will seem to be meaningless and unimportant to speakers of Sesotho sa Leboa who will always concentrate on their real or ultimate (figurative) meaning without thinking of the importance of discovering how the expressions came into existence, which is, as a matter of fact, the cultural root of the meaning of the idiom or proverb. The literal definition of the expression will help to restructure the etymology of the idiom or proverb in the sense that the user will get to know the cultural background as to how, and why, the idiom or the proverb originated.

   The etymological definition of the proverb **Ya se rage e tla tlatša kgamelo** will show that its real meaning, explained in Stage 3 above, is metaphorically compared to the milking of a cow. This real (or figurative) sense has nothing to do with milking cows, or milk containers. In this way, familiar objects or events within the Basotho culture are used metaphorically in the form of idioms or

proverbs which acquire figurative meaning. It is, therefore, familiar to and customary for the Basotho people to own cattle as a major livestock, and to enjoy one of the major 'fringe benefits' of owning cattle, i.e. milk.

The etymological definition of the above-mentioned proverb will help those non-mother-tongue speakers of Sesotho sa Leboa, who only know its real meaning, also to understand its real background. For instance, there are many foreign speakers of Sesotho sa Leboa who know the meaning of the proverb as it is found in dictionaries and other publications such as Rakoma's *Marema-ka-dika*, but who do not know how a cow is milked. This is true of those people who are from highly developed mostly urban areas where milking is done by machines, and who may not know the manual system which is part and parcel of the culture of rural Basotho people. To these people it will not be clear how the cow can kick the milk container. It is however, possible for these people to learn to speak Sesotho sa Leboa and to know the real meaning of this proverb without being familiar with the cultural background from which it originated.

The following analysis of the meaning of the proverb, **Monna ke nku, o llela teng,** and the idiom, **go ba le pelo e botse,** explains the importance of:

— the metaphorical comparison of the literal and the figurative meanings of the idioms and proverbs,

— the etymological (background) knowledge of the idioms and proverbs, and

— the background knowledge of the culture of the Basotho people, in order to get a full understanding of these expressions.

(a)   **Monna ke nku, o llela teng**

Literal meaning 1: 'A man is a sheep, he cries within himself.'
Literal meaning 2: 'A man is like a sheep, he doesn't cry publicly when he is in trouble.'
Real meaning: 'A real man is strong and courageous, he doesn't expose his problems publicly.'
'A real man doesn't go around telling people his problems.'
'One cannot predict the real man's emotional status by his facial appearance.'

The etymological definition of this proverb will answer the following questions:

— Why is **monna** 'man' metaphorically compared to a sheep?

— What is the characteristic of a sheep as explained in this proverb, which a man possesses?

— Are all the users of this proverb aware of this characteristic?

— What is the meaning of the term **monna** 'man' as used in this proverb?

The etymological explanation, or description, of this proverb will show that the word **monna** does not refer to an ordinary man, but to the supposed characteristic of a normal man, which is metaphorically compared to that of a sheep. There are many people who know the real or even the literal meaning of this proverb, but who cannot answer question number one. There are also many Sesotho sa Leboa language users who may not be aware of this specific characteristic of a sheep, namely that a sheep, as compared to some other animals like goats and cattle, will remain quiet even when it is being slaughtered. This silence of a sheep when it is in trouble or in difficulty is metaphorically compared to the behaviour of a real man, who will not tell anyone about his problems. The word 'real man' is used here since not all men have this characteristic. According to the Sotho custom, men who do not have this characteristic are referred to as **basadi** 'women'. That is why we usually hear people saying:

Lebeko le se ke la mmotša, ke mosadi yola

Literal meaning: 'Do not tell Lebeko, he is a woman.'
Real meaning: 'Do not tell Lebeko because he will go around telling other people.'
'Do not tell Lebeko because he has no courage.'
'Do not tell Lebeko the bad news because he might collapse.'
'Do not tell Lebeko because he behaves like a woman.'

Or:

Makikele le ka mmotša, ke monna yola

Literal meaning: 'Makikele, you can tell him, he is a man.'
Real meaning: 'Makikele can be told the news because he can keep the problem to himself.'
'Makikele can be given the news because he has courage.'
'Makikele can be given the news because he has the qualities of a real man.'

(b)   **Go ba le pelo e botse**

Literal meaning: 'To have a good heart.'
Real meaning: 'Not being cruel or cunning but being kind and merciful.'

Most people know the real (figurative) meaning as well as the literal meaning of this idiom, but very few of them can trace its origin. The etymological approach to idioms like this one, especially those using organs of the body, will show the Sotho tradition assigning various functions (or responsibilities) to the various organs of the body. For instance, the heart is regarded as a controlling factor in the life of a person:

—   The heart helps one to make decisions. That is why one usually hears a Sotho person saying: **Pelo ye nngwe e re ke mo rage, ye nngwe ya re aowa, mo tlogele,** 'One heart says I should kick him, the other one says no, leave him.'

— Doing bad things is believed to be the work of a person's heart, that is why we hear the Sotho people saying: **O na le pelo e mpe,** 'He has a bad (evil) heart (He is too bad (evil))'.

The Basotho people regard the heart as a guiding factor as far as security is concerned. This is reflected in statements such as the following:

— **Ke rile ge ke swanetše go sepela pelo ya ka ya gana.**
'When I was supposed to go my heart refused.'

In this way, by stopping a person from making the journey, the heart prevents him from possible danger. These are some of the traditional beliefs which are part and parcel of the Basotho culture associated with the role of the human heart.

5.   Conclusion

This article aimed to reveal the importance of etymology in the semantic description or the definition of proverbial and idiomatic expressions in Sesotho sa Leboa. This means that the background source of the meaning of these expressions can be rediscovered if the lexicographic approach to idioms and proverbs can also consider the importance of etymological analysis of these figurative expressions. It is important to have this background knowledge underlying idioms and proverbs recorded by lexicographers in the form of specialized encyclopedic dictionaries of idioms and proverbs.

References

General

Gouws, R.H. 1999. *Equivalent Relations in Translation Dictionaries*. M.Phil. Study Guide. Stellenbosch: University of Stellenbosch.
Makkai, A. 1969. The Two Idiomaticity Areas in English and their Membership: A Stratificational View. *Linguistics* 50: 44-58.
Svensén, Bo. 1993. *Practical Lexicography: Principles and Methods of Dictionary-making*. Oxford: Oxford University Press.

Dictionaries

Kriel, T.J. 1988. *Popular Northern Sotho Dictionary*. Pretoria: J.L. van Schaik.
Kriel, T.J. and E.B. van Wyk. 1989. *Pukuntšu*. Pretoria: J.L. van Schaik.
Lombard, D.P. et al. 1991. *Sediba*. Pretoria: Via Afrika.
Louwrens, L.J. 1994. *Dictionary of Northern Sotho Grammatical Terms.* Pretoria: Via Afrika.
Rakoma, J.R.D. 1970[5]. *Marema-ka-dika tša Sesotho sa Lebowa*. Pretoria: J.L. van Schaik.
Ziervogel, D. and P.C. Mokgokong. 1975. *Comprehensive Northern Sotho Dictionary*. Pretoria: Van Schaik.

The Transliteration Principle: Is this the Best Procedure in African Language Lexicography and Terminology?

Motlokwe Clifford Mphahlele, *Sesotho sa Leboa National Lexicography Unit, University of the North, Polokwane, Republic of South Africa (motlokwem@unorth.ac.za)*

Abstract: The creation of target language equivalents in bilingual dictionaries have always been a challenge to both lexicographers and terminologists. Lexicographers, terminologists, subject specialists and linguists face a challenging task of supplying term equivalents for foreign international languages. The lack of a sufficient technical vocabulary in the African languages poses challenges to lexicographers and terminologists when supplying technical equivalents. They often regard transliteration as the quickest lexicographical and terminological procedure without taking the practical demands of users into account. This impedes the optimal retrieval of semantic information by the target users of dictionaries. This article discusses possible reasons why transliteration should not be regarded as the first but the last resort in lexicography and terminology. It proposes different lexicographical and terminological procedures lexicographers and terminologists can use to supply accurate and appropriate translation equivalents without making excessive use of transliteration. In this way a scientific language will develop which could assist users to communicate successfully in the mother tongue.

Keywords: TERMS, TERM EQUIVALENTS, TERMINOLOGISTS, SUBJECT SPECIALISTS, LINGUISTS, INTERNATIONAL LANGUAGES, INTERNATIONAL SCIENTIFIC LANGUAGES, UNDERDEVELOPED LANGUAGES, TRANSLITERATION, TERMINOLOGICAL PRINCIPLE, TRANSLATION EQUIVALENTS, SEMANTIC INFORMATION, BORROWING, TARGET LANGUAGE.

Opsomming: Die transliterasiebeginsel: Is dit die beste werkwyse in die leksikografie en terminologie van die Afrikatale? Die skep van doeltaalekwivalente in tweetalige woordeboeke was altyd 'n uitdaging vir sowel leksikograwe as terminoloë. Leksikograwe, terminoloë, vakspesialiste en taalkundiges staan voor 'n uitdagende taak om termekwivalente vir vreemde internasionale tale te verskaf. Die gebrek aan 'n voldoende tegniese woordeskat in die Afrikatale stel probleme aan leksikograwe en terminoloë by die voorsiening van tegniese ekwivalente. Hulle beskou transliterasie dikwels as die vinnigste leksikografiese en terminologiese werkwyse sonder om die praktiese eise van die gebruikers in ag te neem. Dit belemmer die optimale herwinning van semantiese inligting by die teikengebruikers van woordeboeke. Hierdie artikel bespreek moontlike redes waarom transliterasie nie as die eerste nie, maar as die laaste uitweg in leksikografie en terminologie beskou moet word. Dit stel verskillende leksikografiese en termi-

nologiese werkwyses voor wat leksikograwe en terminoloë kan aanwend om juiste en gepaste vertaalekwivalente te voorsien sonder om buitensporige gebruik van transliterasie te maak. Op hierdie manier sal 'n wetenskaplike taal ontwikkel wat gebruikers kan help om suksesvol in die moedertaal te kommunikeer.

Sleutelwoorde: TERME, TERMEKWIVALENTE, TERMINOLOË, VAKSPESIALISTE, TAAL-KUNDIGES, INTERNASIONALE TALE, INTERNASIONALE WETENSKAPLIKE TALE, ONDER-ONTWIKKELDE TALE, TRANSLITERASIE, TERMINOLOGIESE BEGINSEL, VERTAALEKWI-VALENTE, SEMANTIESE INLIGTING, ONTLENING, DOELTAAL

1.   Introduction

The creation of target language term equivalents for new foreign concepts is not as easy as it may seem. This creation of terms should not be done in a haphazard way but systematically according to the general terminological principles. Lexicographers, terminologists, subject specialists and linguists face a challenging if not difficult task of supplying term equivalents for foreign international languages. Underdeveloped languages such as African languages largely lack a technical vocabulary and this necessitates the speedy development of a technical terminology. As English enjoys worldwide recognition as a language of international scientific and technical communication and also due to the fact that many scientific texts are published in English, it is often encounterd as source language in terminology. Therefore, African language terminologists, lexicographers, subject specialists and linguists have to create accurate new term equivalents in their mother tongue for the multitude of terms in English. Trying to cope with this challenge, they apply transliteration as a lexicographical and terminological principle. They often regard transliteration as the quickest lexicographical and terminological procedure without taking into consideration the practical demands of the target users. This therefore impedes an optimal retrieval of semantic information in the target language. The prescriptive nature of the terminological work should not tempt terminologists and their collaborators to supply term equivalents not having any meaning and conceptual equivalence with the source language terms.

The purpose of this article is to propose possible ways to prevent the excessive use of transliteration. Transliteration should not be regarded as the first but the last resort in terminology creation. Different terminological procedures terminologists can use to supply accurate and appropriate translation equivalents that could assist users to communicate successfully in scientific language in their mother tongues will be discussed.

This article aims to show lexicographers, terminologists, subject specialists and linguists how appropriate and accurate term equivalents could carry the same meaning and concept of the original scientific term without the excessive use of transliteration. Duplication of the stem of the source language term (transliteration) cannot, in many cases, assist target users to retrieve semantic

information unless such a source language term has already been established in the target language. This is why transliteration often fails target users because the foreign concept is not easily transmitted into the target language.

2.   Preventing the excessive use of transliteration

Transliteration can be defined as the use of a foreign language stem to coin a target language equivalent. In most countries, English has emerged as a *lingua franca* and also as a language of choice. This has therefore made African language speakers to adopt foreign languages by means of borrowing, neologising, transliterating, etc. The remainder of this article will focus on transliteration as an optional terminological procedure.

As target language speakers, African language lexicographers, terminologists, subject specialists and linguists are tempted to supply a transliteration in a case where there is lack of a target language term equivalent. In most cases, this terminological procedure has not succeeded in assisting terminology users to retrieve the required information. The lack of target language term equivalents entails zero equivalence, i.e. lack of equivalents in the target language. In cases where zero equivalence occurs, many languages have coined and are still coining target language equivalents from foreign words. "Zero equivalence entails a lexical gap," says Gouws (1999: 26). "That is, the lexical item that is supposed to be supplied as a translation equivalent is not present. The absence of a translation equivalent usually indicates the existence of a lexical gap in the target language." In terminology, it would be appropriate to speak of the existence of a terminological or conceptual gap if a ready term equivalent is absent in the target language. According to Al-Kasimi (1977), a major problem confronting the bilingual lexicographer is that the required equivalents in the target language cannot always be found. This is also the case with terminologists, subject specialists and linguists: they cannot always find ready term equivalents in the target language. Al-Kasimi differentiates between two types of vocabulary that contribute to this problem, i.e. culture-bound words and scientific and technological terminology. The fact that a lexicographer cannot find the required equivalents does not mean that borrowing must be utilised (cf. Mphahlele 2001: 27). In cases where zero equivalence occurs, the lexicographer should bear in mind that direct borrowing from the source language cannot serve any semantic purpose. That is, a borrowed word which has not been established in general language use cannot assist the user to communicate successfully (Mphahlele 2001: 27). This is also the case in terminology. When the stems of the source language terms are retained as target language equivalents (as in transliteration), there is no way in which a user could be assisted to comprehend the source language term, unless such a term has been established in the target language.

With regard to transliteration, Roets (2001: 9) observes: "A large number of equivalents are formed through transliteration, particularly in subjects such as

Economics, Medicine and Physics. However, this does not mean that translit-
eration is the best solution. Although English is the language of transit in South
Africa for most scientific and technical subjects, the English language has a
limited use for young African learners." According to Roets, transliteration
should not be regarded as the first solution when supplying target language
equivalents. This is true, because most African learners do not understand Eng-
lish words. In this case, if transliteration is excessively used in terminology
there is no way in which learners could understand the stems of the English
words retained as possible term equivalents. In other words, new target lan-
guage term equivalents would remain foreign even when transliterated. This
means that foreign concepts are not understood in the target language. Trans-
literation could only serve a semantic purpose if the source language terms are
known or well established in the target language. If not, the transliterated
source language forms could only remain foreign term equivalents without any
meaning or concept attached to them in the target language. This procedure
therefore fails the terminology user. The Northern Sotho examples in Table 1
illustrate this point:

Table 1

formula	– fomula
fraction	– frakšin
equilibrium	– ekhwilipramo
pictogramo	– pictokramo

If the new Northern Sotho Mathematics terminology is designed in this way,
the target users would not understand the source language terms. In the above
examples, the English mathematics term **formula** is supplied with a Northern
Sotho transliteration **fomula**. The question here is whether a native speaker of
Northern Sotho has ever heard of an English mathematical term **formula**. If
not, then this transliterated Northern Sotho term cannot assist users in under-
standing the meaning and contextual use of the source language term **formula**.
This is also the case with the other English mathematics terms **fraction**, **equi-
librium** and **pictogram**. If users have not heard of these terms before, translit-
eration could therefore not serve any semantic purpose. That is, in most cases
transliteration does not assist users to retrieve the required information re-
garding the source language terms. Transliteration could only be opted for if
speakers of the target language know or have heard of the source language
terms. If this is not the case, transliteration would therefore just be a mere
duplication of the source language terms. If speakers of the target language do
not know the foreign concept or referent, terminologists should make sure that
the concept is transmitted into the target language by appropriate terminologi-
cal procedures. With regard to these, Roets (2001: 10) gives general guidelines
regarding the acceptance of terminology. According to her, the new terminolo-
gy is acceptable if it has concept correspondence, adheres to orthography, is
appropriate, i.e. keeps to the register of the main corpus of the dictionary, is not

in any way taboo or offensive to any gender or culture group, shows conformity with or similarity to other entries, has aptness and easy identification with the specific subject field, etc. In other words, term equivalents should have correspondence with the original concept, and display linguistic correctness and accuracy by adhering to the orthography. With transliteration it cannot be claimed that the term equivalents have concept correspondence. In most cases, new terms are incorporated into the orthography of a language after being considered as established borrowings in that language. This means that there is no way in which the newly transliterated equivalents can be claimed to adhere to the orthography of a language, because they have not been used in written language before. Although terminologists, subject specialists and linguists are permitted to have a prescriptive approach when supplying term equivalents, this does not necessarily mean that the needs of the target users should be lost sight of. In other words, some general guidelines regarding terminology should be taken into consideration so that appropriate and accurate term equivalents could be supplied.

Given that transliteration does not adhere to most of the general guidelines regarding terminology, it could not be regarded as a balanced terminological procedure. With regard to this Roets (2001: 10) says, "Do not grasp transliteration as a first, but a last resort — if all else fails. No matter if an equivalent may sound and appear like a brief explanation of a term." She believes it is better if terminologists and subject specialists give a brief meaning definition of the source language term as an equivalent. A brief meaning definition of the source language term could assist users to retrieve the required information because a meaning definition carries a concept in itself. This would be far better than the mere listing of a transliteration as term equivalent.

If transliteration is adopted as the first terminological procedure, the technical language of the developing African languages cannot evolve. Excessive use of transliteration may result in heavy borrowing because all speakers of a target language would use the new transliterated terminology. The technical terminology for African languages should be developed so that speakers of these developing languages could use their own terminology for an international technical source language when communicating. The new terminology should coincide with corresponding concepts so that speakers of the developing language could benefit from using international technical language in their mother tongue. In terminology, the source language terms are standardised by means of definitions for each term. Exact communication is possible if there is a one-to-one relation between term and concept. These definitions assist terminologists and their collaborators to supply appropriate term equivalents. With transliteration, definitions supplied for standardisation are sometimes ignored. According to Cluver (1978: 91), a technical dictionary is a bi- or multilingual dictionary which includes the standardised terms of a particular subject field, with or without a full definition for each term. When supplying term equivalents, terminologists and their collaborators should read and understand the

supplied definitions so that their term equivalents could be accurate. Alberts (1990) also describes a multilingual technical dictionary as including an alphabetical list of terms in the source language and equivalents in several languages. According to her, this alphabetical list of terms is standardised for creating better term equivalents.

3.   Towards new independent term equivalents

Roets (2001: 1) states that, for a language to develop, the daily growth of human knowledge in the science and technology of developed countries needs to be shared with developing countries and, to this end, expressed in mother tongue terminology. According to her, this is also important as the status of a language depends on its possession of a technical vocabulary. The technical vocabulary of a particular language develops when technical terms are created in that particular language. New term equivalents in a target language would enable users to learn about science and technology in their mother tongue. If a language has enough technical vocabulary, its users would be able to communicate by means of technical language. According to Roets (2001: 1), the South African terminographer's mission is to stay informed of emerging international developments in all fields of learning and to translate these from a language of transit into mother tongue equivalents for expressing the vast number of new concepts that have to be mastered for technical communication. If international scientific terms are translated into African languages, speakers of African languages would be able to learn and communicate in scientific languages.

Avoiding the excessive use of transliteration, lexicographers, terminologists, subject specialists and linguists should always strive for new independent term equivalents having a one-to-one relation with the source language terms. The new term equivalents should adhere to the orthography of the target language and also have concept correspondence with the source language terms (Roets 2001: 10). These new terms should be the result of terminological, scientific and linguistic decisions taken by collaborators.

The Northern Sotho examples in Table 2 can be considered as illustration:

Table 2

formula	– mokgwataetšo
fraction	– palopalwana
equilibrium	– mokgwatekatekano
pictogram	– tshwantšhotšhupakgopolo

These suggested term equivalents display a balanced terminological procedure. When looking for term equivalents, the terminology user would be able to deduce the meaning of the source language terms because there is a one-to-one meaning relation between the source language forms and the target language equivalents. In this case the terminology user would achieve communicative

equivalence. Because of their self-explanatory nature the supplied term equivalents have transparent meanings and are therefore easily comprehensible. As can be seen, the supplied term equivalents are independent from the original English terms, i.e. the stems of the source language terms have not been retained or borrowed when creating term equivalents. This is what terminologists, subject specialists and linguists should actually do when creating term equivalents. Transliteration should only be opted for if all measures to supply term equivalents have failed.

Rather than adopting the transliteration principle as a lexicographical and terminological procedure, lexicographers, terminologists and their collaborators should try to create new term equivalents having a meaning relation with the source language terms. This would help developing the technical vocabulary of a language.

The examples from Alberts (2001: 2) in Table 3 display a better terminological procedure:

Table 3

English	Afrikaans	Northern Sotho
common law	gemene reg	molao wo o sego wa ngwalwa
common (general) law	gemene reg	molao wa bohle nageng
common law	gemene reg	molao wa setlwaedi
common law marriage	gemeenregtelike huwelik	lenyalo la molaotlwaedi
common law wife	gemeenregtelike vrou	mosadi wa semolao

In legal terminology, the above presentation of term equivalents could enable users of Afrikaans and Northern Sotho to retrieve the required information. In this case, interpreters and translators are enabled to achieve communicative success. Because of its multilingual nature, the above terminology would serve the dual purpose of assisting Afrikaans and Northern Sotho interpreters and translators.

According to Sager (1990: 57), special languages strive to systematise principles of designation and to name concepts according to pre-specified rules or general terminological principles. Using general terminological principles, terminologists, subject specialists and linguists could name concepts in an accurate and systematic way. Sager (1984: 317) says that "terminology is strongly involved in the creation of new designations and is therefore interested in a theory of active, productive term-formation". In creating new terms, terminologists and their collaborators should also take the level (referential skills) of target users into consideration so that the new terminology created could be suitable for and comprehensible to them. If research regarding the target users of terminology is not done, it is therefore likely that term equivalents may be too difficult for the target users. The excerption of terms together with their term equivalents should be done in accordance with the level of target users so that the created terminology could assist them in retrieving the required information. The Terminology Division of the National Language Service (NLS) have

designed different terminology worklists for terminologists and their collaborators so that term equivalents could be suitable for and comprehensible to different levels of learners.

Table 4 shows page 1 of the SCHOOL PROJECT: MATHEMATICS GRADE 1–7 Worklist 1 (National Language Service: 1 September 2001):

Table 4

TARGET LANGUAGE: ……………………………………………………..…
1. Please supply translation equivalents in the target language.
2. Add synonyms where applicable (use comma to separate).
3. Please translate the example sentence.
NOTE: Do not translate the *Definition* at this stage.

abacus – counting frame

abbreviate definition: to shorten a word by leaving out some of its letters or sounds. Example sentence: The centimeter is abbreviated to cm, and ml is the abbreviation for milliliters. (**shorten**)
<term>
<example sentence>

abbreviation definition: A shortened form of a word, phrase or symbol
<term>
<example sentence>

about adv. Definition: Nearly: not exactly, but nearly or roughly. Example sentence: From head to toe, I am about 1 metre tall. (**approximately**)
<term>
<example sentence>

above <preposition>{context: space and position} Definition: A grammatical word indicating a position directly overhead, on top of, or higher than something.
<term>
<example sentence>

abstract Definition: Not relating to concrete objects but expressing something that can only be appreciated intellectually
<term>
<example sentence>

acceptable Definition: considered to be satisfactory.
<term>
<example sentence>

accessory Definition: An additional item that may be fitted to something to perform an extra function or enhance performance. Example sentence. A pencil is necessary accessory for a pair of compasses. (optional part)
<term>
<example sentence>

Terms shown in the above table were excerpted by terminologists who provided them with definitions so that accuracy and concept correspondence could be easy in the target language. These definitions enable terminologists, subject specialists and linguists to supply appropriate term equivalents. Supplying transliteration equivalents for the above terminology would undermine the work done by terminologists. The supplied example sentences (in the source

language) are meant to assist collaborators of terminologists to give term equivalents having concept correspondence with the source language terms.

4.   Conclusions

As becomes clear from the previous discussion, supplying term equivalents for international technical languages is not an easy task for terminologists. In their terminological work, lexicographers, terminologists, subject specialists and linguists should try to adhere to the general terminological principles available. They should try to move away from a dependent approach in the formation of term equivalents and try to be productive and creative. For supplying term equivalents the following conclusions could be useful to terminologists:

— Transliteration should not be regarded as the first but the last solution after all measures for supplying term equivalents have failed.

— The new term equivalents should have concept correspondence, adhere to the orthography of the target language (i.e. display linguistic correctness), show aptness and be easily identifiable with the subject field.

— Transliteration does not necessarily assist in the development of technical language, i.e. if transliteration is used excessively, the technical vocabulary of the target language is not allowed to develop.

— Because the stems of source language terms are retained in the target language, transliteration may result in heavy borrowing.

— If the referent or concept is unknown to speakers of the target language, transliteration could not serve any semantic purpose because even the transliterated term equivalents may still be unknown to the terminology user. In other words, with transliteration a referent or concept may still be foreign to speakers of the target language. In this case, the user is not assisted in retrieving the required information.

— Transliteration could only be opted for if speakers of the target language know the referent or concept. In this case, the retained stem may display concept correspondence with the original source language term.

— Lexicographers, terminologists, subject specialists and linguists should always strive for term equivalents that display accuracy and dependence of the orthography of the source language.

— For a better and balanced terminological practice, definitions supplied for the source language terms should be considered because they serve a standardisation purpose. That is, the supplied definitions assist terminologists, subject specialists and linguists to supply appropriate term equivalents having a one-to-one relation with the source language terms.

Bibliography

Alberts, M. 1990. *'n Bepaling van Afrikaanse vakleksikografiese behoeftes*. D.Litt. et Phil. Thesis. Pretoria: UNISA.

Alberts, M. 2001. *Terminology: Principles and Practice*. Unpublished Document. Pretoria: DACST.

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

Cluver, A.D. de V. 1978. Die Terminologie. Van Schalkwyk, C. et al. (Comp.). 1978. *Diploma in Vertaling*: 67-131. Pretoria: UNISA.

Gouws, R.H. 1999. *Equivalent Relations in Translation Dictionaries*. M.Phil. Study Guide. Stellenbosch: University of Stellenbosch

Mathematics Terminology Worklist. 2001. School Project: Mathematics Grade 1–7. Pretoria: DACST.

Mphahlele, M.C. 2001. *A Model to Achieve Communicative Equivalence in Translation Dictionaries*. Unpublished M.Phil. Thesis. Stellenbosch: University of Stellenbosch.

Roets, S. 2001. *Preliminary Stages of a Terminographic Project*. Unpublished Article. Pretoria: Department of Arts and Culture.

Sager, J.C. 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. Lexicographica. Series Maior 1. Tübingen: Max Niemeyer.

Sager, J.C. 1990. *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins.

The Contribution of Missionaries to Shona Lexicography[*]

Advice Viriri, *Department of African Languages and Culture, Midlands State University, Gweru, Zimbabwe (adviriri2002@yahoo.co.uk)*

Abstract: The article discusses the developments in Shona lexicography during the colonial era and the contribution made by the missionaries in general and Father M. Hannan in particular which later resulted in the on-going process of compiling monolingual dictionaries by the African Languages Research Institute (ALRI) (into which the ALLEX project has been transformed). The missionaries employed various methods that did not only signal the beginning of an economically exploitative relationship between "the West and the rest of us" but also had ancillary cultural consequences (Dathorne 1975: 3). Their motives towards the development of African literature in general and Zimbabwean lexicographic work in particular were primarily evangelical and not to further creative writing. This caused Father Hannan to experience problems not only in his translation of the *Shona Bible* but also in his *Standard Shona Dictionary*. Hannan is used as an example because his translation problems were typical of those experienced by all missionaries. Hannan's dictionary was a welcome new development using the orthography introduced in 1931 and revised in 1955. The aim of his dictionary was "to record Shona words in Standard Shona spelling" (Hannan 1959: ix). Hannan's dictionary has made a considerable contribution to standardising Shona orthography.

Keywords: LEXICOGRAPHY, MISSIONARY, COBUILD, SHONA DICTIONARY, MONO-LINGUAL DICTIONARY, BILINGUAL DICTIONARY, CULTURE, TRANSLATION, STANDARD ORTHOGRAPHY, COLONIALISM, CULTURAL IMPERIALISM, SHONA PROVERBS, AFRICAN LANGUAGES RESEARCH INSTITUTE (ALRI), INFLUENCE.

Opsomming: Die bydrae van sendelinge tot die Sjonaleksikografie. Die artikel bespreek die ontwikkelinge in Sjonaleksikografie gedurende die koloniale tydperk en die bydrae gemaak deur die sendelinge in die algemeen en Vader M. Hannan in die besonder, wat later uit-geloop het op die voortgaande proses van die samestelling van eentalige woordeboeke deur die African Languages Research Institute (ALRI) (waarin die ALLEX-projek omskep is). Die sendelinge het verskillende metodes gebruik wat nie net die begin van 'n ekonomies uitbuitende verhouding tussen "the West and the rest of us" ingelui het nie, maar ook bykomende kulturele gevolge gehad het (Dathorne 1975: 3). Hul motiewe betreffende die ontwikkeling van Afrikaliteratuur in die algemeen en Zimbabwiese leksikografiese werk in die besonder was primêr evangeliserend en nie om skeppende skryfwerk te bevorder nie. Dit het veroorsaak dat Vader Hannan probleme onder-vind het nie net by sy vertaling van die *Shona Bible* nie, maar ook in sy *Standard Shona Dictionary*. Hannan word gebruik as voorbeeld omdat sy vertaalprobleme tipies was van dié wat alle sende-

linge ervaar het. Hannan se woordeboek was 'n welkome nuwe ontwikkeling wat die ortografie gebruik wat in 1931 ingevoer en in 1955 hersien is. Die doel van sy woordeboek was "to record Shona words in Standard Shona spelling" (Hannan 1959: ix). Hannan se woordeboek het 'n aansienlike bydrae gelewer tot die standaardisering van die Sjonaortografie.

Sleutelwoorde: LEKSIKOGRAFIE, SENDELING, COBUILD, SJONAWOORDEBOEK, EENTALIGE WOORDEBOEK, TWEETALIGE WOORDEBOEK, KULTUUR, VERTALING, STANDAARD ORTOGRAFIE, KOLONIALISME, KULTURELE IMPERIALISME, SJONA-SPREEKWOORDE, AFRICAN LANGUAGES RESEARCH INSTITUTE (ALRI), INVLOED

Introduction

Africa's colonial experience had far-reaching implications. Not only did it create disorientation, alienation and economic domination but it also paved the way for cultural invasion, where the imperialists were authors and actors in the process in which the colonised were the subjects. This cultural hegemony sought to serve the interests of imperialism and was a form of cultural domination. For cultural imperialism to succeed, the colonised were to be convinced of their inferiority and the unquestionable superiority of the invaders.

The development of orthographies ushered in an epoch of literary translation that marked the beginning of African literature. Missionaries were the pathfinders and wayfarers of a literary tradition in Zimbabwe. The missionaries' main aim was the development of a Shona writing system, later devised by Doke and from 1931 onwards promoted as "Union Shona" (Chimhundu 1992: 79). Their contribution towards the development of African literature, especially vernacular literature and lexicography cannot be overemphasised. A.M. Chirgwin, a missionary secretary visiting Rhodesia, observed (Chirgwin 1932: 327):

> The books that are in greatest demand are Bibles, hymn-books and catechisms. These are regarded by the people as so clearly a part of the necessary apparatus of a Christian that they purchase them without demur. *The Pilgrim's Progress* enjoys a steady sale in almost every African vernacular into which it has been translated.

The missionaries' motives were primarily evangelical, the intention being to convert Africans and becoming "the shepherds while Africans were the sheep" (Chiwome 1996: 11). While this did not give impetus to creative writing, previously unwritten African languages were reduced to writing. Literacy and literature were introduced to enable the smooth spreading of the gospel while Shona beliefs were not considered a religion.

In order to exert their influence, these missionaries set up stations at geographically strategic positions in most Shona-speaking areas, graphically presented by Chimhundu (1992: 80) as follows:

Dialect	Denomination (Missionary Society)	Mission(s)/Headquarters (Year Established)	Nearest Town (New Name)
Ndau	American Board Mission (American Methodist)	Mount Selinda Chikore (1893)	Chipinge Melsetter (Chimanimani)
Manyika	1. Anglican Church	Penhalonga (St Augustine's) (1898)	Umtali (Mutare)
	2. Methodist Episcopal (United Methodist)	Old Umtali (1892)	Umtali (Mutare)
	3. Roman Catholic Church	Triashill (1896)	Rusape
Zezuru	1. Roman Catholic Church	Chishawasha (1892)	Salisbury (Harare)
	2. Methodist Church (Wesleyan)	Waddilove (1892) Epworth (1892)	Marandellas (Marondera) Salisbury (Harare)
Karanga	Dutch Reformed Church	Morgenster (1891)	Fort Victoria (Masvingo)
Kalanga (and Ndebele)	London Missionary Society	Inyati; Hope Fountain (1859; 1870)	Bulawayo

As a result of missionaries settling at different mission stations, a variety of approaches to the making of dictionaries for use by the indigenous population were devised. These mission stations represented the different Shona dialects spoken in Zimbabwe, namely ChiZezuru, ChiNdau, ChiManyika and Chi-Karanga. The ChiKorekore dialect was not represented in the writing panel because the missionaries shunned the swelteringly hot, tsetse fly and mosquito infested Zambezi valley where the Korekore people live.

A Brief History of Bilingual Dictionaries in Zimbabwe

According to George Fortune (1979: 42), a pioneer Bantu comparativist, W.H.I. Bleek, first made a list of Shona words entitled *The Languages of Mosambique* (1856) and surveyed the Sofala dialect along the Pungwe–Save rivers.

M.E. Weale wrote the first *Matabele and Makalaka Vocabulary Intended for the Use of Prospectors and Farmers in Mashonaland* (1903), representing Karanga. In 1894, Reverend A.M. Hartmann, a chaplain of the pioneer column, compiled *The English–Mashona Dictionary* in the Zezuru dialect. This was followed by W.A. Elliot who in 1897 published the *Dictionary of the Tebele and Shona Languages*. Elliot's work was "an attempt to present a written basis for the Shona language as a whole, from which the peculiarities of the different dialects may be observed".

H.E. Springer of Old Umtali published *A Hand Book of Chikaranga* (1905) and Father E. Beihler followed with his 1906 publication of the *English–Chi-Swina Dictionary with an Outline of ChiSwina Grammar*. The subsequent editions

that culminated in the third which received more attention, was entitled *Chi-Swina Grammar with English–ChiSwina Dictionary, and ChiSwina–English Vocabulary*. The following edition substituted the term *ChiSwina*, possibly because it sounded derogatory, with *ChiShona*, hence the title adopted was *MaShona or ChiSwina Vocabulary*. Finally the fourth edition dropped the designation *Chi-Swina* completely. After Beihler's first edition, *A Dictionary with Notes of the Grammar of the Mashona Language Commonly Called ChiSwina* that dealt with Manyika, was the next compilation by Father H. Buck at St Augustine's Mission, Penhalonga. A small English–Chikaranga Dictionary was produced in 1926.

In 1951, a first vocabulary of the Ndau dialect that came in use was entitled *ChiNdau–English and English–ChiNdau Vocabulary*, the compiler being anonymous. However, it is believed to be the Reverend G.A. Wilder's work. This same year, C.S. Louw published her *Manual of the Chikaranga Language* whose main objective was "to provide for the European student a gradual introduction to the grammar of this important branch of the group of African Native languages known as the Bantu family".

For Christianity to thrive, missionaries needed a mastery of the indigenous languages. It was therefore imperative to develop a writing system in Shona. Christianity was a significant force in African colonisation. It helped to reduce Africans to submission and eventually assisted European powers to take over African land. Various avenues of persuasion were explored to save Africans from their so-called pagan and heathen cultural practices. It is quite evident from this objective that the missionaries' main aim in compiling these dictionaries were to facilitate the translation of the Bible and other religious and educational material from English, and also to help Europeans in using Shona. Fortune (1979: 45) says "the perspective of all these works … had been to assist foreigners to speak local forms of Shona".

Shona lexicography was supposed to play a social role in the standardisation of the Shona language. By standard Shona, Chimhundu (1992: 87) says, is meant: "developing a common system of writing or orthography (that is, spelling, word-division, and also punctuation) which, when implemented, allows people who speak different varieties of the same language to write in the same way, while still allowing for stylistic and other variation, as in the choice of vocabulary".

Apart from developing orthographies, missionaries also established and controlled schools. It was their great desire to propagate Christianity among what they considered African heathens who needed a thorough teaching. The aim was also to establish a literate class who could be functional in colonial administration and commerce. It was out of these schools that a new intelligentsia of first generation writers emerged.

What then, were the consequences of these developments on African literary expression in general and the Shona lexicographic development in particular? Oral art forms were inevitably undermined, though not obliterated, by

the emergence of schooling and literacy. With the transition to a literary culture, the early African writers, being experienced in an oral culture, required a reorganisation of their entire aesthetic sensibilities.

The main institution that worked for a common orthography and lexicography among the Shona was the Southern Rhodesian Missionary Conference which was first held in 1903. The main purpose of this meeting was to secure a translation that could be used in all dialects of Mashonaland.

At the Missionary Conference held in 1906, the common version of the Bible and a common orthography were topical issues. In 1928, the Southern Rhodesian Government decided to support the use of African languages and the missionaries were encouraged, according to Fortune (1979: 46), to "put their heads together to get a common language suitable to the needs of the different Mashona tribes, so that larger and less expensive editions of higher class textbooks be printed".

For a language to develop into a standard language means serious efforts in lexicography. This explains why among Doke's 1931 recommendations was the unification of the orthographies and the vocabularies. The New Unified Orthographies for schools in 1932 and the Standard Shona for the country in 1955 led to the publication of the last series of dictionaries and vocabularies before Hannan's *Standard Shona Dictionary*, called *A Vocabulary of the Dialects of Mashonaland in the New Orthography* compiled by Reverend Bertram H. Barnes, CR.

In spite of this large amount of lexicographic work, the missionaries played a negative role in the development of African literature which should be examined in the tripartite relationship of education, colonial Christianity and African culture. Missionaries viewed African customs and morality as degraded by barbarism and superstition, so their aim was to remake Africans through the Gospel and education. This is why Chiwome (1996: 11) concedes: "Initially the church was fundamentalist, hierarchical, sectarian and authoritarian ... It did not tolerate Shona religion and customs."

Certainly, the church provided the social engineering necessary while schools provided the literacy and literary culture essential to a scripturally based faith. Thus the cultural policy of the mission schools was inimical to the African's cultivation of his traditional art. All Christian converts were expected to renounce their traditional practices. A missionary writer stationed at Zimbabwe's Gokomere Mission in 1949, Father Gruber, did not think that any African art existed in pre-colonial Rhodesia (Peet 1949).

Problems of Bilingual Dictionaries

Any speech community jealously guards the historical associations and institutions peculiar to itself through language. It is language as a cultural asset that is used by people for transmission of their knowledge systems and information. Differences in culture may result in one language using a concept on a higher

level of abstraction than the other, thereby having a gradient between the two closest equivalences. For example, Hannan explains *njuzu* as a "water sprite" while *sprite* is a "fairy". In the Shona cultural context, people should, however, know that if a *njuzu* takes a person, relatives are not supposed to mourn if that person is to be returned alive. Also, on his/her return, he/she will be possessed with great magical powers, as is indicated in Chimhundu's dictionary. The cultural range in translation should have been considered.

Any serious attempt to define *culture* should include, according to Ngara (1991: 2), the following indices: the material culture, the aspect of social relations and communication, the ideological element and the aesthetic component of a group of people. The analysis of these indices in relation to Shona lexicography, in its organic unity, will show that culture is "the expression of a people's social activities in relation to their struggle with their environment, with the elements and with human forces that threaten their survival or their way of life".

With the view that the Shona language enables people's daily experiences to be conveyed fully and vividly, Father Hannan's "contribution towards the study of the Shona language and the development of its written literature" (Fortune 1979: 21), should not be underestimated.

Since any dictionary is considered a reference text generally associated with translation of lexical items, Father Hannan is to be evaluated as a lexicographer and translator through his *Standard Shona Dictionary* (1959). Being a bilingual dictionary, his "considerations regarding its relative usefulness to first and second-language speakers" were necessary during the colonial era.

Many people regard dictionaries as infallible authorities on language usage, which is not true. In Father Hannan's dictionary, the inconsistencies and inadequacies are created by "the complex language situation in the Shona-speaking community" (Chimhundu 1979: 75). A dictionary should record and describe usage. Pinchuck (1977: 225) succinctly sums up the functions of a dictionary when he says:

> Ideally a dictionary should be an instrument of semantic discrimination. It should enable the user to choose between words for a given application; and it should offer information about the position of a word within a series and its value within a lexical structure.

Hannan's Dictionary as Example

Being a bilingual dictionary, Hannan's *Standard Shona Dictionary* was and is quite useful and particularly important for its target users who were fellow foreigners learning Shona. In his effort to translate, Hannan was faced with the major problem of whether to find equivalents that are one-to-one correspondences or single words replacing word groups or compounds and the reverse, hence the creation of collocations or single words or word groups. His explanation of *dare* describes it as a "meeting place for the men of a village". *Dare* can

refer to a court or, within a house complex, to the resting place of males where they can advise adolescent boys on how to contend with growing-up problems. The word can also mean a bell. Unfortunately, he had problems with recognising Shona homonyms.

Hannan's major weakness is that he furnishes standardised translations not corresponding with the full lexical ranges in the two languages and therefore ends up being incorrect because of the shifts of meaning in both languages. In an interview with Hannan a month before he died, Kumbirai (1979: 61) reported Hannan's comments on his difficulties in translating the New Testament into the Shona language:

> The Shona people spoke many dialects and he had, therefore, to take cognizance of this fact in his translation; that about half of the translation was entirely by himself and might therefore be discounted by native Shona speakers as "Chibaba" (which can be freely interpreted as the "Shona spoken by missionaries); and that some missionaries who knew Greek, Hebrew and Aramaic maintained that some of his translation did not do justice to the original, and, therefore, did not measure up to the world's standard of translation (and for this reason some Catholic missionaries even preferred to use the Dutch Reformed Church (Morgenster) translation).

The fact that at one time Father Hannan wished he was a native speaker who could have rendered his translations more intelligible and meaningful (Kumbirai 1979: 62), clearly spells out his linguistic deficiencies particularly for a speaker of both the source and target languages.

Since Father Hannan was a foreign speaker of the Shona language, some of his translations do not give a rendition of the vocabulary as used in normal everyday speech and therefore are not comprehensible and contemporary. While he tried to capture Shona words that occur relatively frequently in daily usage, this have had great value for the recording of the vocabulary. The most common words often have little information value, though they are important functionally. Hannan's main aim to describe the lexical items are not prescriptive in practice. The Shona language is heterogeneous and the social situation in its usage is fluid.

Hannan's dictionary presents the Shona language as an inventory, a list of words unconnected to each other, instead of offering it as a structured and patterned system of words as used in practice (Pinchuck 1977: 225). As non-native speaker of the Shona language Father Hannan encountered a multiplicity of problems. One controversial entry in his dictionary is the following: "Kipa: storekeeper Eng. Kipachitoro and Kipasitoro also occur". For Hannan to define *Kipa* only to mean a "storekeeper" is misleading. He gives examples without first giving a definition. *Kipa*, a loan word from English "keeper", means someone whose job is to look after a particular place, for example a storekeeper, goalkeeper, zookeeper, etc. Chimhundu's *Duramazwi Guru ReChiShona* rightly says that "Kipa is Muchengeti (i) Kipasitoro (ii) Gedhebhoi". Hannan's weak-

ness in this regard shows the need for recognising the crucial role of context. Meaning therefore is a result of context and it is this context that disambiguates.

Hannan's Major Weaknesses

Father Hannan's *Standard Shona Dictionary* should have been both true to the cultural facts and be informative about Shona lexical details. He was supposed to give precise descriptions of meaning. For example, he defines *zhumu* as "big tuft of hair left on shaven head (esp. of boys)". While this is correct, it is incomplete. Even animals like bucks have tufts. Father Hannan gives examples of words found in proverbs but, showing his deficiency in the Shona language, he omitted one of the most interesting proverbial sayings, *mhembwe rudzi inozvara mwana ane kazhumu* which can be translated to mean "a duiker is one of a species, if it bears a child, it will have a tuft on his head". One of his major weaknesses is his inadequate research about the exact uses and meanings of words and expressions.

Hannan's dictionary therefore neglects the reinforcing of some of the Shona people's cultural values. A proverb is a mirror of a cultural context expressing the philosophy of life in society. It is, according to Dorson (1972: 21), "a living recitation delivered to a responsive audience for such cultural purposes as reinforcement of custom and taboo, release of aggressions through fantasy, pedagogical explanations of the cultural world, and application of pressure for conventional behavior". Chimhundu (1980) supports this by adding that proverbs contain the distilled wit and wisdom of past generations because they enjoy users' "traditional popular acceptance" (Doke 1947: 102, quoted by Chimhundu 1980: 41).

Hannan's explanation of *zhumu* does not contain an example sentence of its use as is the case of other words which form part of proverbs, making its treatment inconsistent with these. Hannan's formulations, apart from being confusing, do not faithfully reflect the facts of the natural language and introduces many distortions. Most of Hannan's explanations are incorporated into the "definition". They show the typical short-comings of traditional lexicography. Explanations about what words mean are written but "very little is said about how they are used" (Hanks 1987: 121).

Hannan is found wanting in his converting information from one language or language variety into another. For example, it is not enough to explain *matakanana* as: "Muddy water. Any badly done work." It specifically refers to a children's game as appears from Chimhundu's (1996: 261) definition:

> *Aya matambiro evana vadiki diki vachitevedzera zvinoitwa mudzimba nevakuru sezvakaita kubika, vachishandisa mavhu nemvura.*
> "It is children's play imitating what elders do in carrying out household chores like cooking but they would be using sand and water".

Any dictionary should identify items in terms of what Gleason (1973: 3) calls the "structure of content", the system of interrelated meanings and values which is at the heart of a culture. Any translator is expected to be linguistically competent with a good background in the literature, culture and history of the language from which lexical items are translated.

Fortune (1979: 28) admittedly pointed out that Hannan's dictionary is "deficient in words used in the diction of traditional poetry".

> No doubt many of these words are passing out of the speech of all but the oldest age grades, and this fact points to another aspect of the difficulty in determining what is a complete record. For with changes in culture the lexicon is proving unstable.

Father Hannan furthermore failed to translate some words into Shona or English which makes his definitions difficult and unfamiliar. For a Shona word like *manyemba* to be defined by its scientific name *Smegma praeputtii* is confusing to ordinary users and only understandable to a few learned people. In translating written or recorded material of a natural language, the aim is to reproduce as accurately as possible all grammatical and lexical features of the source language by equivalents in the target language. This is why Zgusta (1971: 257) in his principles of defining says:

> The lexicographic definition (should) not contain words more difficult to understand than the explained word itself.

Through its many editions and printings Father Hannan's *Standard Shona Dictionary*, however, contributed greatly towards the general establishment of the standard Shona orthography.

Conclusion

The development of African and specifically Zimbabwean literature and lexicography shows that mission-mediated literature and lexicography faithfully reflects colonialism. With Christianity, its missions and presses were but loyal partners to the steady tightening of colonial power and the deepening subjugation of the blacks. Mission presses also did not allow the voice of black protest to be heard in their publications. However, these mission presses contributed towards the establishment of a writing and lexicographic tradition for Shona in Zimbabwe.

The lexicographic efforts of the missionaries in general, and that of Father Hannan in particular, have many weaknesses. Basically not meant for the studying of Shona for its own sake, these dictionaries were compiled by non-native speakers of Shona who worked as individuals, unable to consult a wide range of native speakers for the exact uses and meanings of words and expressions.

This early missionary lexicographic tradition later resulted in the on-going process of compiling monolingual dictionaries by the African Languages Research Institute (ALRI) (into which the ALLEX project has been transformed). Shona has now become a field of study in its own right as demonstrated by Chimhundu's dictionaries compiled electronically by a team of professional lexicographers with the help of a comprehensive database.

References

Chimhundu, H. 1979. Some Problems Relating to the Incorporation of Loanwords in the Lexicon. *Zambezia* 7(1): 75-91.

Chimhundu, H. 1980. *Shumo, Tsumo* and Socialisation. *Zambezia* 8(1): 37-51.

Chimhundu, H. 1990. Pragmatic Translation and Language Elaboration. Pongweni, A.J.C. and Thondlana (Eds.). 1990. *Lasu Conference Proceedings 1987*: 140-152. Harare: Lasu.

Chimhundu, H. 1992. Standard Shona: Myth and Reality. Crawhall, N.T. (Ed). 1992. *Democratically Speaking: International Perspectives on Language Planning*: 77-88. Cape Town: National Language Project.

Chimhundu, H. (Ed.). 1996. *Duramazwi RechiShona*. Harare: College Press.

Chimhundu, H. (Ed.). 2001. *Duramazwi Guru RechiShona*. Harare: College Press.

Chirgwin, A.M. 1932. Christian Literature in Africa. *Africa* 5: 323-337.

Chiwome, E.M. 1996. *A Social History of the Shona Novel*. Kadoma: Juta.

Dathorne, O.R. 1978. *African Literature in the Twentieth Century*. Minneapolis University: Minnesota Press.

Doke, C.M. 1931. *Report on the Unification of the Shona Dialects*. Hertford: Stephen Austin and Sons.

Doke, C.M. 1947. Bantu Wisdom-Lore. *African Studies* 6: 101-120.

Dorson, Richard M. 1972. Concepts of Folklore and Folklife Studies. Dorson, Richard M. (Ed.). 1972. *Folklore and Folklife: An Introduction*. Chicago/London: University of Chicago Press.

Fortune, G. 1979. Shona Lexicography. *Zambezia* 7(1): 21-47.

Gleason, D. 1973. *An Introduction to Descriptive Linguistics*. Essex: Longman.

Hanks, P. 1987. Definitions and Explanations. Sinclair, John M. (Ed.). 1987. *Looking Up. An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*: 115-136. London/Glasgow: Collins ELT.

Hannan, M. 1959. *Standard Shona Dictionary*. Harare: The College Press/The Literature Bureau.

Kumbirai, J.C. 1979. Shona Bible Translation: The Work of the Revd Michael Hannan, S.J. *Zambezia* 7(1): 61-74.

Ngara, E. 1991. *Culture for Liberation and Social Transformation*. Unpublished paper read at the Ministry of Education and Culture Seminar on Cultural Awareness, held at the Kadoma Ranch Motel, 15–17 May 1991.

Peet, S. 1949. Report of the Visit to the Range of Fort Victoria, Gutu Reserve 25–31 March 1949. *Central Africa Film Unit File F 268 to 269*. Harare: National Archives of Zimbabwe.

Pinchuck, I. 1977. *Scientific and Technical Translation.* London: André Deutsch.

Zgusta, L. 1971. *Manual of Lexicography*. Janua Linguarum. Series Maior 39. The Hague: Mouton.

Two Decades of *Lexicographica Series Maior*

Reinhard Hartmann, *Exeter, United Kingdom (r.r.k.hartmann@ex.ac.uk)*

No stagnation of lexicography is to be expected,
but on the contrary further flourishing.
Ladislav Zgusta (1991: 3165)

Abstract: This article is a personal account, from the point of view of one of its editors, of the publishing history of *Lexicographica Series Maior*.

Keywords: LEXICOGRAPHICA SERIES MAIOR, PUBLISHING HISTORY, METALEXICO-GRAPHY, LEXICOGRAPHY

Opsomming: Twee dekades van *Lexicographica Series Maior*. Hierdie artikel is 'n persoonlike weergawe, uit die gesigspunt van een van sy redakteurs, van die publikasiegeskiedenis van *Lexicographica Series Maior*.

Sleutelwoorde: LEXICOGRAPHICA SERIES MAIOR, PUBLIKASIEGESKIEDENIS, META-LEKSIKOGRAFIE, LEKSIKOGRAFIE

1. The first volume(s)

It was exactly 20 years ago, while I was editing the proceedings of the *LEXeter '83* Conference for publication as Volume 1 in the *Lexicographica Series Maior* (Hartmann 1984), that I wondered what all this would lead to. At a preliminary meeting in Exeter (not the one that led to the creation of the three-volume ency-clopedia of lexicography, *Wörterbücher/Dictionaries/Dictionnaires*, edited by Franz Josef Hausmann et al. 1989–91), two or three of the conference partici-pants who later became editors decided that the time was ripe for an interna-tional book series specialising in lexicography, and the German publisher Nie-meyer at the same time agreed to take this on, together with the associated periodical *Lexicographica. International Annual for Lexicography*. However, none of us could have foreseen then that by the end of 2003, the book series would have grown to 114 titles since 1984 and the journal to 19 yearly issues since 1985.

In 1988, when LexSM had produced more than 20 volumes (six on average for each of the first four years), I referred to it as a 'basic training library' on the grounds that it seemed to contain "all the major aspects in the theory and prac-tice of dictionary making" (Hartmann 1988: 151). This characterisation is even

more appropriate today, as the published volumes represent what I have called a "wide thematic orientation and coverage" (Hartmann 2000: 4) in terms of the whole range of current perspectives, from dictionary history and dictionary typology to dictionary criticism, and from dictionary use and dictionary structure to computational lexicography. This article[1] asks whether the ambitious aims of spanning the whole spectrum of lexicography and contributing to its 'further flourishing' have been achieved.

2.   Plans and policies

From the start, we have had an editorial team of seven 'leaders' in the field, currently consisting of Ladislav Zgusta (USA), Sture Allén (Sweden), Franz Josef Hausmann, Oskar Reichmann and Ulrich Heid (Germany), Pierre Corbin (France) and myself (U.K.), who rotate the executive editorship for one year at a time (the current calendar year being my own third turn). Although we knew that in terms of sales figures it would be an uphill struggle, we were keen from the beginning (a) to develop an international focus on dictionary research by encouraging a scholarly approach to outstanding issues, (b) to cultivate interdisciplinary contacts with many fields, from linguistics and cultural history to education and information technology, and — perhaps less successfully until very recently — (c) to promote interlingual topics such as pedagogical lexicography and translation.

All of this needed agreement among the editors and with the publisher on proper procedures for attracting, assessing, approving, producing and publicising the books we were to include in the series. From my files I see that by late 1983 we had worked out a contract and a set of guidelines for preparing copy, including the requirement of two positive editorial reports for accepting any manuscript for LexSM. Summaries in English (and sometimes French) are also asked for if the text is written in another language. Since then these rules have been elaborated, refined and partly computerised, and are made available to potential authors.

We have been concerned all along that publicity about LexSM should reach as wide a readership as possible, and the publisher has helped by producing a dedicated brochure (for volume numbers 1-87 in 1998), an annual catalogue as well as half-yearly leaflets, the latter usually with two to three LexSM items in each. Thus, the leaflet for I/2004 is advertising three volumes: Zita Hollós, on the topic of how to design a collocation-oriented German–Hungarian learner's dictionary, Jenny Thumb, on a think-aloud study of dictionary look-up strategies among Hong Kong learners of English, and the proceedings of a German conference on lexicography and its interdisciplinary contacts edited by Thomas Herbst et al. Copies of LexSM publications are regularly sent to periodicals for review and to book exhibitions at selected conferences. I myself have written reports on many of the volumes for the *Annotated Biblio-*

graphy for English Studies[2], and recently initiated an announcement about the series in relevant journals, newsletters and websites[3].

Manuscripts have come from all over (including some from the editors themselves, e.g. Ladislav Zgusta's 1988 bibliography of metalexicography, Hans-Peder Kromann's jointly edited 1995 proceedings of a workshop on contrastive lexicology, Oskar Reichmann's co-authored 1995 book on hidden lexicographic information in Early New High German sources, and Ulrich Heid's 1997 study of how to structure monolingual and interlingual electronic dictionaries), but the European context has been predominant, "although no single profile has yet emerged of the typical country, the typical dictionary, and the typical lexicographer in Europe" (Hartmann 1993: 77). Links have always existed with both EURALEX (Volume 1 had presented the 55 papers of its first Congress, while the proceedings of subsequent congresses have been published by their respective host institutions) as well as the Dictionary Society of North America. But gradually, more and more countries outside Europe, and languages other than German and English, have come within our scope. The authors and editors of submitted manuscripts and published volumes have included both recognised authorities (like Herbert Ernst Wiegand, who had also been instrumental originally in getting both the book series and the year-book series *Lexicographica* off the ground) and relative newcomers, e.g. authors of recent research dissertations, such as Ulrike Rothe's (2001) comparative study of the English and French 'dictionary scenes', and of the way(s) in which they may condition different lexicographic practices.

3.   Texts and topics

LexSM volumes have exemplified several different text genres:

(a)   *Conference papers* constitute a substantial proportion of published titles. At least 10 of these are devoted to the proceedings of the Symposium on Lexicography, a pioneering series of conferences held at Copenhagen every other year since 1982, which had helped to build bridges between theory and practice, Northern Europe and the rest of the world, and East and West (especially during the earlier phases when it offered an opportunity for scholars from both East and West Germany to meet). 11 symposia have been held so far, and the proceedings of all but the very first have been published in LexSM. Number VIII (1998) may be considered typical of this sub-series: edited by the founder, Arne Zettersten, and two of his colleagues, it consists of 22 papers by 24 authors typifying the 10 or so languages and countries that were represented at the 1996 Symposium, addressing such diverse issues as text compression, grammatical and semantic information, and translation equivalence. Of the many other conference proceedings published in LexSM, I will mention only two which are based on conferences held in 2002 and due to be

issued later this year: one edited by CHAN Sin-Wei (Hong Kong) on the mutual relations between translation and bilingual lexicography, and the other edited by Julie Coleman (Leicester) and Anne McDermott (Birmingham) on the development of historical and period dictionaries.

(b)   *Anthologies* and other collections bring together papers on related sets of topics by a single author (e.g. Herbert Ernst Wiegand's on semantics and lexicography, edited by Antje Immken and Werner Wolski 1999) , by two authors (e.g. Jean-Claude Boulanger and Monique Cormier 2001 on personal names in French lexicography), or by a group of authors (e.g. Gerhard Augst et al. 1997 on German spelling dictionaries and Sue Atkins et al. 1998 on dictionary use), or papers delivered by a number of authors in one place over a period of time, e.g. (four) volumes of lectures from the Lexicography Colloquium at Heidelberg (Wiegand 2000).

(c)   *Bibliographies* include Zgusta's (1988), but also a rather unique bibliography of dictionary bibliographies by Margaret Cop (1990), a bibliography of a total of 4863 monolingual and interlingual, general and specialised dictionaries of Dutch by Frans Claes and Peter Bakema (1995), and an annotated list of references on learners' dictionaries and the user perspective by Fredric Dolezal and Don McCreary (1999).

(d)   *Monographs* comprise about half of the total LexSM list. They are typically based on specific research projects, some of which may be worth commenting on here.

  Pedagogical lexicography is one subject which has grown in stature over the years, covering topics such as user surveys, from Turki Diab's (1990) report on dictionary use by Arab nursing staff to Ursula Wingate's (2002) observation of Chinese students of English and German, and learner-oriented dictionaries, from Humberto Hernández' (1989) analysis of Spanish school dictionaries via Ekkehard Zöfgen's (1994) discussion of French dictionaries for German learners to Martin Stark's (1999) study of the design features of the hybrid genre 'encyclopedic learner's dictionary' from the point of view of the user and Hilary Nesi's (2000) account of how foreign learners use and misuse EFL dictionaries.

  Grammar in the dictionary is another recurring theme, which embraces such specialties as morphology and syntax (e.g. Edward McCorduck 1993), valency (e.g. Vilmos Ágel 1988), phraseology (e.g. Mi-Ae Cheon 1998) and collocations (e.g. Peter Howarth 1996).

  Other *information categories* that have been treated in detail are pictorial illustrations (Werner Hupka 1989), usage labels (Michel Glatigny 1998), and translation equivalents (Gitte Baunebjerg Hansen 1990).

Terminology, specialised dictionaries and languages for specific purposes have been examined, inter alia, from the point of medicine (strictly speaking a collection coordinated by Stephan Dressler and Burkhard Schaeder 1994) and accountancy (Franz Schneider 1998).

Computational lexicography is the term often used to refer to the ever-closer relationship between lexicography and information technology, represented by the LexSM volumes by Matthias Kammerer (1995) on how to write dictionary entries on the computer screen, Thierry Fontenelle (1997) on how to extract semantic information from machine-readable dictionaries, and papers in the collection edited by Ingrid Lemberg et al. (2001) on how to support lexicography by electronic means.

Various *lexicographic traditions* have been investigated in detail, e.g. for English (Gabriele Stein 1985), for French (Margarete Lindemann 1994) and several other languages (and language pairs), including Greek (Evthymios Papachristos 1990), Tamil (Gregory James 1991), Hungarian (Rita Fejér 1995), Turkish (Hadi Yüksekkaya 1998), Polish (Agnieszka Frączek 1999), and Latvian (Ineta Balode 2002).

Some famous (and not so famous) *lexicographers* have received the attention of LexSM authors, e.g. John Wilkins and William Lloyd (Fredric Dolezal 1985), Matthias Kramer (Laurent Bray 2000), Antoine Furetière (Dorothea Behnke 1996), and Jacob and Wilhelm Grimm (two volumes prepared by a collective led by Alan Kirkness et al. 1991).

(e)   *Textbooks* include Burkhard Schaeder's (1987) introduction to lexicography within the confines of German studies.

(f)   *Dictionaries* themselves have not been neglected, and include both entirely new ones and reprints of classics, e.g. the Chinese–English dictionary of idioms and proverbs by HENG Xiao-Jun and ZHANG Xue-Zhi (1988), the dictionary of Yiddish loan-words in German dialects by Heidi Stern (2000), and new editions of Daniel Sanders' German thesaurus (Peter Kühn 1985) and of Jean-François Féraud's *Dictionaire Critique* (Philippe Caron and Terence Wooldridge 1994).

4.   Strains and satisfactions

59 LexSM titles have been cited in this article, an arbitrarily chosen cross-section of 50% of those published so far, including two that are forthcoming later in 2004. But progress has not always been even, and relations between our respective offices and the outside world have not always been straightforward. For a start, the editors can bring their attention only to what they are offered,

and they very rarely get the chance of initiating a publication unless they take it on themselves (see Section 2 above) or encourage others, such as colleagues, visiting fellows or doctoral students they have supervised or examined to write up their research (I am proud that more than 10% of the cited titles are the result of my own intervention through such channels). It may well be the case that we have not (yet) dealt with certain topics that may be of particular interest to individual editors. And if a manuscript on such a subject finally emerges, it sometimes has to be turned down because of formal deficiencies. It even happens that an author or editor of a manuscript accepted for LexSM decides to take it to a different publisher. Thus I was personally very disappointed when we lost a volume on the lexicography of American Indian languages (edited by William Frawley et al. 2002).

There have been other losses, too, for instance when one of the editors, Hans-Peder Kromann, died prematurely and after completing only one round of executive editorship. And there are issues that have never been fully resolved, for example, how the loose links we have enjoyed with EURALEX and DSNA from the start should have been formalised, and whether similar links should be established with other continental, regional or subject societies.

Among the positive highlights has been the feeling that we can observe, sometimes even promote, progress in relatively neglected areas. One of these has been dictionary criticism, where we have definitely noticed considerable advances over the years, not least by the contributions of the periodical *Lexicon* published by the Iwasaki Linguistic Circle in Tokyo. Within LexSM, Martha Ripfel (1989) analysed the texts of 736 reviews of five dictionaries, and Herbert Ernst Wiegand commissioned teams of experts to evaluate and assess prototypical dictionaries. Based on the earlier 'grilling' of two German dictionaries for foreign learners (*Langenscheidt* and *De Gruyter*), Wiegand's (2003) meticulous examination, in the first of two volumes dedicated to the *Duden. Das große Wörterbuch der deutschen Sprache in zehn Bänden*, consisted of 29 contributions by 31 authors arranged in six chapters, ranging from comparisons with other 'national' dictionaries and other Duden dictionaries to this dictionary's treatment of semantics, grammar and regional vocabulary. Seven more chapters on multifarious structures and information categories in both the print and electronic versions of the well-known 10-volume general-purpose dictionary are planned for the second volume, which is in preparation.

Sometimes a book takes a long time to materialise, and there is then great personal satisfaction when something that you had almost given up hope for sees the light of day. One example was the long gestation of the book on the methodology of user studies by Yukio Tono (2001), which I had encouraged ever since we had made contact many years ago, but which when it finally appeared set new standards.

Lastly, an illustration of both high and low points. Ladislav Zgusta has been well represented in LexSM (and I have publicly acknowledged my indebtedness to him by reprinting part of his (1971) *Manual of Lexicography* in my

3-volume reader *Lexicography: Critical Concepts*, Hartmann 2003). Some years ago, when I heard that a festschrift was planned to honor this all-time 'great', I was ready to offer my own contribution to the two editors (Braj Kachru and Henry Kahane 1995). However, as this was couched in terms of a transcontinental trialogue (between 'Lex', 'Dox' and 'Fax') rather than presented in the form of a conventional paper, it was turned down [and eventually published elsewhere: Hartmann (1994)].

5.   Some final thoughts

So what has been accomplished? Has LexSM made an impact? Has the series increased general awareness about lexicography and its problems and pleasures? Based on the evidence of impressive sales achieved (6 of the first 20 volumes now out of print) and positive reviews received, I certainly think so. One of the things I have enjoyed most is the privilege of judging some of the best contributions to lexicography and thus supporting its maturation into a scholarly discipline, although it is still considered an inferior minority subject in academic circles, and contacts between LexSM editors and authors are therefore often by word of mouth. What I am most happy about, though, is the fact that the series has become the only truly representative organ for all aspects of our specialism world-wide.

Notes

1.   I am grateful to the editor of *Lexikos*, J.C.M.D. du Plessis, for encouraging me to report on my experiences as one of the editors of LexSM. Although I have shown a draft to my fellow editors for comment, I alone remain responsible for its contents.

2.   Between 1994 and 2004, I compiled more than 20 entries on LexSM titles for this on-line service. Website: <http://abes.swets.nl/abes/>.

3.   *Contact details*: For more information on titles issued since Volume 95, consult the website of our publisher, Max Niemeyer (Tübingen): <www.niemeyer.de>. As one of the seven editors of the series (executive for the current calendar year 2004), I would be keen to hear from you if you have a manuscript to submit or know of a proposal for one. Please write to Reinhard Hartmann by email (above) or: 40 Velwell Road, Exeter, Devon EX4 4LD, U.K.

References

Ágel, V. 1988. *Überlegungen zur Theorie und Methode der historisch-synchronen Valenzsyntax und Valenzlexikographie*. LexSM 25. Tübingen: M. Niemeyer.

Atkins, B.T.S. (Ed.). 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. LexSM 88. Tübingen: M. Niemeyer.

Augst, G. et al. 1997. *Rechtschreibwörterbücher im Test*. LexSM 78. Tübingen: M. Niemeyer.

Balode, I. 2002. *Deutsch–lettische Lexikographie. Eine Untersuchung zu ihrer Tradition und Regionalität im 18. Jahrhundert*. LexSM 111. Tübingen: M. Niemeyer.

Baunebjerg Hansen, G. 1990. *Artikelstruktur im zweisprachigen Wörterbuch*. LexSM 35. Tübingen: M. Niemeyer.

Behnke, D. 1996. *Furetière und Trévoux. Eine Untersuchung zum Verhältnis der beiden Wörterbuchserien*. LexSM 72. Tübingen: M. Niemeyer.

Boulanger, J.-C. and M.C. Cormier. 2001. *Le nom propre dans l'espace dictionnairique général*. LexSM 105. Tübingen: M. Niemeyer.

Bray, L. 2000. *Matthias Kramer et la lexicographie du français en Allemagne au XVIIIe siècle*. LexSM 99. Tübingen: M. Niemeyer.

Caron, P. and T.R. Wooldridge (Eds.). 1994. *Jean-François Féraud: Dictionaire Critique de la Langue Française (1787)*. LexSM 51-53. Tübingen: M. Niemeyer.

Chan, S.-W. (Ed.). 2004. *Translation and Bilingual Dictionaries*. LexSM 119. Tübingen: M. Niemeyer.

Cheon, M.-A. 1998. *Zur Konzeption eines phraseologischen Wörterbuchs für den Fremdsprachler. Am Beispiel Deutsch–Koreanisch*. LexSM 89. Tübingen: M. Niemeyer.

Claes, F. and P. Bakema. 1995. *A Bibliography of Dutch Dictionaries*. LexSM 67. Tübingen: M. Niemeyer.

Coleman, J. and A. McDermott (Eds.). 2004. *Dictionary History and Historical Lexicography*. LexSM 123. Tübingen: M. Niemeyer.

Cop, M. 1990. *Babel Unravelled. An Annotated World Bibliography of Dictionary Bibliographies, 1658–1988*. LexSM 36. Tübingen: M. Niemeyer.

Diab, T. 1990. *Pedagogical Lexicography. A Case Study of Arab Nurses as Dictionary Users*. LexSM 31. Tübingen: M. Niemeyer.

Dolezal, F.T. 1985. *Forgotten but Important Lexicographers: John Wilkins and William Lloyd*. LexSM 4. Tübingen: M. Niemeyer.

Dolezal, F.T. and D.R. McCreary. 1999. *Pedagogical Lexicography Today. A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users*. LexSM 96. Tübingen: M. Niemeyer.

Dressler, S. and B. Schaeder (Eds.). 1994. *Wörterbücher der Medizin. Beiträge zur Fachlexikographie*. LexSM 55. Tübingen: M. Niemeyer.

Fejér, R. 1995. *Zur Geschichte der deutsch–ungarischen und ungarisch–deutschen Lexikographie*. LexSM 60. Tübingen: M. Niemeyer.

Fontenelle, T. 1997. *Turning a Bilingual Dictionary into a Lexical-Semantic Database*. LexSM 79. Tübingen: M. Niemeyer.

Frączek, A. 1999. *Zur Geschichte der deutsch–polnischen und polnisch–deutschen Lexikographie (1772–1868)*. LexSM 93. Tübingen: M. Niemeyer.

Frawley, W. et al. (Eds.). 2002. *Making Dictionaries. Preserving Indigenous Languages of America*. Berkeley: University of California Press.

Glatigny, M. 1998. *Les marques d'usage dans les dictionnaires français monolingues du XIXe siècle*. LexSM 91. Tübingen: M. Niemeyer.

Goebel, U. et al. 1995. *Versteckte lexikographische Information*. LexSM 65. Tübingen: M. Niemeyer.

Hartmann, R.R.K. 1988. IJL salutes *Lexicographica. Lexicographica Series Maior* — The first 20 volumes. *International Journal of Lexicography* 1(2): 151-152.

Hartmann, R.R.K. 1993. General Lexicography in Europe. *Lexikos* 3: 67-82.

Hartmann, R.R.K. 1994. The bilingualised learner's dictionary (A transcontinental trialogue on a relatively new genre). James, G. (Ed.). 1994. *Meeting Points in Language Studies*: 171-183. Hong Kong: The Hong Kong University of Science and Technology Language Centre.

Hartmann, R.R.K. 2000. European lexicography. *Dictionaries. Journal of the Dictionary Society of North America* 21: 1-21.

Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9–12 September 1983*. LexSM 1. Tübingen: M. Niemeyer.

Hartmann, R.R.K. (Ed.). 2003. *Lexicography: Critical Concepts* [3 volumes]. London: Routledge.

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.* Handbücher zur Sprach- und Kommunikationswissenschaft Volumes 5.1–5.3. Berlin: W. de Gruyter.

Heid, U. 1997. *Zur Strukturierung von einsprachigen und kontrastiven elektronischen Wörterbüchern*. LexSM 77. Tübingen: M. Niemeyer.

Heng, X.-J. and X.-Z. Zhang. 1988. *A Chinese–English Dictionary of Idioms and Proverbs*. LexSM 24. Tübingen: M. Niemeyer.

Herbst, T. et al. (Eds.). 2004. *Lexikografie, ihre Basis- und Nachbarwissenschaften*. LexSM 118. Tübingen: M. Niemeyer.

Hernández, H. 1989. *Los diccionarios de orientación escolar*. LexSM 28. Tübingen: M. Niemeyer.

Hollós, Z. 2004. *Lernerlexikographie: syntagmatisch*. LexSM 116. Tübingen: M. Niemeyer.

Howarth, P.A. 1996. *Phraseology in English Academic Writing*. LexSM 75. Tübingen: M. Niemeyer.

Hupka, W. 1989. *Wort und Bild. Die Illustrationen in Wörterbüchern und Enzyklopädien*. LexSM 22. Tübingen: M. Niemeyer. [Section 8.7 'Fallstudie/Case Study' updated, translated and reprinted in Hartmann 2003, Vol. III: 363-390.]

Immken, A. and W. Wolski. (Eds.). 1999. *Herbert Ernst Wiegand: Semantics and Lexicography. Selected Studies (1976–1996)*. LexSM 97. Tübingen: M. Niemeyer.

James, G. 1991. *Tamil Lexicography*. LexSM 40. Tübingen: M. Niemeyer.

Kachru, B.B. and H. Kahane (Eds.). 1995. *Cultures, Ideologies, and the Dictionary. Studies in Honor of Ladislav Zgusta*. LexSM 64. Tübingen: M. Niemeyer.

Kammerer, M. 1995. *Bildschirmorientiertes Abfassen von Wörterbuchartikeln*. LexSM 68. Tübingen: M. Niemeyer.

Kirkness, A. et al. (Eds.). 1991. *Studien zum Deutschen Wörterbuch von Jacob Grimm und Wilhelm Grimm*. LexSM 33/34. Tübingen: M. Niemeyer.

Kromann, H.-P. and A.L. Kjær (Eds.). 1995. *Von der Allgegenwart der Lexikologie*. LexSM 66. Tübingen: M. Niemeyer.

Kühn, P. (Ed.). 1985. *Daniel Sanders: Deutscher Sprachschatz* [Nachdruck der Ausgabe Hamburg 1873–1877]. LexSM 6/7. Tübingen: M. Niemeyer.

Lemberg, I. et al. (Eds.). 2001. *Chancen und Perspektiven computergestützter Lexikographie*. LexSM 107. Tübingen: M. Niemeyer.

Lindemann, M. 1994. *Die französischen Wörterbücher von den Anfängen bis 1600*. LexSM 54. Tübingen: M. Niemeyer.

McCorduck, E.S. 1993. *Grammatical Information in ESL Dictionaries*. LexSM 48. Tübingen: M. Niemeyer.

Nesi, H. 2000. *The Use and Abuse of EFL Dictionaries*. LexSM 98. Tübingen: M. Niemeyer.

Papachristos, E.C. 1990. *Die deutsch–neugriechische Lexikographie von 1796–1909*. LexSM 32. Tübingen: M. Niemeyer.

Reichmann, O. [see Goebel et al.]

Ripfel, M. 1989. *Wörterbuchkritik. Eine empirische Analyse von Wörterbuchrezensionen*. LexSM 29. Tübingen: M. Niemeyer.

Rothe, U. 2001. *Das einsprachige Wörterbuch in seinem soziokulturellen Kontext*. LexSM 108. Tübingen: M. Niemeyer.

Schaeder, B. 1987. *Germanistische Lexikographie*. LexSM 21. Tübingen: M. Niemeyer.

Schneider, F. 1998. *Studien zur kontextuellen Fachlexikographie. Das deutsch–französische Wörterbuch der Rechnungslegung*. LexSM 83. Tübingen: M. Niemeyer.

Stark, M.P. 1999. *Encyclopedic Learners' Dictionaries*. LexSM 92. Tübingen: M. Niemeyer.

Stein, G. 1985. *The English Dictionary before Cawdrey*. LexSM 9. Tübingen: M. Niemeyer.

Stern, H. 2000. *Wörterbuch zum jiddischen Lehnwortschatz in den deutschen Dialekten*. LexSM 102. Tübingen: M. Niemeyer.

Thumb, J. 2004. *Dictionary Look-up Strategies and the Bilingualised Learner's Dictionary*. LexSM 117. Tübingen: M. Niemeyer.

Tono, Y. 2001. *Research on Dictionary Use in the Context of Foreign Language Learning*. LexSM 106. Tübingen: M. Niemeyer.

Wiegand, H.E. 2003. *Untersuchungen zur kommerziellen Lexikographie der deutschen Gegenwartssprache I (Duden)*. LexSM 113. Tübingen: M. Niemeyer.

Wiegand, H.E. (Ed.). 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*. LexSM 100. Tübingen: M. Niemeyer.

Wingate, U. 2002. *The Effectiveness of Different Learner Dictionaries*. LexSM 112. Tübingen: M. Niemeyer.

Yüksekkaya, H.Y. 1998. *Die deutsch–türkische Lexikographie*. LexSM 85. Tübingen: M. Niemeyer.

Zettersten, A. et al. (Eds.). 1998. *Symposium on Lexicography VIII. Proceedings of the Eighth International Symposium on Lexicography May 2–4, 1996, at the University of Copenhagen*. LexSM 90. Tübingen: M. Niemeyer.

Zgusta, L. 1971. *Manual of Lexicography*. Janua Linguarum. Series Maior 39. Prague: Academia/The Hague: Mouton. [pp. 345-357 updated and reprinted in Hartmann 2003, Vol. I: 70-82.]

Zgusta, L. [with the assistance of D.M.T.Cr. Farina]. 1988. *Lexicography Today. An Annotated Bibliography of the Theory of Lexicography*. LexSM 18. Tübingen: M. Niemeyer.

Zgusta, L. 1991. Probable Future Developments in Lexicography. Hausmann, F.J. et al. (Eds.). 1989–1991, Vol. III: 3158-3168.

Zöfgen, E. 1994. *Lernerwörterbücher in Theorie und Praxis*. LexSM 59. Tübingen: M. Niemeyer.

Kirkeby's *English–Swahili Dictionary*

James S. Mdee, *Institute for Kiswahili Research, University of Dar es Salaam, Dar es Salaam, Tanzania (jamesmdee@hotmail.com)*

Abstract: Kirkeby's *English–Swahili Dictionary* is a bilingual dictionary of more than 50 000 entries. The most laudable feature of the dictionary is its attempt to be user-friendly especially in the way the entry words have been arranged and the amount of information given. However, a clear objective for the compilation of the ditionary is lacking. The compilers do not seem to know the lexicographical gap they want to fill, the users they are targeting, and their dictionary-using skills. In discussing the strong and weak points of the dictionary, the article will refer to theories of dictionary criticism. Three criteria set by McMillan (1949) will guide this review article: (1) the quantity of the information in the dictionary; (2) the quality of the information presented; and (3) the effectiveness of the presentation of the information. Questions posed in the course of this article will include: Does the dictionary give the information required by the user? Is the information transparently accessible? How is the information presented?

Keywords: DICTIONARY EVALUATION, USER-FRIENDLY, DICTIONARY-USING SKILLS, LEXICOGRAPHICAL ENTRIES, GRAMMATICAL CATEGORIES, SUBGRAMMATICAL CATEGORIES, WORD COMBINATIONS, COLLOCATIONS, TRANSLATION EQUIVALENTS

Opsomming: Kirkeby se *English–Swahili Dictionary*. Kirkeby se *English–Swahili Dictionary* is 'n tweetalige woordeboek met meer as 50 000 inskrywings. Die mees prysenswaardige kenmerk van die woordeboek is sy poging om gebruikersvriendelik te wees, veral deur die manier waarop die trefwoorde gerangskik is en die hoeveelheid inligting wat verskaf word. 'n Duidelike doelwit vir die samestelling van die woordeboek ontbreek egter. Die samestellers is skynbaar onseker oor die leksikografiese leemte wat hulle wil vul, en die gebruikers vir wie dit bedoel is en hul woordeboekgebruikersvaardighede. In die bespreking van die sterk en die swak eienskappe van die woordeboek sal die artikel verwys na teorieë van woordeboekkritiek. Die kriteria gestel deur McMillan (1949) sal as gids dien vir hierdie resensieartikel: (1) die hoeveelheid inligting in die woordeboek; (2) die gehalte van die inligting aangebied; en (3) die doeltreffendheid van die aanbieding van die inligting. Vrae wat in die loop van hierdie artikel gestel sal word, sluit in: Gee die woordeboek die inligting wat deur die gebruiker verlang word? Is die inligting deursigtig beskikbaar? Hoe word die inligting aangebied?

Sleutelwoorde: WOORDEBOEKBEOORDELING, GEBRUIKERSVRIENDELIK, WOORDE-BOEKGEBRUIKSVAARDIGHEDE, LEKSIKOGRAFIESE INSKRYWINGS, GRAMMATIKALE KATEGORIEË, SUBGRAMMATIKALE KATEGORIEË, WOORDVERBINDINGS, KOLLOKASIES, VERTAALEKWIVALENTE

1. Introduction

Throughout its long history, the dictionary has always been an object of criticism and it is this practice that helped it to grow from its beginning as a glos-

sary to a book of multi-linguistic and multi-cultural information like James Murray's *Oxford English Dictionary*. Some of the early recorded dictionary critics as quoted in Landau (1984: 48-56) include Addison who suggested an English dictionary with quotations from literature. Many critics of 18th century English dictionaries such as Swift, Defoe, Pope, and others, proposed a dictionary attempting to survey and record the English language. Johnson (1755) took into account proposals made by these critics and introduced into his dictionary illustrative quotations as evidence of how a word is used in its different contexts and meanings. Johnson had his critics too. Richardson and Trench, as quoted in Mathews (1933: 63-65) criticized Johnson and others for not adhering to the principles of historical lexicography. It is on the basis of these comments that James Murray with others compiled the *Oxford English Dictionary* according to the principles of historical lexicography. Modern dictionary critics include Lemmens and Wekker (1986) who reviewed grammar in English learner's dictionaries in which they pointed out flaws and inconsistencies in entering grammar in these dictionaries. Recommendations made by the critics helped to improve the treatment of grammar in the *Oxford Advanced Learner's Dictionary of Current English*, the *Longman Dictionary of Contemporary English*, the two volumes of the *Oxford Dictionary of Current Idiomatic English* and the *Longman Dictionary of Phrasal Verbs*. Critical reviews of the coding system of the dictionaries and proposals for an adequate coding system helped much to improve the coding system used in presenting various kinds of grammatical information.

The value of a dictionary is not based on the size of the dictionary alone (Ilson 1988: 10) but also on other more important factors. Jackson (1998: 28) puts these in question form: Is the information required by the user present in the dictionary and is it transparently accessible? How is the information presented?

The following three criteria for evaluating dictionaries set forth by McMillan (1949) will be used in discussing Kirkeby's *English–Swahili Dictionary*:

(1)   The quantity of information the dictionary offers. This includes the number of entries, meanings, new words (when compared with its competitors), synonyms, pronunciation, etymology and the use of subject and usage labels.

(2)   The quality of the information presented in the dictionary. The aspects to examine here are:

(a)   the accuracy of the information,
(b)   the completeness of the information,
(c)   the clearness of the information,
(d)   the simplicity in presenting the information, i.e. the information should be rendered in such a way that it can easily be deciphered, and

(e) the modernity of the dictionary, i.e. the number of current words in the language which has been included in the dictionary.

(3) The effectiveness of the presentation of the information. This includes the order of arranging entries, the placement of etymology, the ordering of senses, and the presentation of pronunciation. Also important is the typography.

McMillan notes that any thorough review should also examine the essays in the front matter and the appendices in the back matter. It is important that a dictionary review should further answer the following questions: What are the purposes of the particular dictionary? What are its properties ? How is its content presented and how are the information categories structured?

2.   Kirkeby's dictionary

If a dictionary were to be judged by the number of its printed pages alone, the number of entries it purportedly contains according to the blurb, the information categories it has and the size of the paper used in printing it, Kirkeby's dictionary would undoubtedly have been the largest English–Swahili dictionary. It has 1 069 pages in comparison with the 924 pages of TUKI (1996). The largest Swahili dictionary is the Swahili–French dictionary of Sacleux (1939) with 1 112 pages. Kirkeby (2000) is the first English–Swahili dictionary that offers a pronunciation guide for English words. Even the paper size is also the largest. Unlike TUKI or Johnson or many English dictionaries whose page numbering in Arabic numerals only begins with the main text, Kirkeby starts off with Arabic numbering from the first page so that the main section of the dictionary only begins on page 13.

Kirkeby (2000) is a replica of an English–Norwegian dictionary by the same compiler. The project started in 1997 when Willy Kirkeby and his assistants/consultants began to replace the Norwegian text with a Swahili text (see Preface). The dictionary was completed in two and a half years. In addition to several assistants, the compiler had the help of a Swahili novelist, a teacher and a Swahili expert who acted as proof reader. Kirkeby is a prolific dictionary compiler. He has to his credit ten other dictionaries which were compiled between 1979 and 1998 and which were published in 1979, 1983, 1986, 1988, 1991, 1992, 1993, 1996 and 1998. Two dictionaries appeared in 1988. Of these dictionaries, six were Norwegian–English while four were English–Norwegian, one of which was used in compiling this English–Swahili dictionary.

Kirkeby does not explicitly identify the user profile of his dictionary except mentioning nurses complaining that they had no dictionary to look up medical terms they came across in textbooks (see Preface). It is important for any dictionary project to identify the users, the information the users need and if they can understand such information. The lexicographer has therefore to

answer the following questions: Who will use the dictionary? What will they use it for? What are their reference skills and what are their expectations? It is important to recall the observation of Barnhart (1962: 161): "It is the function of a popular dictionary to answer the questions that the user of the dictionary asks, and dictionaries on the commercial market will be successful in proportion to the extent to which they answer these questions of the buyer."

Dictionary writing should not be driven by commercial factors alone, e.g. the potential size of the market and current competitor dictionaries in the market, but should be designed with a special set of users and their specific needs in mind. Apart from being guided by the three aspects of evaluation outlined in par. 1, the following discussion will especially consider the needs of the two main user groups for whom this dictionary is possibly meant: English learners of Swahili and Swahili learners of English.

3.   Entries in Kirkeby

Kirkeby estimates that his dictionary contains 60 000 entries within the 1 047 pages of the main dictionary text. But a rough estimate shows that the dictionary may contain a little more than 50 000 entries. The entries are arranged in alphabetical order. An entry in this dictionary could be a basic form, a derived or inflectional form of the basic form, a compound word or verbal phrase and fixed expressions whose first member is the headword. For example:

(1)   **accident, accidental, accidentally, accidental death, accident write-off.**

All these entries are in bold-face. In addition to this, some phrasal terms in bold-face have also been entered as run-ons. For example:

(2)   **tablet** … *n* **1.** … kidoge (vi-); **sleeping** ~ …; **suckable** ~ …
        I. track … *n*. **1.** alama (-) (*of* ya); **cover one's -s** …; *colloq:* **stop dead in one's -s** …; **off the beaten ~** …

It seems as if the dictionary has entered as many word combinations (compound words) of the headword as possible. Although this could be seen as a good practice, the question remains: Which of these are essential in a dictionary of this size? Of the word **family** 24 combinations are given (see p. 339), all of which are entered as headwords.

(3)   **family allowance, family bliss, family business, Family Division, family doctor, family estate, family feeling, family friend, family grouping, family heirlooms, family house, family income supplement, family influence, family law, family likeness, family lines, family man, family name, family planning, family responsibilities, family room, family saga, family silver, family tree.**

Are these combinations sufficient? And could further ones not be added indefinitely? For example:

(4)   family Bible, family car, family dictionary, family loss, family mosque, family school, family table, family teacher.

This question is asked, not because the inclusion of indefinite word combinations is propagated, but because there seems to be too many superfluous entries.

Other compound words entered as headwords include those with **air** (81 combinations), **aircraft** (13), **advertising** (19), **black** (27), **blood** (51), **body** (29), **book** (29), **chief** (34), **child/children** (37), **Christmas** (27), **church** (13), **film** (36), **fire** (61), **hair** (48), **half** (80), **hard** (72), **home** (66), etc.

By giving every lexical item as a main headword, too much valuable space has been wasted. Furthermore, the so-called "senses" of a headword are all moved to the left side of a column, each beginning on a new line regardless of whether the last line of the definition of a "sense" is completely filled or not. So entry words like **get** with 53 "senses" or **make** with 38 "senses" plus 23, the so-called "various connections", add up to 61 unfilled lines for **make** alone. The entry word **take** has a total of 128 cm of unused space. If one line of a column of the dictionary covers 6.5 cm this means that the unused space for the entry **take** equals 19 (i.e. 128 cm ÷ 6.5) lines. A random selection of two pages of the dictionary reveals that page 437 has 130 cm and page 928 has 108 cm of unused space which equals 17 lines per page. Statistically the dictionary has 17 799 (i.e. 1 047 pages x 17) unused lines. Translated into pages the unused space equals 59 (i.e. 17 799 lines ÷ 150 lines per 2 columns of a page = 118 columns ÷ 2) pages. So if all the space had been utilized, the dictionary would have been reduced to 988 (1 047 – 59) pages.

The dictionary was made unnecessarily voluminous. Although by giving headword status to all derived, inflected and compound words, the compiler might have attempted to user-friendliness, this was done at the expense of wasted space whose cost the user will have to bear, because the bigger the dictionary, the more expensive it is.

It is not necessary to enter every combination a word can have but only a few combinations, especially those whose meanings are opaque, i.e. those whose meanings cannot be deduced from the meanings of the words that form them. Combinations like those of **family** given under (3) and even those of **home** listed under (5) should not all have been entered in the dictionary because the meanings of many can easily be deduced from the meanings of their component parts.

(5)   **home address:** *anwani ya nyumbani*
home entertainment: *burudani ya nyumbani*
home life: *maisha ya nyumbani*
home lover: *mtu anayependa sana nyumbani*
home loving: *-enye kupenda sana nyumbani*

The word **home** has 66 independent headwords which are either derivatives of **home** or compound words whose one member is **home** .

Compared with Kirkeby, TUKI (1996) has almost the same lexical stock. The only difference is that TUKI has been more prudent in selecting and arranging entries. As a result TUKI has fewer main entries because most of the word combinations are entered as run-on entries as shown by the following table:

Selected entries in TUKI (1996) and Kirkeby (2000)

	A	F	G	M	T	U	**Total**
	Aa–advocate	Ff–family	Gg–game	Mm–manufacture	Tt–tamper	Uu–unattached	
TUKI	150	50	50	150	50	50	500
Kirkeby	196	212	121	501	132	101	1263

When the entry words under A, between Aa and advocate, are compared, it is found that TUKI has all the items in Kirkeby except the following 77:

(6)   aardvark, aardwolf, abbacy, abbr(ev), ABC, abele, abnormally, aborticide, abortifacient *n*, abortifacient *adj*, about-ship, abreact, abreaction, abseil, absinth(e), acacia, a cappella, accepted, access key, accessories department, acclaim, acclamatory, acclimatization, accompanying, account-balancing, account customer, account executive, account holder, accounting, accounting cost, accounting item, accounting machine, accounting system, accounting term, accounting year, account manager, accounts analysis, accruals, accrued, accumulative, accusing, AC/DC, ace, acerbate, acescent, acetone, acetyl, acetyl-salicylic, achlamydeous, achromatic, achromatism, acicular, acid-base balance, acid count, acid drops, acid-fast, acidhead, acoustic homing, acoustically, acoustic image, acoustic nerve, acoustic perfume, acquaintanceship, acropolis, across-the-board, actionable, action committee, action-packed, action programme, action replay, activation, activator, actress, actuator, acupuncture

Kirkeby, on the other hand, lacks the following lexical items found in TUKI (1996) between Aa–advocate:

(7)   abb, ablate, abnegation, A-bomb, abscise, absentmindedly, absent oneself, absorption, abusively, abusiveness, abysmal, Abyssinia, academically, academicals, academician, accelerando, accidence, accordant, accretion, accrebities, acriflavine.

As regards the method of presenting entries which share a common basic form, TUKI treats all such items as run-on entries to the main entry word as shown in (8).

(8)   **accept** *vt* 1 pokea. ~ bribe pokea rushwa. 2 kubali, kiri. ~ a fact kubali ukweli. ~**able** *adj*. ~**ability** *n* ukubalifu. ~**ance** *n* kukubali, ridhaa.

Suffixes ~**able, ~ability** and ~**ance** represent the verbal derivatives of **accept**, namely **acceptable**, **acceptability** and **acceptance**.

This is in contrast with Kirkeby who enters them all as main entry words as shown in (9) below:

(9)   **accept** … *vb* **1.** -kubali *(ex sby's excuse; an invitation; a task; terms)*; -pokea *(ex students)* ~ **the responsibility** -kubali (*or* -pokea) jukumu (*or* wajibu); …
2. (= *put up with*) -stahimili; -vumilia;
3. *com*: -kubali.
acceptability … *n* ukubalifu
acceptable … *adj*: -a kukubalika; -a kupokewa; -a kukaribisha; …
acceptance … *n* **1.** kukubali; ridhaa; …
2. *com*: makubaliano; kuridhika; kukiri; kukubali; …
accepted … *adj.* **1.** -enye kukubalika *(ex truth; rule)*; …

4.   Lexicographical categories

After the selection and presentation of headwords have been looked at, the lexicographical categories given under these headwords can now be examined. The reference needs of foreign learners are basically two: receptive and productive, and any information given should try to meet these needs. The information usually found in a dictionary entry is of four types: 1. *Internal information* which presents facts about the lexicographical unit: the orthography, pronunciation and form of the headword. The form includes derivational and inflectional aspects as well as compound forms which make up one lexical item. 2. *External information* provides facts about its relationships with other lexicographical units. In this information, the syntax of a word is shown, especially the obligatory combinations. 3. *Usage information* gives facts about the way the lexicographical unit is used. This aspect deals with the context in which the entry word is used in ordinary linguistic discourse, revealing its grammatical and semantic meanings. 4. *Etymological information* provides facts about the history of the headword which includes the historical development from its origin to its current form, and the development of its senses.

Kirkeby makes available the following lexicographical information for the headwords: orthography, pronunciation, grammatical information, usage labels which include status labels, subject or domain labels, regional labels and time labels. Other information categories are meaning and illustrative examples. Etymological information is not provided. Kirkeby does not disclose the function of his dictionary, but from its content it can be deduced that it has decoding and encoding, i.e. receptive and productive functions respectively. It also gives technical language from some specialized fields such as botany, zoology, physics, etc. All this information is condensed in a pedagogical bilingual dictionary for English learners. There is no doubt that the compiler wanted to meet all the needs of different categories of prospective users of the dictionary. But the desire to serve the diverse needs of these users makes the different categories of this dictionary virtually incompatible, often leaving the users very dissatisfied. The information categories can now be examined separately.

4.1   Orthography

The dictionary gives variant spellings of some headwords and variations in British and American English, e.g. **haulier** (US: *hauler*), and **hardware shop** (US: *hardware store*). This is normal practice in lexicography, and the examples have been treated properly except that the variants and the English American variations which are on a par with the headwords, are given in italics instead of the bold type used for the headwords.

Many word combinations or compound words are entered as headwords. Most of them are written as two separate words while others are joined by a hyphen and some are spelt as a single word. For example:

(10) (a) mail sack *n*, mail order *n*, mail carrier *n*, mail van *n*, mailing list *n*, male chauvinism *n*, male chauvinist *n*, male dominance *n*, main line *n*, head teacher *n*, honey badger *n*

   (b) mail-order firm *n,* male-dominated *adj*, male-chauvinist *adj*, man-day *n,* man-eater *n,* man-hour *n*, man-hungry *n*, man-made *adj*, man-hunting *adj*, man-labour *adj,* holiday-maker *n,* home-brew *n,* home-brewed *adj*, home-loving *adj*

   (c) mailbag *n*, mailbox *n*, mailman *n*, manhandled *adj,* manhandle *vb,* mainstream *n*, mainframe *n*, mainline *vb*, homelover *n,* homesick *adj,* headmistress *n*, headmaster *n*, honeybee *n*, honeybird *n*

If the word combinations in (10)(a)–(c) are studied, it is unclear why the compound words are presented in different spellings. Why, for example, are *mail sack, mail carrier* and *mail van* written as two words while *mailbag*, *mailbox* and *mailman* as one word, especially when bearing in mind that *mail sack* and *mailbag* are synonyms as is the case with *mail carrier* and *mailman*, both meaning *postman*. Likewise the adjectives *male-dominated* and *male-chauvinist* are joined by a hyphen while the nouns *male chauvinist, male dominance* and *main line* are written as two words, and *mainframe* and *mainstream* are written as one word without a hyphen. Although some adjectival forms are joined by hyphens as can be seen in (10)(b), nominal forms like *man-eater* and *man-hour* are also joined by hyphens. Forms written as one word without a hyphen cut across all grammatical categories: adjectives (*homesick* and *manhandled*), nouns (*homelover* and *honeybee*) and verbs (*manhandle* and *mainline*).

4.2   Pronunciation

Pronunciation in a dictionary is rendered in phonetic symbols which represent the sounds of a language. Provision of the symbols is based on the assumption that the user has already mastered the sound inventory of the target language and is able to produce the acceptable sounds as they occur in familiar words. Transcribing phonetic symbols to represent the sounds of a word is difficult to

a non-phonetician, but to interpret them is much harder because many dictionary users do not have the background in theoretical phonetics needed to interpret the transcription system.

Kirkeby uses the phonetic symbols of the IPA to represent the different sounds in English words, e.g. /ɑː/, /æ/, /ʌ/, /e/, /θ/, /ð/, /ŋ/, /ʒ/, /ʃ/, etc. To be able to decipher these symbols one needs a pronunciation guide. Kirkeby's guide is not very useful in helping to pronounce the words. One would have expected that, since this is a dictionary intended to help Tanzanians whose "knowledge of English has greatly deteriorated" (see Preface), the compiler would have made a serious study of Swahili words in order to find suitable examples for producing certain sounds instead of referring to English, German or French words as shown in (11).

(11) x in words like 'loch' and 'ugh' is pronounced like **ch** in the German word 'nach'.
~ above a vowel indicates nasalisation (only in some French words).

How can Tanzanian English learners be expected to use German or French to help them with the pronunciation of English words? One would even question whether a non-linguist Tanzanian can really make any distinction between the following set of vowels: /ɑː/ and /æ/, /e/ and /ə/, /o/ and /ou/, or /u/ and /ʌ/ as in the following example:

(12) far-flung [ˈfɑːˈflʌŋ]

Likewise it is difficult for them to decipher phonetic symbols such as /ŋ/, /θ/, /ð/, /ʃ/, /tʃ/, /ʒ/ and /dʒ/.

A lexicographer intending to provide this information has to think of how best to help the users. Indeed he has to be sure whether dictionary users will actually need this information. Lexicographical research has shown that pronunciation is information that users look up the least in a dictionary.

However, the guide would have been meaningful if examples of Swahili were used. For example: **ng'** as in *ng'ombe* could have been used instead of the phonetic symbol /ŋ/, **th** as in *thamani* instead of /θ/, **dh** as in *dhana* instead of /ð/ and **kh** as in *khatibu* instead of /x/. The Swahili phonetic equivalents would have been more meaningful to the Swahili learner than the IPA. Although this would have been breaking away from the tradition, it is user-friendly as far as the Swahili English learner is concerned. It should be recalled that the English dictionary had also used the respelling method before the IPA was adopted. This can be done if one is willing to involve consultants in the field. The native speakers, journalists and amateur lexicographers used as consultants by Kirkeby have limitations.

When one looks at the pronunciation guide and the pronunciation symbols in the entries, it becomes obvious that the information was copied from another dictionary disregarding its relevance for this specific dictionary.

Further examination of the other symbols in the guide reveals some pho-

nological information such as main stress, secondary stress and nasalisation which does not form part of the general knowledge of the anticipated users and cannot be deciphered by them. This is known only to students of phonology. Provision of a pronunciation guide requires the relevance of the symbols for the anticipated users, the users' ability to interpret these symbols and the lexicographer's innovation of the pronunciation guide to make it suitable for the users.

4.3   Grammar

Grammar is one of the important information categories in a dictionary, especially one designed for learners of a language.

4.3.1   Grammatical categorization

Grammatical information is more useful to the learner of a foreign language than to the native speaker. Pedagogical and bilingual dictionaries provide more grammatical information than those for native speakers (Landau 1984). Information given in a dictionary should be helpful to the user and the user should not be confused with obscure conventions. More pragmatic information is needed about how people communicate with each other and more attention must be given to the structure of the spoken language (Sinclair 1987).

Most of the headwords in this dictionary have been classified and labeled with their word-class markers, viz. noun, verb, adjective, adverb etc., but some headwords have not been marked:

> (13) tax arrears, tax avoidance, tax bill, tax burden, tax certificate, tax concession, tax consultant, tax collector, tax dodger, tax evader, tax expert, jack plug, jack point, jack tar, jack towel, jet engine

The following items with similar forms have, however, been classified and labelled:

> (14) tax-exempt *n*, tax-finding *n*, tax-free *adj*, rag-and-bone-man *n*, jack-in-the-box *n*, jack-in-house *n*, jack-of-all-trades *n*, jet-propelled *adj*, jet-black *adj*

There are no linguistic explanations that discriminate (13) from (14). Although one could argue that the examples under (14) (with a hyphen) tend to develop towards (becoming) one-word compounds, it could still be argued that the two sets (with or without a hyphen) do not differ because all of them are lexical units representing one concept just like single lexical units.

4.3.2   Subgrammatical categorization

Subgrammatical categorization is very important for a pedagogical dictionary. It is not enough to classify words into seven or eight categories as if words in

each category behave the same morphologically and syntactically, especially the major categories viz. nouns, verbs and adjectives. Language learners need to be informed whether a noun is countable or uncountable and whether it has a plural and what the plural is. Users need to be told whether a verb is transitive or intransitive and what its conjugation forms are. Information is also needed about attributive and predicative adjectives or their comparative and superlative forms. Morphological aspects of lexical items such as inflections and derivational affixes is furthermore important information that should be given.

Kirkeby's dictionary lacks information about verbal conjugation in its entries. This information is necessary if it is to be used efficiently for production purposes. Kirkeby does indicate plural inflections for the irregular nouns, e.g. *wife (wives), thief (thieves)*, but this has not been done consistently. Some nouns e.g. *directory, fly, family* etc., do not have their plural inflections indicated. Moreover, the compiler does not mark nouns which are countable and those which are uncountable, cf. *furniture, sympathy, butter, water*, etc. This information is useful to learners of English because it helps them to select the appropriate agreement between a noun and the verb.

As for the comparative and superlative forms of adjectives, the dictionary is very inconsistent. At some adjectives like *big* and *bad* this information is given while many adjectives such as *fast, poor, pretty, quick, safe, tall, thin* etc. lack this information. Even the irregular adjectives which do not take *-er* and *-est* but are preceded by *more* and *most* for the comparative and superlative forms such as *beautiful, handsome, competent*, etc. have also not been provided with this information. Even the so-called regular forms undergo morphological changes when the comparative and superlative inflections are suffixed to them, hence it is important to indicate them. For example:

(15) tall, tall**er,** tall**est;** safe, safe**r**, safe**st;** pretty, prett**ier,** prett**iest;** thin, thin**ner,** thin**nest**; beautiful, **more** beautiful, **most** beautiful; competent, **more** competent, **most** competent; little, **less, least**.

The adjectival examples show forms which take ~*er* and ~*est*; others take ~*r* and ~*st* or ~*ier* and ~*iest* while still others double the last consonant and then add ~*er* and ~*est.*

Because all this information is important for the English learner, a pedagogical dictionary should supply it. Kirkeby completely neglected this information.

4.3.3   Swahili grammatical information

Kirkeby has attempted to enter some Swahili grammatical aspects in the Swahili equivalents of the English headwords. These are the plural affix of nouns. For example**:**

(16) **friend** *n* rafiki (ma-); **hand** *n* kiganja (vi-) cha mkono

The affixes in brackets after the Swahili equivalents (see *rafiki* and *kiganja* under (16)) are the plural affixes of the headwords. This means that the plural of *rafiki* is **ma**rafiki and that of *kiganja* is **vi**ganja. Legere (1990) has done the same in his dictionary. This is relevant only when dealing with single lexical units like *rafiki* and not with multilexical units like *kiganja cha mkono* because, when this lexical unit changes into plural every item in it is affected, as shown in (17).

> (17) kiganja cha mkono *n > vi*ganja *vy*a *mi*kono

There is no *viganja cha mkono* as implied in (16).

If the compiler felt that it was necessary to show the plural form of the Swahili noun, he should have considered how best to deal with this type of lexical unit. Should all the plural affixes and concordial agreements for the headword be indicated as in (18)?

> (18) kiganja cha mkono *n* (vi- vy- mi-)

Or should he rewrite the whole multi-lexical unit with its plural prefixes as in (19) if he feels that the affixes are not communicative enough?

> (19) kiganja cha mkono *n* (viganja vya mikono)

Alternatively, the plural prefixes can be eliminated and instead an elaborate front matter with a sketch of Swahili grammar, giving the classification of nouns can be included. Hopefully, with such an alternative, the user will consult the front matter for clarification when in need. What is important here is that any information in a dictionary should be communicative.

4.3.4   Word combinations and collocations

Syntax of the headword is usually shown through verbal phrases, collocations and illustrative examples. This information is important because it shows words found in combination with the headword. In Kirkeby's dictionary many verbal phrases, most of which are idiomatic in character, have been arranged alphabetically as headwords and translated. Because English is rich in verbal phrases and idiomatic expressions, many verbs have a significant number of these multilexical items. For example *look* has 30 verbal phrases such as:

> (20) *look about for sth, look across, look after, look ahead, look around, look as if, look at, look back, look before you leap!, look down, look for, look forward to, look here!, look in, look into, look like, look lively!, look nearer home, look on, look out, look out for, look over, look round, look sharp, look through, look to, look up, look up to sby, look sby up and down, look upon.*

Whereas these word combinations are important in encoding and decoding, not all deserve a place in the dictionary because their combinations are neither obligatory nor have an idiomatic connotation. For example, *look at*, *look across*,

look down, *look into* or *look up* are phrases that should not have been listed because the meaning of each phrase is the total sum of the meaning of the words that form the phrase. Indeed one could replace *look* with another verb and get the meaning of the two. For example:

> (21) wave at sby, walk across, go down, put into, climb up.

Although collocations and especially verbal phrases are important in a dictionary, Kirkeby should have selected and entered only those which have fixed meanings. Words which take certain prepositions, should be indicated, e.g. *look* (*~ at, ~ up*).

4.4   Illustrative examples

Illustrative examples are given to support and in some cases to supplement definitions. They help foreign learners to understand the meaning of a word. They serve to show the usual collocations or contexts in which each sense is used, thus providing a variety of grammatical information such as whether a verb takes an indirect object or is usually used in the passive voice, or whether an adjective is predicative or attributive. Examples are therefore given under headwords which have collocations that provide specific syntactical information or can help to distinguish different senses of a polysemous lexical item. Illustrative examples are also given to function words which have no lexical but only grammatical meaning. Familiar words are usually not exemplified because their meaning and usage is known unless they have developed new senses. For an example sentence to be useful, it must give a clear idea of the meaning (which has become polysemous). It must be short and to the point but sound authentic. Familiar words must also have few collocations and a fair amount of grammatical information.

Kirkeby provides many examples of usage that cut across different word classes. For example:

> (22) **appea**r … *vb* **1.** -jitokeza; **he -ed out of the blue** alijitokeza ghaf(u)la; … **4.** *on the radio, on TV or in the theatre:* -onekana; **he first -ed on a television variety show** yeye alionekana kwa mara ya kwanza katika kipindi cha michezo cha televisheni (*or* runinga); **5.** -onekana; **you can't ~ in** (= *be seen with*) **that hat** huwezi kuonekana na hiyo kofia; **6.** *law:* ~ (**in court**) -enda mbele ya mahakama; **7.** (= *seem*) -elekea kuwa; **he -s to be wrong** (= *it seems as if he's wrong*) anaelekea kuwa mwongo; … **8.: it -s from what he said that** … kutokana na alichosema inaonekana kwamba …; **9.** *formal* (= *be mentioned*) -tajwa; **my name must not ~ on the cases** jina langu lisitajwe kwenye kesi.

Illustrative examples in this dictionary are no doubt very useful to the user. Many of them show collocations of the headword or its valence i.e. its syntacti-

cal patterns. The headword **appear** under (22) is a case in point. This example shows the different prepositions the headword takes:

(23)   **appear:** ~ out, ~ on, ~ in, ~ to, ~ from.

However, these examples should have a purpose. A critical examination of the dictionary reveals superfluous illustrative examples. Compare the examples of *hand*, *shake*, *potato*, *immediate*, *jacket* and *jeer* under (24).

(24)   **hand: hold -s** -shikana mikono; … **~ hold -s with a girl** -shikana mikono na msichana; **shake -s** -peana mikono; **shake -s with sby** -peana mikono na mtu; **shake -s on the deal -**peana mikono baada kwa kukubaliana jambo

immediate: 1. ~ help msaada wa mara moja; **take ~ action** chukua hatua za mara moja; **2. the ~ family** ndugu wa karibu; **in the ~ future** siku za karibuni; **my ~ neighbours** majirani zangu wa karibu; **her ~ surroundings** mazingira yake ya karibu

jacket: potatoes baked in their ~ viazi vilivyookwa na maganda yake; **potatoes boiled in their ~** viazi vilivyochemshwa na maganda yake

jeer: ~ at sby -dhihaki mtu; **~ at sby's stupidity** -dhihaki ujinga wamtu

The entries cited above show cases of examples which do not offer new semantic or grammatical information but repeat information already provided in some of the other examples.

5.   Meaning

Meaning of an entry word in a bilingual dictionary is rendered by an equivalent lexical item in the target language. In the absence of exact equivalents in the two languages, glosses can be used as translation equivalents (Bartholomew and Schoenhals 1983). Thus the number of senses in the source language are usually reflected in the target language in the dictionary. Likewise, the phrases and fixed expressions in the source language are given their equivalent translations in the target language. In this section, the senses given as Swahili equivalents of the English headwords, the manner in which they have been presented and the degree of their comprehension to learners of English will be examined. The entry words in the dictionary have one or more than one Swahili equivalent. For example:

(25)   **alcohol** … *n:* alikoholi; kileo; pombe; kilevi; …
   alcohol abuse *(= abuse of alcohol)* unywaji wa pombe kupita kiasi.
   I. alcoholic … *n:* mlevi (wa-) sugu.
   II. alcoholic *adj:* -a kulevya; -a kileo; -enye asili ya kulevya.
   alert *adj* **1.** *(= wide awake; fully awake)* -liyemacho; -liyeamka *(ex I 'm not feeling very alert this morning because of lack of sleep);*
      2. *fig (= wide-awake; sharp)* -liye makini *(ex some alert readers noticed the mis-*

> *take*); **he is a very ~ child** yeye ni mtoto aliye makini; **he's an ~** *(= a wide-awake)* **person, intelligent, resourceful and not afraid to act** yeye ni mtu mwenye akili, mwenye mawazo mazuri, asiyeogopa kufanya jambo; **with all senses ~** makini kabisa; …

From the examples under (25) it can be seen that *alcohol* has only one sense which is translated by four Swahili synonyms: *alikoholi*, *kileo*, *pombe* and *kilevi*. The headword **alcoholic** *n* has only *mlevi sugu* while **alcoholic** *adj* has the synonymic equivalents: *-a kulevya*, *-a kileo* and *-enye asili ya kulevya*. The Swahili equivalent *alikoholi* unlike its other synonyms is not a common but a technical word, hence it needs to be marked in order to alert the user that it has a re-stricted usage. Although the compiler has extensively marked words with a special usage, the marking of such words has not been consistent. Moreover, *alikoholi* should appear last in the list of synonyms because it is a rare word. Common synonyms should appear first because they are the ones which users may most often want to look up in the dictionary.

The presentation of Swahili equivalents has in some cases been preceded by synonymic paraphrases of the senses. This can be illustrated by *alcohol abuse* and the senses 1 and 2 of *alert* under (26).

> (26) **alcohol abuse** *(= abuse of alcohol)* unywaji wa pombe kupita kiasi.
> **alert** *adj* **1.** *(= wide awake; fully awake)* -liyemacho; …
>   **2.** *fig (= wide awake, sharp)* -liye makini …

The paraphrases are meant as elaborations (glosses) of the headwords before the Swahili equivalents are given. With the help of these paraphrases, the reader is expected better to understand the headword in the source language before coming to the Swahili equivalents. This is probably possible for the speaker or learner who already knows English and understands fixed expressions such as *wide awake* or *fully awake*. A Swahili speaker learning English will find the paraphrases useless because he does not know what the fixed expressions mean. Even for the English speaker, 'wide awake' in 1 and 2 would not make any difference. Indeed, *-liyemacho* can be used figuratively as well. Moreover, the illustrative sentence in *alert* sense 2 (see (27)) is neither helpful to the reader nor a good example of a well-constructed grammatical English sentence.

> (27) **he's an ~ person, intelligent, resourceful and not afraid to act** yeye ni mtu mwenye akili, mwenye mawazo mazuri, asiyeogopa kufanya jambo.

The Swahili translation is also confusing. Do the phrases following 'mtu mwenye akili' qualify this phrase or are they different interpretations of the English sentence: 'he's an alert person'. Another feature in the entry **alert** which is also found throughout the dictionary, is the English illustrative examples in brackets preceded by '*ex*' that are juxtaposed in the entry after the Swahili equivalent. These examples are not translated into Swahili. See senses 1 and 2 of **alert** under (28):

(28) **alert** *adj* **1.** *(= wide-awake; fully awake)* -liyemacho; liyeamka; (*ex I'm not feel-*
   ing very alert this morning because of lack of sleep);
   2. *fig* (*= wide-awake; sharp)* -liye makini (*ex some alert readers noticed the mis-*
   take) …

The elaboration principle indicated by the equation marks is either restating
the headword or paraphrasing it. For example:

(29) **alcohol abuse** *(= abuse of alcohol)*
   alert *(= wide-awake; fully awake)*
   fame *(= celebrity; renown)*
   famed *(= well-known; famous)*
   family allowance *(= child benefit)*
   family silver *(= silver heirlooms)*
   fancied *(= imaginary)*

Even if the words in brackets were elaborating the headwords, such words are
often more difficult than those being elaborated. The words *celebrity* and *re-*
nown are no more common than the headword *fame.* Likewise *silver heirlooms* is
more difficult than *family silver* and *child benefit* than *family allowance.*
    A critical examination of the Swahili equivalents reveals a number of
shortcomings in the equivalents selected and the manner of presenting them. In
the following section, the Swahili equivalents and how they have been pre-
sented will be examined.

5.1   Sentences used as Swahili equivalents

The main objective of a bilingual dictionary in rendering the lexical meaning of
the headword is to provide equivalent words in the target language. Phrasal
equivalents are accepted as translation equivalents in the absence of one-to-one
lexical units. The significance of using a lexical or multilexical unit is that such
items can easily be applied in translations from one language to another. Defi-
nitions of headwords are confined to monolingual dictionaries because their
main objective is to paraphrase the headword in such a way that one can un-
derstand its meaning. This is in contrast with a bilingual dictionary where one
is expected to know the meaning of a lexical item in one of the languages and
only wants the equivalent word. Kirkeby has in many cases used full-fledged
definitions as Swahili equivalents for English headwords. For example:

(30) **buffet 2.** *in restaurant or hotel:* chakula (vya-) kilivyowekwa mezani ili walaji
   wajihudumie
   call loan mkopo (mi-) unaoweza kudaiwa wakati wowote
   call mark namba (-) ya kitabu inayoonyesha mahali kitabu kinapopatikana
   call meter mita (-) ya kuhesabu dakika katika matumizi ya simu
   call sign fungu (ma-) la maneno na tarakimu zinayotambulisha utangazaji
   redioni

> **cable television** huduma ya televisheni kwa njia ya kebo bila antena
> **callboy** mtu (wa-) aliye pembeni ya jukwaa ili kuwahimiza waigizaji
> **chaps** US: ovaroli za ngozi zivaliwazo na wachunga ng'ombe katika tam-
>   barare za Amerika

There is no reason for such long descriptions which could have been shortened. If the language does not have words for the headwords in the source language, simple coinages would have been better. For example:

(31) **buffet**: chakula cha kujihudumia
      call loan: mkopo wa kulipa haraka
      call mark: alama bainishi
      call meter: mita ya simu
      call sign: ishara ya kutangaza
      cable television: televisheni ya kebo
      callboy: mhimiza waigizaji
      chaps: US: bwelasuti la ngozi

5.2   Superfluous equivalents

The dictionary has some examples of good Swahili equivalents for certain English headwords. Unfortunately these are followed by further paraphrases of these shorter equivalents. In most cases such descriptive paraphrases are repetitive and long-winded.

(32) **calf love** mapenzi (-) ya kitoto; mapenzi ya kijingajinga kwa vijana
      caller mgeni (wa-); mtu (wa-) ajaye kusalimu
      bug *colloq:* virusi *(pl)*; chembe maradhi (vye-)

These further paraphrases of the headwords under (32) which have already been rendered by the Swahili equivalents *mapenzi ya kitoto*, *mgeni* and *virusi* are unnecessary.

5.3   Ambiguous or misleading Swahili equivalents

Some Swahili equivalents are ambiguous or misleading to users, not representing exactly what the headword denotes. They share some semantic features of the headword but cannot qualify to be the equivalent of the headword.

(33) **hard currency** fedha za kigeni
      hard palate the ~ kaakaa (ma-) la juu
      hard sell ugumu wa kuuza
      off camera nje ya kamera

The Swahili equivalent of *hard currency* is not *fedha za kigeni* (foreign currency) but *convertible currency*, i.e. currency whose value cannot fall suddenly (*Oxford*

Advanced Learner's Dictionary of Current English 1988: 568). The term *fedha za kigeni* was coined from the Tanzanian perspective, because foreign currencies such as the US dollar, British pound or European Union's euro are to the Tanzanians *fedha za kigeni* (foreign currencies). The term cannot be used in the same way by Americans, British or Europeans when referring to their own currencies. To them the currency is not *foreign currency*, but *hard currency* because it is convertible.

Hard palate is a term contrasted with *soft palate*. In Swahili the equivalents of these phonetic terms are *kaakaa gumu*, and *kaakaa laini* respectively. It is therefore not correct for *hard palate* to be translated as *kaakaa la juu* (the upper palate) as if there is *kaakaa la chini.* Moreover, the dictionary lacks the term *soft palate*.

Hard sell means "difficult to sell". This is used when referring to goods which sell very slowly. The Swahili neologism cannot be *ugumu wa kuuza* because this is a literal translation of the words *hard* (ugumu) and *sell* (kuuza). The headword in English is *sell* and the qualifier is *hard*. The headword in the Swahili equivalent should have been *kuuza* and the qualifier *gumu*. So the Swahili equivalent could be *kuuza kugumu* which is better than *kuuzika kwa shida.*

Off camera is incorrectly translated as *nje ya kamera*, literally meaning "outside the camera" while it actually means "away from the camera" or "not to be photographed". An appropriate equivalent could have been *bila kupigwa picha*.

Other headwords which have not been given appropriate Swahili equivalents are given under (34).

> (34) **camouflage** majificho
> **Caesarean birth** uzazi wa kupasuliwa
> **cafeteria** mahali pa kulisha watu wengi

5.4   Incorrect description of headwords

Some descriptions of headwords provide false information. For example:

> (35) **drama** drama; riwaya
> **hair root** kisiki cha nywele; mzizi wa nywele
> **hair piece** kipisi cha nywele
> **I. Gambian 2.** *language:* Kigambia
> **Senegalese 2.** *language:* Kisenegal

Swahili has adopted *drama* as the equivalent of the English word *drama* also given as headword. The term *riwaya* was coined to represent the English word *novel.* It is therefore misleading to translate *drama* with *riwaya* as well.

Hair root cannot be defined as *kisiki* (stump) *cha nywele* (of hair) because *kisiki* is neither the equivalent of *root* in Swahili nor the appropriate translation of *hair root*. The appropriate equivalent should have been *mzizi wa nywele.* The translation *kipisi cha nywele* as equivalent of *hair piece* cannot be accepted because *kipisi* is not a Swahili word. *Kigambia* and *Kisenegali* as languages of Gambia and Senegal respectively are questionable. There are no such languages.

Many African countries were created and given names which do not conform with the name of the people and languages spoken in these territories, cf. Tanganyika, Kenya, Uganda, Sudan, Namibia, Mozambique, etc. This is in contrast with many European countries, e.g. Ireland (Irish), Spain (Spanish), England (English), etc.

5.5   Long descriptions as Swahili translation equivalents

The dictionary has some headwords with long descriptions which are appropriate for a monolingual dictionary but not for a bilingual dictionary whose objective is to help the reader to translate from a source language to a target language. Compare the examples under (36):

(36) **flatmate** *n* mtu (-) wa kuishi naye fleti moja *(= friend with whom one shares a flat)*

flat race mbio-za farasi ambazo hakuna kuruka vihunzi *horse race without hurdles*

flight crew wafanyakazi (wa-) ndani ya ndege *workers in a plane*

flight attendant mhudumu (wa-) ndani ya ndege *attendant in a plane*

flight handler mfanyakazi wa kupakia na kupakua mizigo na abiria uwanja wa ndege *worker who handles passenger's luggage at the airport*

flight instructor mwalimu (wa(a)limu) anayetoa mafunzo ya kurusha ndege *a teacher who teaches how to fly a plane*

flight lieutenant rubani (ma-) wa kijeshi mwenye cheo cha luteni *pilot of the airforce with the rank of lieutenant*

harbour charges ushuru wa kutia nanga *fees for anchoring at a port*

These paraphrases could be shortened to multiword equivalents as shown under (37):

(37) **flatmate** mkazi mwenza wa flati *flatmate*

flight crew wafanyakazi wa ndege *flight crew*

flight attendant mhudumu wa ndege *flight attendant*

flight handler kuli wa ndege *flight handler*

flight instructor mkufunzi wa kurusha ndege *flight instructor*

flight lieutenant luteni wa jeshi la anga *flight lieutenant*

harbour charges ushuru wa bandari *harbour charges*

5.6   Use of Swahili proverbs to describe headwords

Proverbs are fixed expressions with figurative meanings which cannot be deduced from the meanings of the words which form them. It is therefore unwise to use them to describe headwords because their meanings would not be understood. For example:

(38) **flatter** *vb:* **~ sby** -visha mtu kilemba cha ukoka; -paka mtu mafuta kwa mgongo wa chupa *make sby feel big*

> **flattery** *n:* kilemba cha ukoka; … kupakwa mafuta kwa mgongo wa chupa
> *be made feel big*
> **meat 5.** *colloq: easy meat (= an easy target)* kuku mweupe

These headwords could simply be described as follows:

> (39) **flatter** *vb* **~ sby** sifu mno; sifu kupita kiasi *praise falsely*
> **flattery** sifa za uwongo, kusifu kupita kiasi *false praise*

The glosses in (38) are correct except that they are proverbs which cannot easily be understood. Therefore equivalents in plain English as in (39) are proposed.

5.7   Unacceptable Swahili equivalents or translations

Some Swahili equivalents cannot be regarded as appropriate translations because they are ungrammatically constructed and do not accurately represent the meaning of the English headwords.

> (40) **I. change** *n:* **1.** mabadiliko …; **complete ~** mabadiliko kabisa; <u>mabadiliko mazima</u> *whole change*
> **chaplain** … **army ~** *(= chaplain to the forces)* mchungaji (wa-) aliyeambati-shwa kwenye taasisi *pastor attached to an institution*; **prison ~** mchungaji wa kanisa <u>wa kwenye</u> gereza.
> **character assassination** uharibifu wa sifa *destroying of character*

Acceptable equivalents would have been:

> (41) **complete change** mabadiliko kamili *complete change*
> **army chaplain** kasisi wa jeshi *army pastor*
> **prison chaplain** kasisi wa jela *prison pastor*
> **character assassination** kuharibu sifa ya mtu *tarnish sby's character*

6.   Miscellaneous errors

Some typographical and grammatical errors appear in the dictionary. There are also factual errors as well as omissions of some entry words. For example:

> (42) **alert** *adj* **1.** *(= wide-awake; fully awake)* -<u>liyemacho</u>
> **cajole** *vb* <u>raiai</u>
> **booking** *n* **3.** **the player received a ~** *(= the player was booked)* mchezaji <u>aliyeandika</u> jina kwa kuvunja sheria
> **chandalua** *n;* (in East <u>frica</u>) *(= mosquito net)* chandarua
> **half-bound** *adj; bookbinding* liojadiliwa kwa <u>ngoz</u>
> **disappointment** *n* … **a bitter ~** masikitiko makali; **it was a great ~** <u>ilikuwa</u> jambo la kukatisha tamaa sana; **a bitter ~** masikitiko makali; **it was a great ~** ilikuwa jambo la kukatisha tamaa sana; … **her ~ was obvious from her face** kukata tamaa kwake <u>kulionyesha</u> bayana usoni mwake

Some typographical errors are the following: The equivalent of **alert** should be written as two words *-liye* and *macho* and not one word *-liyemacho* as given in the dictionary. The entry words above show examples of typographical errors such as *raiai* (instead of *rairai* under **cajole**), *frica* (instead of *Africa* under **chanda-lua**) and *ngoz* (instead of *ngozi* under **halfbound**). Under the entry **disappoint-ment** two illustrative examples are repeated: **a bitter disappointment** and **it was a great disappointment**.

Some grammatical errors include the following: For the wrong Swahili translation, e.g. the player received a booking *mchezaji aliyeandika jina kwa kuvunja sheria* (the player who wrote his name for breaking the rule), the proper translation is *mchezaji aliandikwa jina kwa kuvunja sheria* (the player's name was written for breaking the rule). Likewise the translation of the illustrative sentences: **Her disappointment was obvious from her face** *kukata tamaa kwake kulionyesha bayana usoni mwake* (his disappointment was shown clearly on his face) and **it was a great disappointment** *ilikuwa jambo la kukatisha tamaa sana*, should have been *kukata tamaa kwake kulionekana bayana usoni mwake* (his disappointment was obvious on his face) and *lilikuwa jambo la kukatisha tamaa sana* (it was very disappointing).

Omissions include the entries between headwords **great** and **grow**.

Usually errors in such a large-scale work are inevitable. However, some simply show that the compiler was not careful enough in constructing the entries and editing the dictionary manuscript. Given the time frame for this project, two and a half years, it is obvious that the work was done rather hurriedly.

7.   Dictionary and standardization

A dictionary is a historical document and a standardizing tool. It records the language as it was used when it was first written and/or as it is used today. The dictionary also helps to standardize a language and show or teach the user the accepted (standard) form of the language if that has been determined. This explains why people always make reference to the dictionary whenever they are in doubt about the spelling, meaning or usage of words.

A lexicographer should be able to determine and show the most accepted form of a word where there may be variants. Kirkeby, however, instead of determining the most accepted form and giving it as standard, has entered all variants:

(43) **cake** *n* US: **apple ~** keki ya tofa(h)a
      I. **calm** 2. *naut:* utulivu; ush(u)wari
      calumny *n* kash(i)fa
      calumnious *adj* -a kash(i)fa
      flight instructor mwalimu (wa(a)limu) anayetoa mafunzo ya kurusha
        ndege

These entries show different variants of Swahili words: *tofaha* and *tofaa*, *ushwari* and *ushuwari*, *kashifa* and *kashfa*, and *walimu* and *waalimu*. Any Swahili scholar knows that *walimu, tofaa, kashfa* and *shwari* are the most accepted forms and the compilers of Kirkeby should have been aware of this and not introduced little-used variants. Since Swahili has dialects, variant pronunciations are inevitable. A pedagogical dictionary should aim at the standard form and avoid the inclusion of variants unless there is evidence that all variants are regarded as accepted standard forms, or unless the stated policy of the dictionary is to record all variants. Kirkeby's dictionary does not have such a policy.

8.   Sense discrimination and arrangement

Senses of an entry word are usually introduced by numbers, letters or semicolons. Numbers show that the senses are more distant in relation to each other, while letters and semi-colons suggest that the senses have a closer relationship. Kirkeby uses numbers and semicolons to distinguish different senses of the headwords. Example (44) shows four senses of *depress* as found in Kirkeby.

> (44) **depress** … *vb* **1.** *(= press down)* -inamisha chini; -gandamiza chini *(ex a pedal)*;
>   **2.** *fig (= sadden)* -huzunisha; -sikitisha; -vunja moyo; …
>   **3.** *com*: -shuka; *of price:* -anguka;
>   **4.** *med.; of heart activity*: -shusha mapigo ya moyo.

The dictionary also shows other examples of headwords with multiple senses. Cf. the senses of *devil* in example (45).

> (45) **devil** … *n* **1.** *also:* shetani (ma-); *rel*: ibilisi; shetani;
>   **2.** *fig*: adui (-) mbaya; …
>   **3.** *derog:* **he's a lazy ~** yeye ni mvivu sana; …
>   **4.** *sympathetically:* **the poor ~** *(= bugger)* **died of heart failure** maskini ali-kufa kwa ugonjwa wa moyo; …
>   **5. give the ~ his due** -pa mtu haki yake hata kama mwenyewe ni mbaya; …
>   **6.** *proverb:* **better the ~ you know (than the one you don't)** zimwi likuju-alo halikuli likakwisha
>   **7.** → **16.**

The headword **devil** is said to have 16 senses but in reality it has only two, 1. and 2. in the example above. The rest (3. → 16.) are fixed expressions in which the headword is used. It is therefore misleading to indicate that the word has sixteen senses. It is important to differentiate senses of an entry word, and fixed expressions, idioms or proverbs in which the headword occurs. A dictionary can introduce the senses with numbers, and then give the fixed expressions, idioms or proverbs under a subtitle.

9.   The use of labels in the dictionary

Labels are used in a dictionary to draw the attention of the user to the limited use of some words. The labels could be categorized into the following: (a) *currency label* which indicates the dimension of time, e.g. obsolete/archaic; (b) *domain label* which shows the field to which the word belongs, e.g. law/music; (c) *evaluation label* which indicates the speaker or writer's attitude, such as pejorative/appreciative; (d) *figuration label* which shows the type of meaning, e.g. literal/figurative; (e) *regional label* which shows where an item is mainly used (US/Brit. or Tanzania/Kenya); (f) *register label* indicates the manner of speech and writing in which the word is used, such as formal/informal; and (g) *status label* which shows whether a word belongs to the standard language or to a subset (slang/colloquial/dialect/offensive, etc.).

Kirkeby applies labels to mark the usage restriction of words. These labels are important and useful to the user. However, they need to be presented consistently. When the headwords in the dictionary are looked at more closely, it becomes obvious that some words of the same status have been marked and others not. For example:

(46)   **machine-cut** *engin:* -liokatwa kwa mashine
machining allowance *engin* posho ya kufanyakazi kwa mashine
machine-finished *of paper* -liomalizwa kwa mashine
machine-glazed *of paper* -liong'arishwa kwa mashine (liongarishwa)
machine part kipuri cha mashine
machine-readable -enye kusomeka kwa mashine
mediocre *derog.* -a hivi. hivi, -a kadiri; sio-zuri sana
mediocrity *n:* uduni; uhafifu; ukadiri
social democracy *polit*: demokrasia ya jamii
social democratic *polit* -a demokrasia ya jamii
socioeconomic -a kijamii na kiuchumi
social intercourse (*formal*) mwingiliano wa kijamii
social ladder mfumo wa kitaba katika jamii

Words whose domain seems to be easily deduced have been labeled, e.g. *machine allowance* or *social democracy*, but those such as *machine part* or *social ladder* whose domain is less easily determined, have not been marked. Words of the same status like *machine-cut* on the one hand and *machine-finished* and *machine-glazed* on the other have been treated differently. The former has been labeled but not the latter.

It seems as if the compilers could not make a distinction between words with restricted usage and ordinary words. Almost every word has a domain, even the everyday words we use. Cf. *cook, boil, fry, roast, bake,* etc. (cookery); *ugali, nyama choma, biriani, makande, ndizi choma,* etc. (foods); *dengelua, buza, mbege, tembo, kangara, chimpumu, ulanzi, gongo,* etc. (drinks); or *plane, gauge, T-square, square, mullet, hammer, screw driver, chisel,* etc. (carpentry). Such words need not be labeled because they are ordinary words. Users know their domain and that they are no longer restricted to specific users. Kirkeby has also labeled words of this type. For example:

> (47) **solar eclipse** *astr* kupatwa kwa jua
>   **lip** *anat* mdomo (mi-)
>   **meeting house** *rel* jengo (ma-) la mikutano
>   **melon** *bot* tikiti (ma-) maji

Looking at the examples under (47), one is tempted to ask whether users really need to be told that *solar eclipse*, *lip*, *melon* and *meeting house* are specialized words with restricted usage. Actually, many of the labeled words in this dictionary are no longer specialized. Cf. (48):

> (48) **machine accounting** *bookkeeping* uwekaji mahesabu kwa kutumia mashine
>   **machine-cut** *engin* -liokatwa kwa mashine
>   **machining allowance** *engin* posho ya kufanya kazi kwa mashine
>   **social democracy** *polit*: demokrasia ya jamii
>   **social democratic** *polit* -a demokrasia ya jamii

A general dictionary like this one should rather mark status, region, currency, figuration or register.

10.   Discussion

Kirkeby's dictionary has most of the properties required of a dictionary. It contains more than 50 000 entries which are arranged alphabetically within 1 047 pages. It is indeed the first English–Swahili dictionary which provides pronunciation of the English headwords. However, the first major shortcomings of the dictionary are: its failure to identify (a) the users for whom the dictionary has been compiled, (b) the dictionary needs of that group and (c) their dictionary-using skills. The compilation of a dictionary of this magnitude should have been preceded by a limited survey to try and find out the information categories suitable for the expected users. After having examined different aspects of the dictionary, some of the questions raised at the beginning of this article can now be answered.

(a) Is the information required by the users present in the dictionary?

Although the purpose of Kirkeby is not stated, its structure shows that it is a pedagogical dictionary. It contains illustrative examples and grammatical categorization of the headwords. However, it lacks subgrammatical information necessary for the learner of English (see par. 4.3.2). As far as grammatical information for the learner of Swahili is concerned, it is not very useful either because even this is not fully rendered (see par. 4.3.3). As for the Swahili equivalents, the dictionary fails the user here too because many of the headwords are defined instead of being given equivalent lexical or multilexical units (see parr. 5.1 and 5.5). Furthermore proverbs have been used to define entry words (see par. 5.6).

(b) Is the information transparently accessible?

Although pronunciation is important for the user, its presentation in this dictionary is very cumbersome, especially for one who is not well versed in the IPA. Moreover, reference to German or French words to help pronounce English words as shown in the pronunciation guide does not assist a Swahili learning English. It seems that the compilers did not try to find out the dictionary-using skills of the expected users of the dictionary before giving lexicographical information through graphic symbols. As a result, a user cannot decipher these symbols to reproduce the sounds. The provision of pronunciation information considered as an advantage to this dictionary in comparison with other similar dictionaries, therefore reflects negatively on its design because it is information not easily accessible.

(c) How is the information presented?

The information is orderly presented except when other information inserted in between distorts the sequence of presentation. For example:

(i)   A paraphrase of the headword has sometimes been inserted between the grammatical/subject label and its equivalent:

(49) **merchant** *n* (= *wholeseller*) mfanyabiashara ya jumla
    merchantable *adj* (= *saleable*; *marketable*) -a kuuzika, -a kununulika.
    merchant ship *n* (= *merchant vessel*) meli (-) ya bidhaa

(ii)   At times additional English illustrative sentences introduced by *ex* (meaning *example*) which are not translated into Swahili are inserted between the Swahili equivalent and another illustrative example:

(50) **meet: ~ the eye** -ona *(ex a bleak landscape met the eye)*; **our eyes met** macho
    yetu yalikutana
    measure *n* **1.** kipimo (vi-) *(ex a glass measure for holding liquids)*; **common ~**
    kipimo cha kawaida

This practice is contrary to the dictionary-making principle in bilingual lexicography which requires a sentence in the source language to be translated into the target language.

Sometimes such examples are not sentences but phrases which indicate the context in which the headword can be used.

(51) **measure: ~up** (= *measure out*) -pima *(ex ten metres of material)*

This juxtaposition of extra information overloads the entry to such an extent that it becomes cumbersome for the user to decode it and thus to use the dictionary efficiently.

(iii) As a rule, senses of a headword are usually introduced by a number, and users can easily tell the number of senses of a word. In this way, they can

scan through the entry to find the sense they want. Kirkeby's dictionary applies this principle, but mixes senses with phrases. Distinction is not made between senses of a word and word combinations/fixed expressions, hence all are introduced sequentially by numbers without differentiating senses from word combinations/fixed expressions (see par. 8).

The quality of the information presented and the effectiveness of the presentation of the information can now be critically examined. The work has some typographical errors (see par. 6). The compilers failed to determine forms which are more accepted than others, especially where there are variants (see par. 7) or headwords with similar characteristics (compound words) spelt differently, as one word, two hyphenated words or two separate words (see par. 4.1). The dictionary also contains inaccurate information (see parr. 5.3, 5.4 and 5.7). These all show that the compilers often did not seriously attempt to collect correct data.

The omission of the entries between the headwords **great** and **growth** is considerable: about 80 entries if the missing entry words are counted in TUKI (1996), but compared with *Oxford Advanced Learner's Dictionary of Current English* (1990) the omitted words are 180. The incomplete grammatical information (see parr. 4.3.2 and 4.3.3) is also a serious shortcoming because this information category is important for a productive dictionary. Very little is known about the consultants and advisers of this dictionary because it is just a replica of an English–Norwegian dictionary. But for the Swahili part, the consultants and advisers are identified as two native speakers of the language, a journalist, a school teacher, a banker and a computer expert. The compilers should have consulted more people than these, especially linguists and lexicographers.

11.   Conclusion

Having been compiled according to the principles of lexicography, the dictionary is a good contribution to the development of Swahili. However, its quality and effective presentation of the lexicographical categories show some flaws. The collection and processing of the data have not been carried out carefully and rigorously. This is partly because it was done in a limited timespan and without qualified experts. This work, if improved, can be a good dictionary for the learner of English.

Bibliography

Barnhart, C.L. 1962. Problems in Editing Commercial Monolingual Dictionaries. Householder, F.W. and S. Saporta (Eds.). 1962: 161-81.

Bartholomew, D.A. and Schoenhals L.C. 1983. *Bilingual Dictionaries for Indigenous Languages.* Mexico: Instituto Linguistico de Verano.

Courtney, R. (Ed.). 1983. *Longman Dictionary of Phrasal Verbs*. Harlow: Longman.

Cowie, A.P. and R. Mackin (Eds.). 1975. *Oxford Dictionary of Current Idiomatic English*. Oxford: Oxford University Press.

Hornby, A.S. et al. (Eds.). 1974. *Oxford Advanced Learner's Dictionary of Current English.* Oxford: Oxford University Press.

Householder, F.W. and S. Saporta (Eds.). 1962. *Problems in Lexicography.* Bloomington: Indiana University Press.

Ilson R.F. 1988. How Big is Your Dictionary? *Verbatim* 15 (2): 9-10.

Jackson, H. 1998. How Many Words are in YOUR Dictionary? *English Today* 14 (3): 27-28.

Johnson, F. (Ed.). 1939. *A Standard Swahili–English Dictionary*. Nairobi/Dar es Salaam: Oxford University Press.

Kirkeby, Willy A. (Ed.). 2000. *English–Swahili Dictionary*. Dar es Salaam: Kakepela Publishing.

Landau, S.I. 1984. *Dictionaries. The Art and Craft of Lexicography*. New York: The Scribner Press.

Legere, K. (Ed.). 1990. *Wörterbuch Deutsch–Swahili*. Leipzig: VEB Verlag Enzyklopädie.

Lemmens, M. and H. Wekker. 1986. *Grammar in English Learner's Dictionaries.* Tübingen: Niemeyer.

Mathews, M.M. 1933. *A Survey of English Dictionaries*. London: Oxford University Press.

McMillan, James B. 1949. Five College Dictionaries. *College English* 10 (4): 214-221.

Procter, P. et al. (Eds.). 1978. *Longman Dictionary of Contemporary English.* London/Harlow: Longman.

Sacleux, C. (Ed.). 1939. *Dictionnaire Swahili–Français*. Paris: Institut d'Ethnologie.

Sinclair, J.M. (Ed.). 1987. *Looking Up. An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London/Glasgow: Collins ELT.

Sinclair, J.M. 1987. Grammar in the Dictionary. Sinclair, J.M. (Ed.). 1987: 104-115.

TUKI. 1996. *English–Swahili Dictionary*. Dar es Salaam: Tuki.

Willem Botha (Redakteur). *'n Man wat beur. Huldigingsbundel vir Dirk van Schalkwyk.* 2003, xii + 384 pp. ISBN 0-9584401-4-X. Stellenbosch: Buro van die WAT. Prys: R150.

Dr. Dirk van Schalkwyk het sedert sy aanstelling as Hoofredakteur van die *Woordeboek van die Afrikaanse Taal* op 1 Augustus 1991 rigtinggewende leksikografiese leiding op nasionale en internasionale terrein gegee. Met *'n Man wat beur* verleen die Buro van die WAT treffende erkenning aan dr. Van Schalkwyk by sy aftrede. Die nuwe Hoofredakteur van die WAT, dr. Willem Botha, het as redakteur van die huldigingsbundel opgetree.

Die bundel is in twee afdelings verdeel. Die eerste afdeling plaas Dirk van Schalkwyk, die *mens*, onder die loep. Uit die toegeneë en soms humoristiese bydraes van 'n dertigtal kollegas, vriende, familie- en gesinslede, ontwikkel die beeld van Dirk van Schalkwyk, gebore Namibiër en 'n uitsonderlike ménsmens. Karin Benjamin-van Lierop vleg die bydraes in hierdie afdeling knap inmekaar met 'n treffende inleiding wat die woorde *mens* en *menslik* leksikografies ontleed. Hierna verskyn 'n twintigtal gedateerde waarhede uit Van Schalkwyk se persoonlike spreukeskat, gevolg deur sy *curriculum vitae*, 'n toonbeeld van sy lewenslange akademiese arbeid op Afrikaanse taalgebied en veral sy konstruktiewe rol in die uitbouing van die WAT. Prof. Andreas van Wyk, oudrektor van die Universiteit van Stellenbosch, noem Van Schalkwyk in sy voorwoord " 'n man uit een stuk" en dat dit 'n genot was "om met iemand saam te werk wat sy vak ken en in beheer van sy taak is". Die voormalige Hoofredakteur van die WAT, mnr. Niël Hauptfleisch, konstateer onomwonde in sy huldeblyk aan sy opvolger: "Van die vyf hoofredakteurs van die WAT, was dr. D.J. van Schalkwyk die uitmuntendste." Uit die verskillende perspektiewe van kollegas, vriende, familie- en gesinslede kristalliseer daar sekere herhalende hoedanighede ten opsigte van Van Schalkwyk uit, soos *gesinsman, taleman, digter, natuurmens, leier, mensekenner, denker, emosiemens, humoris, storieverteller, diplomaat, strateeg, optimis, lojale vriend,* en dan ook herhalende eienskappe, soos *nederigheid, openhartigheid, deursettingsvermoë, hardwerkendheid, selfgedrewenheid, resultaatgerigtheid, perfeksionisme, goeie menseverhoudings,* en *organisatoriese en bestuursvermoëns.*

Die tweede afdeling bestaan uit artikelbydraes van leksikograwe en taalkundiges van Afrika en Europa, 'n verdere aanduiding van die internasionale aansien wat dr. Van Schalkwyk as akademikus en voorstander van die koöperatiewe leksikografie in Afrika en oorsee geniet. Leksikograwe en taalkundiges uit Afrika en Europa (van so ver as Nederland, België, Duitsland, Noorweë, Tanzanië en Gaboen) het bydraes vir die huldigingsbundel gelewer in 'n verskeidenheid tale, wat 'n onmiskenbare internasionale karakter aan dié bundel gee. Dit sluit akademici en leksikograwe in soos Mariëtta Alberts, Willem Botha, Albina Rogati Chuwa, Jac Conradie, Hans den Besten, Gilles-Maurice de Schryver, James D. Emejulu, Ilse Feinauer, Rufus H. Gouws, Hans Heestermans, Kathryn Kavanagh, Joep Kruijsen, Johan Lubbe, Willy J.R. Martin, Fons

Moerdijk, F.F. Odendal, Anna Nel Otto, Anton Prinsloo, Daniël Jacobus Prins-loo, Michael Schlaefer, Jacques van Keymeulen, P.G.J. van Sterkenburg, Lars S. Vikør en Herbert Ernst Wiegand.

Volgens Willem Botha in "Die renaissance van die *Woordeboek van die Afri-kaanse Taal*" het die WAT in Van Schalkwyk se termyn, sedert Augustus 1991, vanweë sy kennis van beplanningstrategieë, 'n produksieversnelling onder-vind, deurdat dele IX, X, XI en 'n groot deel van deel XII in 'n redelike kort tyd afgehandel is, in vergelyking met die tempo waarteen die eerste agt dele ver-skyn het. (Weens oorbewerking het die voltooiing van die letter K byvoorbeeld meer as 20 jaar in beslag geneem!) Dit het gepaard gegaan met 'n nuwe redak-sionele stelsel as 'n doelbewuste plan van aksie om die beplanning te imple-menteer, rekenaartegnologie te optimaliseer en produksienorme vir redaksie-lede daar te stel. Die uiteinde: 'n moderne woordeboek wat die taalwerklikheid ten beste weerspieël en sodoende die belange van die gebruiker en Afrikaans dien.

Rufus Gouws skryf in sy inleiding "Oor patriotte en ander leksikografiese vernuwers" die geskiedenis van leksikografiese suksesse en mislukkings toe aan deeglike teenoor gebrekkige beplanning. Die uiteindelike sukses-era van die WAT loop parallel aan dié van die tydperk wat Van Schalkwyk aan die WAT verbonde was. Gouws wys ook op die ooreenkomste tussen die benade-ring van S.J. du Toit met sy *Patriot-woordeboek* (1902/1904) en dié van die WAT tydens die era van Van Schalkwyk. Sy beplanningswerk en strategiese ge-moeidheid met die primêre en sekondêre leksikografiese proses het in hierdie tydperk gelei tot koöperatiewe leksikografie, die destigmatisering van Afri-kaans en die vestiging van die WAT as nasionale én internasionale rolmodel.

Fons Moerdijk se bydrae "Metalexicografie en betekenisbeschrijving in het WAT" evalueer die aard en invloed van die metaleksikografiese begeleiding ten opsigte van die inhoudelike betekenisbeskrywing in die WAT. Teen die agtergrond van die kritiek teen vroeëre dele van die WAT, naamlik die vermel-ding van te veel ensiklopediese elemente in die betekenisomskrywing op mikro-strukturele vlak, vergelyk Moerdijk WAT-definisies (met behulp van begin- en eindetikette) met definisies soos voorgestel deur resensente. Hy oordeel onder andere dat die komponensiële analise nie die oplossing vir leksikografiese woordbewerking is nie. Dit word duidelik uit sy bespreking dat die WAT, wat aanvanklik sonder veel metaleksikografiese bemiddeling op die gebied van definiëring gefunksioneer het, in die Haupfleisch- en Van Schalkwyk-eras ten goede verander het.

Die bydrae van D.J. Prinsloo en G.-M. de Schryver "Effektiewe vordering met die *Woordeboek van die Afrikaanse Taal* soos gemeet in terme van 'n multidi-mensionele liniaal" bepaal die vordering van die werk aan die WAT met behulp van 'n wetenskaplike, multidimensionele meetinstrument (liniaal) wat duide-like leemtes in die relatiewe balans ten opsigte van reeds voltooide dele uitwys en terselfdertyd vooruitskouings maak. Daar is wel deurgaans volumegewys vordering met die samestelling van die WAT gemaak, maar *effektiewe vordering*

met die voltooiing van die woordeboekgeheel in gedagte toon egter 'n afplatting op die grafiek vanaf die middel-sestigerjare tot in die vroeë negentigerjare. Daar het dus 'n radikale, positiewe ommekeer in *effektiewe vordering* in die Van Schalkwyk-era (1991–2003) plaasgevind. Die vooruitskouing is dat, indien die WAT daarin kan slaag om dié effektiewe vordering vol te hou, die verwagte voltooiingsdatum van die WAT, in terme van die *liniaal,* op ongeveer 2020 bereken kan word.

In "Vaktaalontwikkeling en -opleiding in Suid-Afrika" wys Mariëtta Alberts op die groot behoefte aan gestandaardiseerde, meertalige, veeltematiese en veelvlakkige vaktaalontwikkeling en -opleiding, beskou in die lig van die voorsiening vir meertaligheid en taalerfenisontwikkeling in die grondwet van die nuwe Suid-Afrika. Sy verskaf ook agtergrond oor werksaamhede van die Terminologiekoördineringafdeling (TKA) van die Nasionale Taaldiens (NTD) as huidige meertalige Nasionale Termbank, en die afdeling se samewerking met onder andere die nuwe PanSAT-strukture. Die TKA gaan binnekort al sy ou, bestaande en lopende vaktaalprojekte op die internet plaas. Daar word ook 'n blik gegee op die langsame historiese ontwikkeling van vaktaal in Suid-Afrika en 'n pleidooi gelewer vir die toekomstige betrokkenheid van tersiêre instansies, staatsdepartemente en die private sektor (die eindgebruikers van gestandaardiseerde terminologie) by die nasionale terminologiekantoor. Die verlangde betrokkenheid sal daartoe lei dat terminologiese inligting versamel, gedokumenteer en uiteindelik gedissemineer word.

A.R. Chuwa fokus in die artikel "Meaning Extension and its Impact on Language Vocabulary: The Case of Kiswahili" op die kompleksiteit van woordeskatuitbreiding in Swahili deur middel van die betekenisverruiming van bestaande woorde. Wanneer daar na taboekonsepte in 'n veranderende wêreld verwys word, beperk die stigma wat aan nuwe betekenisse kleef, soms die gebruik van die woord se ou betekenis. Nuwe betekenisse word deur die hele spektrum van 'n spraakgemeenskap tot die woordeskat toegevoeg, en só 'n proses verlang samewerking van beide spreker en hoorder in 'n bepaalde konteks. Die bespreking word toegelig met 'n tabel van Swahiliwoorde waaraan nuwe betekenisse toegevoeg is en wat soms as homonieme funksioneer as gevolg van die afwesigheid van enige betekenisooreenkoms.

Jac J. Conradie vergelyk in sy artikel "Voltooide deelwoorde in verklarende Afrikaanse woordeboeke" 'n sestal verklarende Afrikaanse woordeboeke om te bepaal hoe Afrikaanse voltooide deelwoorde grammatikaal beskryf word. Hy verwys onder andere na etikette as voorbeeld van hulpmiddelgebruik. Volgens hom kan daar meer aandag gegee word aan die leksikografiese beskrywing van 'n werkwoordstelsel met unieke eienskappe onder Germaanse tale. Ter illustrasie wys hy op WAT IX se vollediger morfologiese inligting in dié verband, teenoor dié in WAT III.

In "Pidgin (and English) Roots of Afrikaans *booi* and *Boesman*" demonstreer Hans den Besten die vroeëre interrelasies tussen voorbeelde van "Cape Dutch" en "Cape Dutch Pidgin" aan die hand van veral onomastiese leenwoor-

de soos *booi* en *Boesman* deur die invloede van 'n verskeidenheid tale (soos Nederlands, Khoekhoen, Engels, e.s.m.) te verreken.

James Emejulu lewer in "Challenges and Promises of a Comprehensive Lexicography in the Developing World: The Case of Gabon" getuienis in sy *Erkennings* van die nuwe horisonne wat vir leksikografie in Gaboen oopgegaan het deur bemiddeling van Van Schalkwyk en die Buro van die WAT. Hieruit vloei sy ondersoek na die omvang van omvattende leksikografie en die beloftes wat dit inhou vir ontwikkelende lande met 'n komplekse, veeltalige en multi-kulturele opset, met spesifieke verwysing na Gaboen. Volgens hom is 'n beleg-ging in opvoeding (kennisbestuur) en die taalindustrie die sleutel tot ontwik-keling. Multifunksionele produkte kan tot meertaligheid lei en toegang tot globale kennis ten beste ontsluit. Strategiese beplanning en bestuur van lek-sikografiese eenhede verskaf 'n stewige fondament.

Ilse Feinauer se bydrae "TB OR NOT TB? The Communicative Success of Translated Medical Texts in South Africa" fokus op die verstaanbaarheid van informatiewe en instruktiewe Engelse brontekspamflette oor tuberkulose (TB), die vertaling daarvan na die doeltaal en die teikengebruikers se uiteindelike begrip (of afwesigheid daarvan) as aanduiding van die uiteindelike suksesvlak van die kommunikasie. Die funksionele model is in hierdie geval verkieslik bo die ekwivalensiemodel, aangesien sodanige vertaling van die bronteks tot 'n wyer omvang teikengebruikers spreek. Sy bespreek mediese pamflette aan die hand van die eksterne struktuur (plasing van belangrike feite) en interne struk-tuur (aanspreekvorm, jargon, passiewe vorme, nominalisasies, sinstruktuur en skakeling).

Hans Heestermans verwys in "*Officieel*: over betekenis en betekenisveran-dering" na die definisie (die semantiese onderskeidende omskrywing van woordbetekenis) as die hoogste, maar moeilikste vorm van kuns in die leksiko-grafie. Hy toon die stelselmatige betekenisontwikkeling in enkele Nederlandse lemmas aan en sien historiese leksikografie as 'n vervullende ervaring vir die nougesette leksikograaf.

In "Co-operative Lexicography — Developing Bilingual Dictionaries" ver-wys Kathryn Kavanagh na Van Schalkwyk as advokaat vir koöperatiewe leksi-kografie en die totstandkoming van leksikografiese eenhede vir die ander tale in Suid-Afrika. Sy fokus veral op algemene faktore van die beplanningsfase en ontwikkelingsfase van tweetalige woordeboeke (met Engels as een van die tale). Die beplanningsfase behels byvoorbeeld insette van teikengebruikers, uit-gewers, spesialiste en terminoloë en die ontwikkelingsfase die saamstel van 'n kernwoordeskat en 'n stylgids, die inagneming van onderskeie teikengebrui-kers se taal- en kultuurverskille en ook uitspraakleiding. Die hantering van die lemmas *figure* en *live* word vervolgens vergelyk in vier bestaande Engels–Afri-kataalwoordeboeke.

Joep Kruijsen lewer in sy huldeblyk "Limburgse bloemlezing voor Dirk van Schalkwyk" verslag oor die drie regionale woordeboeke (van die Brabant-se, Limburgse en Vlaamse dialekte) se aflewering oor blomme. Hy gee 'n oorsig

van die verskillende benamings vir die sering, die sleutelblom en die madelie-
fie, telkens toegelig met kaarte waarop die verspreiding van die benamings
aangedui word.

Johan Lubbe neem die Afrikaanse en Nederlandse byvoeglike naamwoord
aardig se positiewe en negatiewe gekonnoteerde betekenisse onder die loep in
sy artikel "*Aardig* in Afrikaans en Nederlands — 'n diachroniese beskrywing
van die betekenisverskille". In dié proses bring hy hulde aan Van Schalkwyk se
aandeel in die afhandeling van die *Etimologiewoordeboek van Afrikaans* (2003).
Die bespreking begin met 'n historiese oorsig van die verskillende teoretiese
benaderings met betrekking tot die beskrywing van betekenisverandering, hoe
dit vervolgens in Afrikaans neerslag gevind het, hoe *aardig* in Nederlandse en
Afrikaanse woordeboekinskrywing (wat betekenis betref) ooreenkom of verskil
en hoe dit verklaar kan word in die lig van gegewe teoretiese benaderings.

Willy J.R. Martin verwys in sy artikel "Lexicography, Lexicology, Linking
and the Hub-and-Spoke Model" na die onderskeid wat Nuccorini maak tussen
leksikologie en leksikografie en vermeld 'n perspektiefskuif tussen die twee
begrippe, met 'n gepaardgaande klemverskuiwing na leksikografiese infra-
struktuur by die konstruksie van tweetalige woordeboeke. Die rol wat data
(gekoppelde, omkeerbare databasisse), datamanipulasie (omkeringsmoontlik-
hede) en datamodelle (breër raamwerke soos die naaf-en-speek-model) speel,
word ook belig. Hy kontrasteer die ou (vertalende) en nuwe (gekoppelde/om-
keerbare) benaderings tot tweetalige woordeboeke aan die hand van die lemma
glimmer. Twee leksikale databasisse word voorveronderstel om as kontrole te
dien by die gekoppelde/omkeerbare benadering. Vormeenhede en leksikale
eenhede word onderskeidelik deur koppeling- en vertalingskakels verbind. Dit
impliseer dat leksikale eenhede indikatore soos resumés, semantiese tipes en
definisies van die vormeenhede moet bevat. Hy verwys ook kortliks aan die
hand van die lemma *river* na die rekenaarprogram OMBI as instrument om
data te manipuleer. Ten slotte beklemtoon hy die meriete en potensiaal van die
naaf-en-speek-model wat inter- en intralinguale verhoudings ontgin ('n interes-
sante model vir 'n meertalige land soos Suid-Afrika) en die gepaardgaande kruis-
bestuiwing tussen leksikologie en leksikografie wat só 'n benadering teweeg-
bring.

In sy artikel "Bevry ons tog leksikografies van Nederlands" bespreek F.F.
Odendal die verhouding tussen Afrikaans en Nederlands en Van Schalkwyk se
bydrae om Afrikaans se onafhanklikheid en eiewaarde (veral in die latere
WAT-dele) te versterk. Hy vermeld dat verklarende en vertalende Afrikaanse
woordeboeke sterk op Nederlandse woordeboeke steun en in dié proses onno-
dige Nederlandse bagasie saamdra, omdat die woorde nie meer in hedendaag-
se Afrikaans gebruik word nie. Hy lig sy bespreking toe met tientalle lemmas
wat, volgens hom, "in aanmerking kom vir skrapping as onafrikaans" (p. 291).

Aan die hand van verskillende persepsies van leksikograwe bespreek
Anna Nel Otto die aard en waarde van illustrasies in woordeboeke in haar arti-
kel "Woordeboekillustrasies, met spesiale verwysing na illustrasies in aanleer-

derwoordeboeke". Daar word verwys na illustrasies van 'n enkele selfstandige naamwoord, verskeie naamwoorde (kollektiewe illustrasie) en ander woordsoorte (byvoorbeeld die werkwoord *damage*). Praktiese aspekte met betrekking tot illustrasies word ook in die artikel verreken. Aanvullende illustrasies en definisies sáám is onontbeerlik in aanleerderwoordeboeke en behoort sistematies opgeneem te word.

In Anton Prinsloo se artikel "*Huidiglik* — 'n sondebok (bet.1)" bespreek hy 'n aantal woorde in Afrikaans wat as sondebokke beskou kan word, byvoorbeeld *agte(r)losig*, *aanwen(d)sel*, en veral *huidiglik*. Gegewens uit die Naspersargief oor die gebruik van dié woorde word ook verskaf. Hy vergelyk die hantering van dié lemmas in verskillende woordeboeke deur die lemma se opname, nie-opname of opname met voorbehoud of afwysing. Uit sy bespreking word dit duidelik dat woorde soms as taalfoute ontstaan — sommige waarvan mettertyd weens gebruiksfrekwensie as standaardvorm aanvaar word (mits die taaldebat daaroor nie te hewig is nie, of die woord nie te veel sinonieme as ander gebruiksopsies het nie). Dit laat die vraag ontstaan of sodanige woorde nie dalk slagoffers van "dubbele norme" in die Afrikaanse leksikografiepraktyk geword het nie.

In sy artikel "Wortgeschichte in panchronischen Wörterbüchern" wys Michael Schlaefer op die kritiek wat gedurende die laaste jare gelewer is op historiese woordeboeke wat oor verskillende tydperke in die geskiedenis van Duits handel. Dit gaan onder andere oor die onvoldoende verklaring van leksikografiese konsepte, probleme met die keuse van lemmas en die beskrywing van die semasiologiese ontwikkeling. Hierdie kritiek word aan die hand van die Duitse lekseem *Feder* (veer) bespreek. Schlaefer kom tot die gevolgtrekking dat die Duitse leksikografie, anders as dié van ander Europese tale, nie net kan hersien nie. Nuwe korpusse moet saamgestel en nuwe etimologiese konsepte moet ontwikkel word. Dataverwerking sal 'n belangrike rol in die ontginning van linguistiese inligting speel.

Jacques van Keymeulen dink in sy artikel "De introductie van het Nederlands in Vlaanderen: taalstrijd van de middenklasse" met genoeë terug aan sy kontak met Van Schalkwyk. Hy verwys in sy inleiding na gesprekke in 2002 met leksikograwe in Afrika met betrekking tot die groeiende kulturele selfbewussyn van eie (inheemse) tale en woordeboeke as uiteindelike taalmonumente. Hy skets verbasende ooreenkomste tussen die talesituasie in Vlaandere en dié in Afrika en verwys ten slotte na Nederlands as 'n plurisentriese taal. Volgens hom gaan aspekte soos taalemansipasie, kulturele emansipasie en ekonomiese vooruitgang, soos in Vlaandere, hand aan hand en lei dit tot 'n groeiende selfbewussyn (soos emansipatoriese taalbewegings ook nou in Afrika begin ontstaan met die skryf van woordeboeke vir inheemse tale).

P.G.J. van Sterkenburg bestudeer etikette soos *scherts* en *schertsend* in sy artikel "Humor in woordenboeken: het label 'schertsend' " ten einde vas te stel hoe 'n leksikograaf, vanuit metaleksikografiese gesigspunt, sodanige etikette wetenskaplik en funksioneel vir die woordeboekgebruiker kan toepas. Dit word

stelselmatig duidelik dat 'n grammatika van *skerts* nog nie bestaan nie. Eers wanneer sodanige grammatika beskikbaar is, sal die leksikograaf in staat wees om woorde of woordgroepe se skertsende gebruik aan te toon en, indien gewens, van 'n etiket te voorsien. Enkele aspekte wat vir 'n moontlike voorgestelde strukturele oplossing uit die terreinverkenning geïdentifiseer is, is dat gesproke taal in die materiaalversameling meer verreken behoort te word, dat die kontekste, die emosionele funksie (sosiale en interaktiewe aspekte) en die taalniveau van *schertsen* aangedui en beskryf moet word.

Lars S. Vikør verwys in sy artikel "Ethnonyms in Norwegian — Between Neutrality and Abusiveness" na die sosio-politieke aspekte (waaronder etniese terme) waarmee rekening gehou moet word in bepaalde gebruikskontekste, hul gepaardgaande konnotasies en die sosiale bewussyn wat hulle omring. In sy bespreking verwys hy na die ooreenkomste tussen Noorweë en Suid-Afrika wat betref onderdrukte etniese groepe en hoe emansipasie in taalterme tot uiting kom. Die neutrale aanslag tipies van Noorweegse woordeboeke word ook vermeld. Hy vergelyk verder die wyse waarop woorde soos *Negro, Nigger, Neger*, e.s.m. leksikografies verskillend hanteer word in woordeboeke. Sosiopolitieke faktore verander mettertyd die betrokke woorde se konnotasies van neutraal na negatief, wat dan die nodigheid van 'n etiket beklemtoon.

H.E. Wiegand gee 'n algemene beskrywing van die alfabetiese vakwoordeboek *Wörterbuch zur Lexikographie und Wörterbuchforschung* (WLWF). Hy konsentreer veral op die vorm, maar ook op die inhoud en die meertalige aanbieding daarvan. Op die vlak van die buitetekste en die vakleksikografiese definisies is dit tweetalig: Duits en Engels; op die vlak van die lemmata is dit veeltalig deurdat elf tale verteenwoordig is. Op 'n algemene karakteristiek van die woordeboek volg 'n meer spesifieke karakteristiek ten opsigte van die teikengebruikers en bronneomvang. Na 'n bespreking van die belangrikste tekskomponente word 'n voorbeeldartikel gegee om die teksposisies en hul funksies te illustreer. Aandag word ook geskenk aan die mikrostruktuur en gegewens betreffende die mediostruktuur.

Oorsigtelik beskou, toon hierdie huldeblyk hoe dr. Dirk van Schalkwyk, op professionele vlak, met sy bestuurs- en beplanningsvaardighede, waarvoor hy inderdaad internasionale erkenning verwerf het, die transformasie van die Buro van die WAT en die modernisering van die *Woordeboek van die Afrikaanse Taal* volgens metaleksikografiese beginsels bedryf het. Met sy publikasies oor die beplanning en bestuur van leksikografiese eenhede (soos die Buro van die WAT) het hy leksikografiese horisonne verbreed (sien Emejulu, p. 210). Verder het hy vorendag gekom met werkbare langtermynbeplanning in 'n tyd toe die staatsubsidie gekrimp het, die WAT-wet herroep is en die sosiale, kulturele en politieke speelveld verander het.

Buiten werk aan die WAT het dr. Van Schalkwyk die Buro van die WAT ook gelei om ander inisiatiewe te onderneem, byvoorbeeld 'n reeks opleidingskursusse in die leksikografie aan nasionale én internasionale studente en ook personeel van die nasionale leksikografie-eenhede, die totstandkoming van die

geakkrediteerde internasionale vaktydskrif *Lexikos* en die tweemaandelikse Woordpret, wat 'n podium vir Afrikaanse woord- en musiekkunstenaars bied. Die totstandkoming van woordeboekeenhede vir die amptelike landstale deur sy betrokkenheid by PanSAT, sy begeleiding met die beplanning of herbeplanning van woordeboekprojekte in Suid-Afrika en Europa, en, vanweë sy ideaal van koöperatiewe leksikografie, die sluit van samewerkingsooreenkomste met projekte en universiteite in Afrika en Europa, het internasionaal van die Buro van die WAT 'n voorbeeld gemaak. Die WAT is deur sy inisiatief tans ook gerekenariseerd beskikbaar (die *Elektroniese WAT*) en buite-outeurs het verder as medewerkers met die Buro saamgewerk om publikasies soos die *Etimologiewoordeboek van Afrikaans* (EWA), *Afrikaanse idiome en ander vaste uitdrukkings* en die *Woordkeusegids: 'n kerntesourus van Afrikaans* die lig te laat sien.

Hierdie boek is 'n móét vir elkeen wat belangstel in die moeisame leksikografiese pad van die WAT sedert die 1920's. Onder dr. Van Schalkwyk se leiding het die vaartversnelling en karakterverandering van die WAT as omvattende woordeboek van Afrikaans, gelei tot die WAT se hernieude geloofwaardigheid, relevansie en vermoë om die hele Afrikaanse taalwerklikheid getrou te weerspieël. Ter afsluiting van die bespreking van hierdie huldigingsbundel, kan 'n persoonlike spreuk (1992) van dr. Van Schalkwyk aangehaal word: " 'n Mens moet nie jou beroep so beoefen dat jy eers na aftrede lus raak vir werk nie." Hy kan nou egter met voldoening terugkyk, want hy was 'n man wat sy personeel deur sy navolgenswaardige voorbeeld geïnspireer het om onvermoeid te bly beur in belang van Afrikaans ...

Corlietha Swart
Afrikaanse Taalkunde
Departement Afrikaans en Nederlands, Duits en Frans
Universiteit van die Vrystaat
Bloemfontein
Republiek van Suid-Afrika
(*SwartC.blms@mail.uovs.ac.za*)

J.D.U. Geldenhuys en E.E. Viljoen-Smook. *New Business Dictionary/ Nuwe Sakewoordeboek*. Fourth, revised and expanded edition/Vierde, hersiene en uitgebreide uitgawe. 2003, xxi + 315 pp. ISBN 1 86890 038 X. Cape Town/Kaapstad: Pharos Dictionaries/Pharos Woordeboeke. Prys: R140.

Die *New Business Dictionary/Nuwe Sakewoordeboek* van Geldenhuys en Viljoen-Smook (hierna G en VS) kan tereg bestempel word nie net as 'n *nuwe sakewoordeboek* nie, maar ook as 'n *nuwesakewoordeboek*. Om dit te verklaar, moet die sakeagtergrond in ag geneem word waarteen hierdie nuutste aanvulling tot die Afrikaanse vakwoordeboekkorpus saamgestel is.

Die eerste probleem wat ondervind word, is veranderinge in die beursbedeling gedurende die laaste dekade. Daar bestaan verskille tussen Britse en Amerikaanse saketerme op dié gebied — die beginsel van een begrip, een term geld dus nie aan beide kante van die Atlantiese Oseaan nie. Wat maak 'n mens dus in Afrikaans? Tot redelik onlangs is Suid-Afrika beskou as 'n deel van die dampkring van Groot-Brittanje: die Britse finansiële terminologie is as bron aanvaar. Hierdie toedrag van sake het die laaste tyd aansienlik verander en selfs die Engelse aanvaar nou die losser jargon van die Amerikaners wanneer dit by finansiële dokumentasie kom. As voorbeeld: Wat die Britte *shares* genoem het, het die Amerikaners as *stocks* geken. *Shares and bonds* in Britse Engels het *stocks and bonds* in die ander land geword.

Hierdie oneweredigheid is in Suid-Afrika weerspieël in die Afrikaanse terminologie vir die *Johannesburg Stock Exchange* (soos dit vroeër bekend gestaan het), naamlik die *Johannesburgse Effektebeurs*. *Aandeel* is die Afrikaanse ekwivalent vir *share*, en *effek* vir *stock* — dit beteken dus dat die Britse patroon in Afrikaans gevolg is. In die vroeë negentigerjare het die *Bond Exchange of South Africa* sake begin doen. Met dié naam het die Amerikaanse invloed geseëvier: *stocks* is gebruik vir die Britse *shares* (*aandele*), en *bonds* vir *stocks* (*effekte*). In Afrikaans het die *Bond Exchange* die *Effektebeurs* geword, en die *Stock Exchange* die *Aandelebeurs.*

Nog is het einde van die verwildering niet: Die *South African Futures Exchange* het ook in die negentigerjare sy vleuels begin sprei en handel gedryf hoofsaaklik in termynkontrakte, maar ook in sekere opsies. Die Johannesburgse Aandelebeurs het intussen ook in aandeleopsies begin handel dryf — en dié probleem is vir die Engelse uitgestryk toe die Johannesburgse Effektebeurs en die *South African Futures Exchange* vroeg in dié eeu saamgesmelt het onder die nuwe naam: *JSE Securities Exchange South Africa.* Maar: Afrikaans sit toe weer in 'n kluif deurdat die ingeburgerde vertaling van *securities* lank reeds *effekte* was. Die Effektebeurs was nie deel van die samesmelting nie — nog 'n nuwe naam moes gekry word. Die *JSE Securities Exchange South Africa* is toe in Afrikaans gedoop as die *JSE Sekuriteitebeurs Suid-Afrika.* Die onderliggende gedagte was dat die term *sekuriteite* geskik is om *aandele, effekte, ekwiteite, termynkontrakte, vooruitkontrakte, opsiekontrakte, ruilopsiekontrakte* en enkele ander elemente te omvat.

Die *New Business Dictionary/Nuwe Sakewoordeboek* begelei die vakman uit hierdie duister termdoolhof.

Wat met die eerste oogopslag opval, is dat Nederlandse spore in ons vroeëre (ook algemene) woordeboeke wat finansterme opgeneem het, in hierdie uitgawe grootliks uitgewis is (selfs miskien geheel en al — dit is nie moontlik om binne 'n redelike tyd al die terme in die *New Business Dictionary/Nuwe Sakewoordeboek* deur te kyk nie). Om enkele voorbeelde te noem: *annual salary* staan vandag nog in sommige Afrikaanse woordeboeke bekend as *jaarwedde* (*wedde* in Nederlands is "bezoldiging, vaste beloning van een hoofdarbeider" — maar beslis nie in Afrikaans nie; die *Woordeboek van die Afrikaanse Taal* etiketteer dit reeds amper 'n halfeeu gelede as *weinig gebruiklik*); *private sector dishoarding* (*privaatsektorontpotting*) het baie lank as die (bykans eksotiese!) *onthamstering van die partikuliere* bekend gestaan; *proxy form* as *volmagformulier*; *drive a hard bargain* as *skerp marsjandeer, tot die uiterste beknibbel*; *denominations of coin* as *muntkoepures*; *floating rates* (*swewende koers*) as *vervarende tariewe*; *free of commission* (*vry van kommissie*) as *franko kommissie* en dergelike meer.

Die samestellers se opname (in vroeëre uitgawes van hulle woordeboek) van verskeie Engelse beursterme wat daagliks in die nuus voorkom, het nou ook al neerslag gevind in ander Afrikaanse publikasies. Van die bekendstes is die (in Engels) ondeursigtige terme soos *bear* en *bull*. Die Afrikaanse ekwivalente weerspieël in hoë mate die definisies daarvan, byvoorbeeld:

bear: 'n Pessimis, iemand wat 'n sekuriteit verkoop het in die hoop om dit later teen 'n laer prys weer aan te koop (G en VS: *daalspekulant*);

bull: Beleggers wat aandele koop met die verwagting dat die prys later sal styg (G en VS: *stygspekulant*).

Op soortgelyke wyse is die Afrikaanse vertalings van die Engelse *jobber* (*kitsspekulant*), *piker* (*kleinspekulant*) en *stag* (*kortspekulant*) ook deursigtiger.

Twee interessante terme is *contango* en *backwardation* — wat by eerste aanblik lyk na skeppings wat 'n eendagsvliegdood moes gesterf het, maar wat stewig op dié gebied ingeburger is. Volgens die Engelse woordeboekdefinisies is een van *contango* se betekenisse "die koste van die uitstel van 'n vereffening van 'n sekuriteit wat gewoonlik by die prys van die sekuriteit gevoeg word". G en VS se weergawe is *prolongasie.* Een betekenis van *backwardation* is "die persentasie koste wat 'n verkoper van aandele moet betaal om aflewering te vertraag". In hierdie geval is *deport* (soms ook *effektehuur* genoem) as vertaalekwivalent ingespan. Hoewel dié term nie vir die leek deursigtig is nie, het dit goeie voorsate in Nederlands (en, terloops, ook in Frans).

Een van die besware wat dikwels deur taalvosse geopper word, is die sogenaamde "anglisistiese" gebruik van *depresiasie* in plaas van die "suiwer Afrikaanse" *waardevermindering.* In *New Business Dictionary/Nuwe Sakewoordeboek* ontrafel G en VS die subtiele betekenisverskille tussen *depresiasie, devaluasie,*

waardeafname, waardedaling en *waardevermindering* wat nodig is om eenduidig-heid in die terme van die finansiële sektor te verkry. Om net bolangs hierna te kyk: Onder die ouditeure word *waardedaling* en *verswakking* vir *impairment* ge-bruik, maar G en VS het dit met *verswakking* vertaal (dit is iets wat nou met die nuwe rekeningkundige standpunt, RE133, baie aandag geniet en standaardi-sasie nodig het). As bates *impair* word, het die waarde nie noodwendig *gedaal* nie, maar kan die bates byvoorbeeld as gevolg van geldeenheidbewegings *verswak* wees. Dit kom basies daarop neer dat *impairment* nie deel van die *de-cline in value*-konstellasie in Engels uitmaak nie, terwyl *waardedaling* wel deel van die *waardedaling*-konstellasie uitmaak en *verswakking* eerder in dié verband gebruik moet word.

Verder: die aanvaarding van die woord *franchise* in Afrikaans het ook in puristiese geledere opslae gemaak, nieteenstaande die verduideliking dat *fran-chise* as superordinaat talle kohiponieme soos *konsessie, vergunning, handelsreg* en etlike ander omvat. In *New Business Dictionary / Nuwe Sakewoordeboek* word die gebruik van *franchise* weer bevestig, wat die voorskriftelike aard van dié woordeboek weerspieël. Dit bly eienaardig dat *franchise* die wraakgevoelens van talle gaande maak, maar 'n ander ontlening aan Engels, *dumping*, ongesiens en onaangekondig op die vakgebied aanvaar is.

Besonder interessant vanuit 'n taalkundeoogpunt is woorde uit Omgangs-engels wat mettertyd as terme op finansiële gebied aanvaar is, soos *catapult* (in Afrikaans *vals uitbraak*); *cough up* (betaal ▬ en in Afrikaanse sleng *hoes*); *dawn raider* (*vinnige stroper* ▬ op die effektebeurs); *flipover pill* (*heilsame pil* ▬ by oor-name); *high-street sale* (*formele verkoop*); *swoption* (*ruilopsie*), en die beeldryke *dead-cat bounce* (*tydelike oplewing* op die sekuriteitebeurs waar die tendenslyn skerp na onder gedaal het en dan skielik weer 'n effense styging vertoon, soos 'n dooie kat wat van 'n groot hoogte val en dan so 'n klein bietjie opwip).

'n Leksikografiese vernuwing is ook die opname van geykte finansiële uitdrukkings (meerwoordige lemmas) onder die aanvangsletter van die eerste woord daarvan, en nie onder wat soms as die *hoofwoord* bestempel word nie, en waarmee dikwels die eerste selfstandige naamwoord in die uitdrukking bedoel word. In algemene woordeboeke sal byvoorbeeld 'n meerwoordige vakuit-drukking soos *by private treaty* as 'n inskrywing onder *treaty* behandel word; *casual work, casual worker* onder *work* en *worker* onderskeidelik; *galloping inflation* onder *inflation*; *lift a leg* (!) met die betekenis "'n been van 'n transaksie volkome afwikkel" onder *leg* ensovoorts. Ook in die deel met Afrikaans as brontaal word dié beginsel gevolg, byvoorbeeld die Afrikaanse weergawe van *lift a leg* hierbo is opgeneem as die meerwoordige lemma *been van 'n transaksie volkome afwikkel, 'n*. (By hierdie spesifieke inskrywing is dit enigsins eienaardig dat die onbe-paalde lidwoord agteraan die inskrywing gevoeg is terwyl dit 'n vaste deel daarvan uitmaak; in die geval van uitdrukkings met die bepaalde lidwoord is dié beginsel nie gevolg nie, byvoorbeeld *die allernodigste; die Effektebeurs, Lon-den, die kabel* ("die sterling-dollar-lokokoers in Londen") is by *die* opgeneem.) Voorbeelde soos *bo prys* by *bo* en *drukkende skuld* by *drukkende* opgeneem kan by

die vleet vermenigvuldig word. Hierdie wyse van opname vergemaklik soektogte en bespoedig naslaanwerk, en moet as 'n wins beskou word.

Die bynaam vir die *Financial Times-effektebeursindeks* (*Footsie*) is opgeneem; maar drie ander ewe bekende terme wat daagliks in Afrikaanse ekonomiese nuusbulletins gehoor word, ontbreek. Hulle is die *CAC40* (die *Compagnie des Agents de Change* — die effektebeurs van Parys), die *DAX* (*Deutsche Aktienindex* van Frankfurt) en die *Nasdaq* (die *National Association of Securities Dealers Automated Quote System* van die VSA). Miskien kan die samestellers in 'n volgende oplaag daaraan aandag skenk.

Dié woordeboek is 'n waardevolle aanwins tot die literatuur op die gebied van finansies, 'n vakgebied waaroor daagliks in alle media, en ook in Afrikaans, wyd berig word. Dit maak dit waarskynlik een van die mins esoteriese vakgebiede en een waar 'n Afrikaanse woordeboek soos dié onmisbaar is. Verder bewys dié boek weer eens dat Afrikaans volwaardig vaktaal is. Die *New Business Dictionary/Nuwe Sakewoordeboek* dui op grondige navorsing oor die onderwerp en toon ook tekens van verbeeldingryke skepping by die vertaling van nuwe Engelse terme wat daagliks die land deur middel van die wêreldmedia binnespoel. As 'n steekproef is 'n professor in die ekonomie genader en gevra om sekere spesifieke nuwe terme te noem waarvan hy bewus is — almal is in dié woordeboek opgeneem.

A.F. Prinsloo
Posbus 44924
Linden 2104
Republiek van Suid-Afrika
(prinsa@worldonline.co.za)

Manfred Görlach. *English Words Abroad*. 2003, xii + 189 pp. ISBN 90 272 2331 9 / ISBN 1 58811 435 X. Amsterdam/Philadelphia: John Benjamins. Price: €70, US$70 (Hb.).

Since the Second World War, the infiltration of English into the Romance languages has remarkably increased, notably diffused by various media such as newspapers, radio and television broadcasts, popular music, and the internet. This comprehensive, thorough and authoritative work, whose data has been collected from these different sources, strives to quantify the spread of Anglicisms within individual foreign languages.

In French, for instance, Anglicisms are used daily as trendy slang which lends an exotic feel to the language; they also refer to concepts otherwise lacking in the foreign tongue. The ease of and preference for using Anglicisms lie in their condensing of otherwise long paraphrases, such as the use of *start-up* which translates in French as "une entreprise en pleine expansion".

Most Anglicisms are fairly recent acquisitions, dating from the 19th and 20th centuries. Certain European languages like Dutch, German and French have been frequently scrutinized, but little is known about the linguistic influence of English on languages such as Czech, Croatian and Polish.

Since 1988, Görlach has been working on the *Dictionary of European Anglicisms* project (2001), a lexicographical study on the diffusion of English in sixteen languages from different language families: Romance languages (French, Spanish, Italian, Romanian), but also German, Dutch, Norwegian, Icelandic (Germanic), Russian, Polish, Croatian, Bulgarian (Slavonic), Finnish, Hungarian (Finno-Ugric) and Albanian and Modern Greek. Two sequels, *An Annotated Bibliography of European Anglicisms* (2002), followed by *English in Europe* (2002), serve as companions to the dictionary. In *English Words Abroad*, Görlach explores the methods used for the compilation of this dictionary showing the present-day impact of Anglicisms as well as the degree of integration of loanwords and the transformation, whether phonological, graphemic, morphological or semantic, they might undergo in the receptor language. He also gauges their chance of survival in the short and long run: some are adopted in the vocabulary of younger people, while others have become dated and are only used by the older generation, the calques of which sometimes gain a certain popularity.

A similar investigation has already been conducted in the 1970s by the Croatian precursor Rudolf Filipović who elaborated a large corpus of Anglicisms in twenty selected languages in order to establish the growing contact of these languages with English. Görlach's *English Words Abroad* is a collection of conference papers and articles showing his own work's progress. The author discusses data analysis, accounts for linguistic prescriptive attitudes to language and presents the methodological and terminological problems faced when compiling a database of Anglicisms and building a multilingual dictionary. The enterprise is ambitious and the results of almost twenty years of research cautions us against potential difficulties and obstacles.

Görlach makes a distinction between *loanwords* and *aliens*. The former relate to lexical items that are fully integrated into a language, the latter refer to words which, due to phonological and morphological differences with the receptor language or other factors such as stigmatization or lack of mediatisation, will not be successfully integrated. To be included in the dictionary, Görlach targeted words that "were recognizably English in form" (Görlach 2003: 44) and frequent enough to be considered part of the national language and cultural history of a country. Calques are often competing with loanwords and will be considered separately.

Although some of the material of this book seems to overlap, this study is far from being repetitive and it can actually be seen how the work has progressed throughout the years. Chapter 1 lays the foundations of the project, gives definitions and also presents a survey of current research. Chapter 2 focuses on the impact of English and on data collection. Chapter 3 is an update of the work in progress. Chapter 4 deals with etymological issues, chapter 5 with marginal lexis such as foreignisms (words like *pub* and *shilling* which strictly refer to Anglo-Saxon culture), technical words and archaisms. Chapter 6 focuses on the graphemic, phonological, stylistic and social development of Anglicisms. The integration of Anglicisms may imply some phonological, stylistic and social changes. Preliminary statistics made on 876 lexical items show that in decreasing order German, Dutch and Norwegian were the languages which most attested the investigated Anglicisms, followed by Italian, French and Finnish. Albanian, on the contrary, was the language which proved least susceptible to these Anglicisms.

The problem of distinguishing between pseudo-Anglicisms and true English coinages makes this study all the more difficult. As languages are constantly evolving, the vocabulary is at times fleeting and it can hardly be foreseen whether such items will become obsolete or will survive in the linguistic repertoire of native speakers. Chapter 7 investigates the integration of Anglicisms and the ensuing grammatical problems such as gender attribution, pluralization, and compounding as well as marginal morphemes or phonemes which do not exist in a receptor language. This is the case of *-ing* and *-er* derivatives which become integrated into Croatian, Polish, Bulgarian and French for instance. Chapter 8 analyses semantic changes (the extension or reduction of meaning) which affect loanwords in the receptor language. The borrowed items can acquire particular connotations which convey a derogative, metaphoric or colloquial meaning. Görlach looks at varieties of calques and semicalques in chapter 9 and shows how the translation of foreign lexical items is part of a long-lasting tradition by purists and language legislation (such as the 1994s *loi Toubon* in France) to preserve languages from foreign intrusion. He also attempts to establish the effectiveness of such legislation in safeguarding the purity of the receptor language. The following chapter is dedicated to the issue of acceptability and analyses the gradual process whereby a word of foreign origin becomes a component of the receptor language. Some words are

fully integrated, while the usage of others is restricted. Chapter 11 summarizes a large number of recent research conducted on Anglicisms and chapter 12 presents a contrastive analysis. Görlach encourages researchers to work on Gallicisms, Germanisms and neoclassical lexis in order to complement this overall study of Anglicisms, and gives suggestions on how to elaborate dictionaries recording such items. Using two pilot studies conducted on selected representative items in both French and German to illustrate his results, Görlach shows the dominance of French in the languages of educated Europe, mostly in German, Dutch and English. The function of French before 1900 can be paralleled with the current role of English. He equally considers the expansion of German in Eastern countries and in Scandinavia. In the last chapter, statistical data obtained from a CD-ROM is opening new perspectives for research.

Clearly illustrated with grids and graphs and substantiated with rich footnotes, this fascinating, multicultural study provides both tools for research and invaluable advice in the elicitation and reading of data. It will undoubtedly appeal to linguists and students alike, as well as to those simply interested in languages.

References

Filipović, R. 1983. An Etymological Dictionary of Anglicisms in European Languages, Theory, Methods and Models of Contact Linguistics. Nelde, P.H. (Ed.). 1983. *Theorie, Methoden und Modelle der Kontaktlinguistik/Theory, Methods and Models of Contact Linguistics/Théorie, Méthodes et Modèles de la Linguistique de Contact*: 59-68. Plurilingua 2. Bonn: Ferd. Dümmler.

Görlach, M. 2001. *A Dictionary of European Anglicisms: A Usage Dictionary of Anglicisms in Sixteen European Languages*. Oxford: Oxford University Press.

Görlach, M. 2002a. *An Annotated Bibliography of European Anglicisms*. Oxford: Oxford University Press.

Görlach, M. 2002b. *English in Europe*. Oxford: Oxford University Press.

Michaël Abecassis
Faculty of Medieval and Modern Languages
University of Oxford
Oxford
United Kingdom
(michael.abecassis@modern-languages.oxford.ac.uk)

Kyo Kageura. *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth.* 2002, viii + 323 pp. ISBN 90 272 2328 9 (Eur.) / 1 58811 314 0 (US). Amsterdam/Philadelphia: John Benjamins. Price: €92/US$92 (Hb.).

This monograph was issued as Volume 5 in the series Terminology and Lexicography Research and Practice, edited by Marie-Claude L'Homme and Ulrich Heid and assisted by Juan C. Sager as Consulting Editor. The Japanese author, Kyo Kageura, does research work at the National Institute of Informatics in Tokyo. His theoretical background and methodological experience in term formation patterns and processes were influenced by intensive cooperation with European terminologists since the 1990s, notably the British linguists Juan C. Sager (Manchester, UMIST) and Yorick Wilks (Sheffield), and R. Harald Baayan in the Netherlands.

The book is devoted to aspects of the description of the dynamics of term formation in Japanese, with focus on intra-term relations of nouns in the domain of documentation, where the author, in his own words (p. 4), is the subject specialist with an "in-depth knowledge". His competence and experience in mathematical statistics are demonstrated in Chapter 7 on "Quantitative Analysis of the Dynamics of Terminology".

The data for the corpus analysis were drawn from the Japanese version (1984) of the joint volume *Terminology of Documentation,* edited by G. Wersig and U. Neveling (1976) in Paris, Unesco. The author gives sensible reasons why he — in contrast to recent terminological studies — decided in favour of a reference book on documentation instead of actual written sources of various text types, particularly articles (cf. p. 4, footnote). Since his aim is to elaborate typical conceptual intra-term relations and also tendencies of terminological growth, random examples of terms from documentation texts were thought insufficient for a representative set of data. Thus, his corpus amounts to 1 228 Japanese lexical items which have a terminological status. They are exclusively nouns consisting of a head or "nucleus" and a modifier element, a "determinant". These nouns constitute the core of the Japanese terminology of documentation. Other word classes or parts of speech such as verbs, adjectives, and adverbs which may equally have a terminological status (J.C. Sager, 1990, speaks of "Entities", "Activities", "Qualities" and "Relations", quoted p. 66f.) have been deliberately discarded.

As his guiding principle the author emphasises that the *whole terminology*, not individual terms, of a target domain should be studied in depth, with the necessary "granularity of description" (p. 32).

It must be taken into account that in the Japanese language constituent elements or morphemes of a lexical item are not indicated "by boundaries between linguistic units such as spaces or hyphens" (p. 4). So the author had first to identify the minimal meaning-bearing linguistic elements which correspond to a morpheme in term formation.

The majority of the terms under analysis consist of two morphemes (50,6%); three morpheme structures amount to 23,0%, and one morpheme terms to 20,0% (p. 5). For all types and models of intra-term conceptual relations the English translation is given. A contrastive analysis of the morphemic structure of documentation terms was clearly not the author's intention.

The book is composed of four parts and divided into ten chapters each of which is closely linked by cross-referencing textual devices which support reader-friendly coherence, such as topic sentences, advance organisers, intermediate summaries and the repetition of key arguments. In this respect, the author follows the tradition of British and American academic writing.

In Part I Kageura provides the "Theoretical Background" of his procedure, gives working definitions of "term", "terminology", "lexical unit", "vocabulary", "concept", "conceptual structure", and "characteristic". The designations "domain" and "subject field", however, appear to be accompanied by a circular definition (p. 10) as both are described as "an area of knowledge", even if the latter is characterised by "the purpose of grouping into conventional categories the concepts considered as belonging together" (p. 10). Unfortunately, the author fails to give a definition or at least an explanation of the domain *documentation* which has yielded the terminology corpus under analysis. Instead, he drops an occasional remark (p. 264) that it differs from "library science". On the other hand, the examples in the term lists quoted in the appendices clearly show numerous overlaps between the two fields.

In a subchapter the author critically assesses "the traditional theory of terminology", in which he shares, to a certain degree, the attitude of R. Temmerman (2000). He deplores that the granularity of conceptual systems described so far in traditional terminology "does not exceed the granularity of semantic or conceptual systems established in non-terminology-related studies" (p. 23). It is therefore his main concern to elaborate a far more subtle and intricate system of conceptual elements which are based on a type-token principle and can be minutely described.

In the following chapter, "The Theoretical Framework for the Study of the Dynamics of Terminology" the author makes an important stipulation which implies that terms as "empirical objects" and "functional variants of words" from the angle of epistemological conditions constitute a category. As linguistic units they exist on the level of parole. As his point of departure he declares "the epistemological precedence of the concept *terminology* over the concept *term* (which implies a criticism of traditional Viennese terminology theory and practice). Thus, terminology as the preceding linguistic inventory, and *not* individual terms as "empirical objects" should be the subject of systematic terminological studies (p. 31).

Kageura then goes on to develop his ideas of "the target and method of the dynamics of terminology", first and foremost the "systematicity or motivated patterns, observed in the construction of the totality of existing terms of the target domain" (p. 34). The term formation patterns may vary in the course of

time so that studies of different synchronic slices of term material will become imperative. This is also an indicator of terminological growth on a diachronic level.

Part II deals with "Conceptual Patterns of Term Formation". The author formulates the assumption "that the main feature determining the systematicity of terminological dynamics is the system of concepts of a domain" (p. 45). After comparing methods in describing word and term formation, he sets out to elaborate conceptual elements suitable for the description of term formation. The formal structure of a complex term is the determinant which precedes the nucleus or head. By way of what traditional word (and term) formation has called "left-hand branching" (Kageura does not use this expression, but means this process), a complex term may be enlarged by another determinant. The constituents are connected by closer or looser relationships (in traditional terminological studies known as "bracketing" which also holds for Japanese).

In the past, conceptual categories were suggested by a large number of authors, as early as by Aristotle. Kageura has recourse to categories set up by Sager, Dungworth and McDonald (1980), R. Jackendoff (1983, 1987, 1990), by J.M. Pugh (1984) and a few other authors. The "correspondences of intra-term relations" are listed in an instructive table on p. 97. One wonders why Kageura, as a supportive aspect, has not taken into account the onomasiological tradition of European authors who composed thesauri of words and phrases of *the general language* (not of terminology) by working out a system of conceptual or notional indicators (P.M. Roget, 1852, for English and F. Dornseiff, 1933, for German are outstanding representatives of philosophical and linguistic attempts to classify concepts underlying the words and phrases of the general language).

As Kageura demonstrates in his corpus of documentation terms, the concepts can be neatly divided into subcategories which constitute the terms. The conceptual elements are represented by a complicated set of abbreviations and listed in a large number of tables including the Japanese terms and their English translations, cf. Chapters 5 and 6.

Part III is aimed at the "Quantitative Patterns of Terminological Growth" and introduces mathematical and statistical methods which are visualised by numerous formulae and diagrams opening up an interdisciplinary perspective. The dynamic patterns of the growth of nuclei, determinants, conceptual categories of determinants and conceptual specification patterns are presented in detailed diagrams. The author draws the conclusion "that binomial interpolation and extrapolation provides a very useful and powerful tool for describing the patterns of terminological growth, which can complement the description of the conceptual patterns of term formation" (p. 245).

Part IV, entitled "Conclusions", resumes and further elaborates a number of arguments discussed in the previous parts. The author repeats his dictum: "we distance ourselves from the precedence of concepts over terms" (p. 250), recommends empirical studies of terminology and locates terminology "some-

where in the middle between general natural language vocabulary and artifi-
cial nomenclatures" (p. 251). He emphasises that terminology, due to the sys-
tematicity of concepts, has a systematic aspect.

Kageura's book is concluded by several extensive appendices. Appendix A
presents the "List of Conceptual Categories" as developed in a subtle diversifi-
cation in the previous chapters. Appendix B includes "Lists of Intra-Term Rela-
tions and Conceptual Specification Patterns". Appendix C provides a "List of
Terms by Conceptual Categories" and actually has the character of a Japanese–
English glossary. The English equivalents of the Japanese terms may be com-
plex lexemes such as compounds *(lending library)* and derivations *(programmer;
translator; archivist)* or simple lexemes/simplicia *(library; serial).* Appendix D
consists of a "List of Morphemes by Conceptual Categories" and resembles an
abridged thesaurus. This list includes the "conceptual category tag", the Japa-
nese morpheme, the English translation, and the frequency of occurrence in the
corpus.

The "Bibliography" covers international publications on terminology,
word formation and language statistics, chiefly of the past 40 years. All Japa-
nese names of authors and titles of books appear as transliterations in the Latin
alphabet. The last item of the appendix is a combined author and subject index.

The descriptive value and explanatory power of Kageura's conceptual
analysis of term formation are of general importance, reaching far beyond the
Japanese language and the domain of documentation. As nouns occupy a
prominent place in the terminology of any domain, their internal conceptual
relations established in the term-formation process, may, in principle, also be
expected in other languages, irrespective of their typological or genealogical
properties.

Moreover, the author offers the possibility of making predictions based on
statistical data as to which term-building patterns will be preferred and become
productive for neologisms in a particular domain.

The practical applicability of Kageura's findings is still open to discussion.
The term lists in the appendix may be a stimulus to lexicographers and termi-
nographers, and, after a reasonable selection and didactic presentation, the
term formation patterns might even lend themselves to teaching terminology in
classes of languages for special purposes (LSP).

Rosemarie Gläser
Am Staffelstein 6
D-01328 Dresden
Germany

Hans-Dieter Kreuder. *Metasprachliche Lexikographie: Untersuchungen zur Kodifizierung der linguistischen Terminologie*. 2003, X + 271pp. Lexicographica. Series Maior 114. ISBN 3-484-39114-6. Tübingen: Max Niemeyer. Price: €78.

This publication, which results from Kreuder's "Habilitationsschrift" of 2001 at the University of Marburg, investigates the development of terminological reference works in Germany within the field of linguistics, and approaches in the compilation of these works. Kreuder (2003: 3) defines "special-field terms in linguistics" or "linguistic terminology" as those terms belonging to a special class of lexical items on a metalinguistic level by means of which information about language can be conveyed. His main aim is to determine how linguists and lexicographers presented special-field linguistic terms up to the late 1990s, and to establish the usability of a wide range of especially German linguistic dictionaries with regard to the needs of the intended users. In this investigation, the conceptual strong points and weak points of the dictionaries under discussion are specifically mentioned.

Kreuder's study specifically describes the history and development of linguistic terminological dictionaries aimed at German-speaking users. It is therefore of primary importance to German-speaking readers who want to follow the history of this particular dictionary genre. Terminologists, lexicographers and linguists especially who regularly work with this type of terminology will benefit from taking note of this publication. For other users who do not work within the realm of German linguistic terminography, Kreuder's work may also be valuable, because it discusses specific problems of terminological dictionaries, notably with regard to linguistic terminology, which might also be valid for dictionaries of this nature in other languages. His findings and recommendations at the end of the book are of special interest.

In addition, the book gives the reader a good idea of the field of dictionary criticism, one of the components within the greater field of dictionary research (cf. Wiegand 1984). The second part of this book is particularly helpful in this regard, when Kreuder for instance investigates whether the titles and prefaces of the dictionaries correspond with the needs of the intended users, and with the resulting contents. In addition, he compares the lemma collections of selected dictionaries, as well as their way of presentation.

Kreuder also gives the reader an idea of how difficult it is to establish the terminology in a special field; the situation of German linguistic terminology can therefore serve as example for all those special fields in other languages where terminology has to be established. He sketches the debates, the problems, the failures of some projects, and finally the results of others, whether they were successful or not, and, if not successful, the reasons for their failure. One lesson that can be learnt from Kreuder's book is that work on a terminological dictionary, and especially a linguistic one, is never concluded. He describes many instances of dictionary projects which started out with great enthusiasm, but which were abandoned after some time, or others which only came into being very slowly.

The book is divided into two parts; the first part gives an overview of the development of some European and mostly German linguistic terminological reference works. Kreuder (2003: 169) aims to present a step-by-step description of the development of linguistic terminological dictionaries in Germany, in which he also takes into account the social-historical circumstances and the relationships between the various dictionaries.

Chapter 1 surveys the use of linguistic terminology in the nineteenth century, specifically with reference to the then current linguistic approaches such as the comparative-historical method (Kreuder 2003: 2, 5), which strove towards establishing grammatical research as a scientific discipline. The main problem of this period was the coining of numerous new terms in Germany (and elsewhere), and the question was hotly debated whether one should use the new German terminology, or adhere to the Latin, or even use both alongside each other. This debate reminds one of contemporary debates about the all-pervasiveness of world languages such as English in comparison with smaller, marginalised languages which have not established their own terminology, but wish to do so. Kreuder sketches the endless discussions about the pros and cons of using either Latin or German which took place among linguists (2003: 9-15), educators (2003: 15-22), and within the language society the Allgemeiner Deutscher Sprachverein (2003: 23-27).

Chapters 2 and 3 deal with the planning of linguistic terminological dictionaries in the first three decades of the twentieth century in Germany and within a wider European context. This period signifies the first attempts to codify linguistic terminology, in order to establish a unified terminology for use in schools and universities. This process started in 1906 in France, at a conference where several national committees were formed to work together on a "terminology which will suit all languages", but each committee concentrated only on the terms in its own language (Kreuder 2003: 29). A Joint Committee on Grammatical Terminology, formed in England in 1908 after a conference of the Classical Association, was explicitly commissioned for finding terms in five languages, namely English, German, French, Latin and Greek (Kreuder 2003: 31). These large-scale efforts faced resistance in several countries. German-speaking countries were, however, inspired to deal with terminological issues more intensively, as can for instance be seen in the formation of the Neuphilologischer Verein in Vienna (Kreuder 2003: 33). There were plans for a *Wörterbuch der sprachwissenschaftlichen Terminologie* (WST) in Germany and a *Léxique de la terminologie linguistique* in France. It seems, however, that in the German-speaking world, the debate for or against the use of German instead of Latin terms became a big, and even political, issue during the years between the two World Wars (cf., for example, Kreuder's description of Krüger's plea for "real German", and against "gibberish" which "poisons the German nation, its nature as well as its language" (Kreuder 2003: 39)).

In the 1920s, differences in the aims of linguists versus philological scholars in general came to the fore, when the former group wanted to make allow-

ance for the entire special field of their discipline, while the latter group was more interested in the nomenclature of school grammar (Kreuder 2003: 53). At the First International Linguistic Conference in The Hague in 1928, this effort of linguists to avoid parochialism was stressed, and it was explicitly mentioned that terminological problems could never be solved if researchers from all the countries did not cooperate. Each country had to work on its own terms, which would then be scrutinised by a central committee. All these plans took a very long time to be realised. In September 1932, however, the central committee had its first session in Frankfurt, where the first specimen articles were scrutinised. It became clear that this enterprise was a very difficult one, because the semantic delineations of the terms in the different languages were so different. After this Frankfurt conference, the chairperson of the commission, Schrijnen, circulated a questionnaire in which different equivalents had to be evaluated. Interesting facts became evident from the outcomes of this questionnaire. It was revealed that many languages, amongst which Afrikaans, lacked a large number of these terms or their terminology was based on completely different systems of approach (Kreuder 2003: 54-61). Later, in the 1930s, the political situation in Germany made international cooperation too difficult, and the commission dissolved after Schrijnen's death and Jakobson's exile in America.

The first comprehensive work on linguistic terminology appeared in 1933 from the hand of the French linguist and Latinist Marouzeau under the title *Léxique de la terminologie linguistique* (Kreuder 2003: 65-69). He already started working on this dictionary before the outbreak of the First World War. Schmitt compiled some specimen articles for the planned WST and published them in an addendum to the 51st volume of the journal *Indogermanistischen Forschungen* (Kreuder 2003: 70-73).

Chapter 4 treats linguistic terminological dictionaries from 1940–1960, specifically within German-speaking countries. By the beginning of the 1940s, the influence of the comparative-historical method in linguistic studies had waned, and structuralism became popular, which meant that there were various national schools in the different countries, each with its own conceptual tools. Because of the political circumstances, Germany was isolated, and this also had its impact on German linguistics (or "Sprachwissenschaft" as it was mostly called) — an isolation which was only broken in the 1960s. During this period, however, many linguistic terminological dictionaries appeared in Germany (Kreuder 2003: 74).

The long-awaited WST, for which the plans had already started before the First World War, never came into being, even though several scholars worked on it, and in spite of several efforts to revive it (Kreuder 2003: 75-78). The first German terminological dictionary of linguistics by Hofmann and Rubenhauer appeared in 1950 in Heidelberg under the title *Wörterbuch der grammatischen und metrischen Terminologie*, which contained very short explanations (according to Kreuder 2003: 81, "in telegram style", therefore often leading to misinterpretation). After this dictionary, preparations were made by Wissmann in Ber-

lin to undertake the compilation of the *Historisches Wörterbuch der sprachwissen-schaftlichen Terminologie* (HWST). His concept included all European languages with regard to their grammar; to its descriptive-definitorial function, he also added a historical one (Kreuder 2003: 85). He realised that this would be a very scholarly publication, and therefore also aimed at publishing an additional "smaller one for students". This effort, however, got stuck in methodological problems related to the historical basis of the project, so that by 1967 it was abandoned (Kreuder 2003: 85-89).

Kreuder furthermore discusses the *Sprachwissenschaftliches Wörterbuch*, of which a first group of lemmata was published in 1961 by Knobloch from Innsbruck. It was intended as a type of textbook, with useful information on the origins and usage of terminology. In the last forty years, additional sections had appeared from time to time, and the first complete volume was published in 1986. By 1998, however, this work had only treated terms up to G in the alphabet (Kreuder 2003: 89-94, 154-155).

Chapter 5 discusses the changes in approach to linguistic terminological dictionaries during the 1960s and 1970s, and the rapid development of linguistic terminography in Germany. While by the mid-1960s in the United States of America the phrase-structure grammar had already been replaced by a standard model of the transformational grammar, structural linguistics was vehemently gaining ground in Germany. This brought about the coining of many new terms, and a different approach in defining these. Kreuder (2003: 96-131) discusses various works which were published during this period.

Dictionaries which appeared in the German Democratic Republic range from, e.g Lang (1967) who took the terminology of the transformational generative grammar into account, via Helbig (1969) who included terms from all linguistic areas, to Conrad with his *Kleines Wörterbuch sprachwissenschaftlicher Termini* (1975), which had second and third editions in 1978 and 1981 respectively, and his *Lexikon sprachwissenschaftlicher Termini* (1985), and its second edition of 1988. In the Federal Republic of Germany, there were dictionaries such as the ones by Ludewig (1969) that was much criticised; Bohusch (1972) who included a wider range of terms than Ludewig, but whose dictionary was not very informative; Rucktäschel (1971ff) who, over a period of time, published terms in consecutive volumes of the journal *Linguistik und Didaktik*; Ulrich (1972) whose dictionary had a second edition in 1975, a third in 1981, and a fourth in 1987; Heupel (1973) that had a second edition in 1975, and a third in 1978; Lewandowski (1973–1975) who gave detailed references to literature, including that of Eastern Europe, making it the best dictionary with regard to the richness of information it provides, so that the second edition of 1976 was followed by a third in 1979–1980, a fourth in 1984–85, and a fifth edition in 1990; Welte (1974); Abraham (1974) with a second edition in 1988; and Stammerjohan (1975).

Chapter 6 presents an overview of linguistic terminological dictionaries published in the 1980s and up to the 1990s. New dictionaries which were added

to the wide range of further revised editions of the existing ones, were those by Spiewok et al., one in 1976 with the title *Wörterbuch grammatischer Termini*, and one in 1977 with the title *Wörterbuch stilistischer Termini*. Others were compiled by Bräuer and Bartels (1979) and by Bußmann (1983), the latter having had a second edition in 1990.

Kreuder concludes the first part of his book by stating that all these efforts resulted in a wide range of dictionaries differing from each other with regard to their lemma selection, way of presentation, and the quality of the information given. A comprehensive dictionary which would be informative to all students and scholars of German linguistics does not yet exist (Kreuder 2003: 164). Most of the discussed dictionaries not only exhibit conceptual flaws, because of the individualistic selection principles on which they are based, but also show many tendencies towards experimentation with regard to the textual presentation of data.

According to Kreuder (2003: 167), German linguistic terminology is still in its infancy in spite of the many publications on lexicography which tried to lay the foundation for this type of dictionary. To illustrate this claim, the second part of Kreuder's book presents an analysis of certain features of selected dictionaries published between 1967 and 1990. For example, in Chapter 7, viewpoints are discussed on what exactly is understood by the terms *lemma* and *dictionary article* in the different dictionaries. Here, Kreuder also refers to scholarly works dealing with lexicographic terminology, such as Zgusta (1971) (cf. Kreuder 2003: 173-174 and 177-179), and Wiegand (1983), who especially aims to provide a conceptual framework within which lexicographers can plan their dictionaries better by using the terminological tools he introduces. Wiegand identifies an entire network of terms relating to the notions *lemma* and *dictionary article* (Kreuder 2003: 174-175).

In his empirical analysis, Kreuder not only investigates the outer form of the dictionaries, the manner of arrangement and the presentation of the special-field contents, but also the lemma selection and the quality of the explanations (Kreuder 2003: 1). He wants to demonstrate the unique characteristics of the lemma collection, of the article structures, of the information presented, and of the relative emphasis of the different dictionaries on certain areas within linguistics (Kreuder 2003: 169). His aim is to show the strong and the weak points of the dictionaries, and contradictions which may present themselves.

Chapter 8 introduces several obligatory and optional building components of dictionary articles which Kreuder considers as important in special-field dictionaries. The lemma is one such obligatory component (2003: 181-182), as is the lexicographic definition (2003: 182-185). Optional components are, for example, items indicating pronunciation, grammatical items, foreign-language equivalents, abbreviations of terms, synonyms, items indicating usage, antonyms, etymological information, examples, and bibliographical details.

In Chapter 9, Kreuder looks at the titles and the prefaces of the above-mentioned dictionaries in order to establish whether the contents in the dic-

tionaries fit their descriptions. It becomes clear that the titles of most of the dictionaries create problems, because in most cases, the compilers do not explicitly state their viewpoints on whether "linguistics" is meant as a generic term, and whether the German term "Sprachwissenschaft" is synonymous with "Linguistik" or rather stands in a hyponymic relationship with it. Kreuder gives a graphic illustration in which the dictionaries under discussion are classified according to their titles (2003: 199). For example, seven of the dictionaries in question used titles indicating the specific area within the special field (e.g. Spiewok et al.'s titles *Wörterbuch grammatischer Termini* (1976) and *Wörterbuch stilistischer Termini* (1977)). The rest of the dictionaries have titles which do not specify the exact special field. In this category, Kreuder distinguishes between the ones which use the term *Sprachwissenschaft/sprachwissenschaftlich* (e.g. Conrad's *Kleines Wörterbuch sprachwissenschaftlicher Termini*) and the ones using the term *Linguistik/linguistisch*. The latter category is divided into those with specifying delimitation of the topic (e.g. Abraham's *Terminologie zur neueren Linguistik* (1975)) and those without a delimitation (e.g. Lewandowski's *Linguistisches Wörterbuch*).

Even when the prefaces to the dictionaries are analysed, it seems that most of the lexicographers did not give explicit specifications about the title and its relationship to the contents of the dictionary or to the potential users. By means of a graphic illustration, Kreuder (2003: 203) classifies selected dictionaries according to the use of *Sprachwissenschaft* or *Linguistik* in their titles, comparing these descriptions with the statements in their prefaces. There are those, for example, which consider *Linguistik* as a synchronic study only (e.g. Heupel 1973), and those which see it as both a synchronic and diachronic study. The latter group can be subdivided into those which present a selective lemma collection (e.g. Stammerjohan (1975), who concentrates on specific themes, and Helbig (1969), who concentrates on certain aspects of terms), and those which strive towards presenting a universal lemma collection. Of the latter class, some are still in progress (e.g. Knobloch 1961ff), and others are already completed. These completed dictionaries can be divided into those with a specific circle of addressees (e.g. Ulrich 1972) and those with a more general circle of addressees. Once again, this category can be subdivided into those concentrating on older linguistic research, newer linguistic research, or the latest linguistic research.

Chapter 10 compares the lemma selection of several closely related dictionaries. When one looks at Kreuder's exposition of lemma list samples (2003: 204-205, 206-208), it is remarkable how many lemma gaps occur in the various dictionaries. This is all the more disturbing when one considers that some of the dictionaries claim to be "comprehensive" with regard to their lemma selection. This becomes even more complex when Kreuder states that most of the lexicographers did not mention their selection criteria at all — seemingly most of them were simply intuitive in their selection (2003: 210). In addition, when they claim that terms in "general use" were included, there is no mentioning of

criteria according to which the frequency of usage was determined (Kreuder 2003: 211). He suggests that, to determine frequency of use, the methods used in contemporary corpus lexicography can also be applied in linguistic terminology.

Chapter 11 presents a comparative study of a specific lemma (the lemma *freie Angabe*) as it was treated in selected dictionaries. The notion *freie Angabe* stems from valency theory. Of the nineteen dictionaries Kreuder investigated, six do not present the lemma *freie Angabe* at all. Six others only have the lemma *Angabe*, but do not use it in the sense it has in valency theory. The remaining nine dictionaries do refer, under the lemma *Angabe*, to the meaning of *freie Angabe*, as determined in valency theory, but only one (Abraham 1974) specifically mentions the expression *freie Angabe* as such. The definitions differ widely, and the semantic content of the term *freie Angabe* is therefore also presented in various ways. This applies not only to the obligatory components of the dictionary articles, but even more so to the optional components. In a graphic illustration, Kreuder (2003: 234) points out that Lewandowski's fourth and fifth editions exhibit the most items with optional components, namely terminological synonyms, antonyms, area of usage, etymology, examples, and bibliographical references which could elucidate the term *freie Angabe*.

Finally, Kreuder draws some conclusions by means of which he hopes to give some guidelines for future dictionaries of linguistic terminology. He fears that the picture is very negative with regard to a systematic presentation of linguistic contents in the dictionaries under consideration (Kreuder 2003: 238). In order to obtain dictionaries which serve their purposes better, and which are compiled according to a more systematic approach, Kreuder suggests the following:

(a)   In the planning of terminographies such as these, the target group has to be specified more explicitly. The wider the intended user group, the more difficult it is to plan the lemma selection and the special-field content of the dictionary articles.

(b)   The title and subtitle should reflect the function of the planned dictionary, and the needs of the intended users. In addition, the preface should explicitly mention the intended user group and the related criteria for the lemma selection and presentation of the special-field data.

(c)   Bibliographical references are important for this type of dictionary. There should not only be a list of authors, but also useful references with dates of publication and page numbers.

(d)   The ways of presenting cross-references should be carefully considered. These should not confuse users.

(e)   In the front matter of such dictionaries, guidelines should be included about the way of using the dictionary and the manner of the data presentation. A glossary of foreign language terms is useful, and a bibliography is necessary.

(f)   For this type of dictionary, teamwork is better than operating as an indi-
vidual. The field of linguistics is so specialised that no single person has
a mastery of all its aspects. Cooperation should be established, with strict
rules and good communication to ensure consistent treatment of the
data.

(g)   Dictionaries of this type should not be planned to have a very long life.
Rather, the possibility should be created to revise them frequently, and
for this, the computer is an excellent medium, because dictionaries could
be placed on the internet or on CD-ROM. If this is not possible, a card
filing system is also acceptable, where cards for newer editions could
replace older ones.

(h)   Publishers and lexicographers should be realistic in their planning. They
should always take the needs of the potential users into account.

Bibliography

Wiegand, H.E. 1983. Was ist eigentlich ein Lemma? Ein Beitrag zur Theorie der lexikographischen
Sprachbeschreibung. Wiegand, H.E. (Ed.). 1983. *Studien zur neuhochdeutschen Lexikographie III.
Germanistische Linguistik* 1–4/1982: 401-474. Hildesheim/New York: Georg Olms.

Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hart-
mann, R.R.K (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on
Lexicography at Exeter, 9–12 September 1983*: 13-30. Lexicographica. Series Maior 1. Tübingen:
Max Niemeyer.

Zgusta, L. 1971. *Manual of Lexicography*. Janua Linguarum. Series Maior 39. The Hague/Paris:
Mouton.

Maria Smit
*Department of Music
University of Stellenbosch
Stellenbosch
Republic of South Africa
(msmit@sun.ac.za)*

Johann R.E. Lutjeharms. *Oseanologiewoordeboek Engels–Afrikaans/Dictionary of Oceanology English–Afrikaans*. 2004, xvi + 334 pp. ISBN 186890041X. Kaapstad/Cape Town: Pharos Woordeboeke/Dictionaries. Prys: R175.

Een van die opvallende kenmerke van die onlangse ontwikkeling in die teoretiese leksikografie is die klem op die onderskeid tussen algemene leksikografie en gespesialiseerde leksikografie en die daaruitspruitende opbloei in die leksikografiepraktyk van vak- of gespesialiseerde woordeboeke. Hierdie nuwe generasie vakwoordeboeke is leksikografiese naslaanbronne waarvan die samestelling voorafgegaan is deur die opstel van 'n volwaardige woordeboekplan wat voorsiening maak vir sowel 'n vakkundige basis as 'n stewige metaleksikografiese grondslag. Soos in die algemene leksikografie staan die gebruiker en sy/haar behoeftes ook sentraal in die opstel van vakwoordeboeke, en in die vasstelling van die gebruikersprofiel word die aard en omvang van die leksikografiese bewerking grootliks bepaal deur leksikografiese funksies. In hierdie verband is multifunksionele woordeboeke dikwels die teiken van die leksikografiese proses met 'n aanbod wat vir sowel kennisgerigte as kommunikasiegerigte funksies voorsiening maak.

Hierdie ontwikkeling in die vakleksikografie lei tot produkte wat gekenmerk word deur veel uitvoeriger bewerkings as wat vroeër tipieserwys die geval was. Daar word dikwels wegbeweeg van 'n vakwoordeboek wat slegs 'n lys brontaalitems met hulle vertaalekwivalente in die doeltaal bevat ten gunste van betekenisverklarings of bykomende konteks- en koteksleiding. Maar die ontwikkeling van die vakleksikografie loop ook parallel met die behoeftes van spesifieke teikengebruikers en hierdie behoeftes mag juis soms die daarstelling van 'n vakwoordeboek noop wat in sy omvang of in die aard van sy leksikografiese bewerking verskil van talle vakwoordeboeke wat in die onlangse verlede gepubliseer is.

Suid-Afrikaanse oseanoloë het toegang tot vakpublikasies waarin die internasionale terminologie van die vakgebied uiteengesit word. Vir begripsprobleme is hierdie bronne beskikbaar en toeganklik. Die terminologie wat in hierdie bronne aan die orde gestel word, is egter slegs in Engels. Met die *Oseanologiewoordeboek Engels–Afrikaans/Dictionary of Oceanology English–Afrikaans* (voortaan afgekort as OW) poog Lutjeharms om aan die behoefte van Suid-Afrikaanse oseanoloë te voldoen wat vertroud is met die Engelse vakterme maar wat op soek is na Afrikaanse ekwivalente daarvoor. En dit is 'n werklike behoefte van werklike gebruikers wat bepalend was vir die aard en omvang van, asook die leksikografiese aanbod in die OW. Die eenrigtingkarakter van die OW, te wete 'n bewerking wat slegs vanuit Engels na Afrikaans strek, moet ook in die lig van die gebruikersbehoeftes verstaan word. Op hierdie vlak lewer OW 'n bydrae wat nêrens betwis mag word nie.

Die oseanologie is 'n vakgebied wat uit verskeie onderafdelings bestaan, onder meer 'n fisiese, 'n chemiese, 'n biologiese en 'n geologiese komponent. Dit verbreed die omvang van die relevante vakterminologie aansienlik. Daarbenewens oorvleuel die oseanologie in sy terminologie met 'n verskeidenheid ander

vakgebiede, bv. die waterboukunde en kusingenieurswese, terwyl skeeps-, seemans- en navigasieterme eweneens vir die oseanoloog ter sake is. Binne die vakleksikografie met sy onderskeiding tussen subveld-, enkelveld- en multiveld-vakwoordeboeke vra die oseanologie om 'n multiveldvakwoordeboek. Een van die groot probleme vir die opsteller van so 'n woordeboek is die keuse van terme vir opname en weglating. In sy voorwoord stel Lutjeharms dit duidelik dat dit die mikpunt van hierdie woordeboek is om 'n gebruikersvriendelike en doelmatige bron te wees. Ook in sy lemmakeuse het hy hom hierdeur laat lei en die gebruiksfrekwensie van terme as 'n bepalende kriterium gebruik vir opname al dan nie. In hierdie verband sluit sy werkswyse regstreeks aan by 'n dominante benadering in die korpusgedrewe leksikografie waar gebruiks-frekwensie so 'n belangrike rol speel.

Die OW vertoon 'n raamstruktuur en die belangrikste funksionele buite-tekste is 'n inhoudsopgawe, 'n voorwoord en 'n toeligtingsteks as voortekste (al drie hierdie tekste word in Afrikaans en Engels gegee). Die raamstruktuur van die woordeboek word voltooi met 'n enkele agterteks, nl. 'n teks met bronne en verwysings. Die toeligtingsteks is beperk tot 'n verklaring van enkele van die konvensies wat in die woordeboek gevolg word. Vir die kundige gebruiker bied dit genoegsame leiding om die verlangde inligting aan die woordeboek te kan onttrek.

Met 'n eerste oogopslag lyk dit of die sentrale teks met sy, volgens die agterbuitebladreklame, meer as 30 000 oseanologiese en tegniese terme slegs 'n lysting van Afrikaanse lemmata met hulle Engelse vertaalekwivalente bevat. Nadere ondersoek lewer tog bewys daarvan dat die bewerking dikwels veel verder as dit strek. Die verpligte mikrostruktuur, dit is die struktuur van die verstekartikels met die minimum leksikografiese bewerking, bied weliswaar as bewerking van die lemmateken slegs 'n enkele vertaalekwivalent, maar daar-naas vertoon talle artikels 'n uitgebreide mikrostruktuur waar die enkele ver-taalekwivalent aangevul word met ander inskrywings. Hier gaan dit veral om doeltaalsinonieme wat aan die gebruiker 'n keuse van vertaalekwivalente bied. Die lysting van vertaalekwivalente is nie arbitrêr nie, maar word in 'n volgorde van verkieslike gebruik gegee; al is dit, soos die leksikograaf in die toeligting sê, 'n subjektiewe volgorde. Hier kan die gebruiker hom/haar ten minste ver-laat op die subjektiewe oordeel van 'n deurwinterde vakkundige.

Vir die voorkoms van divergensie as ekwivalentverhouding soos blyk uit die aangawe van meerdere vertaalekwivalente in die bewerking van een lemma maak die toeligtende aantekeninge voorsiening deur aan te dui dat die optrede van meerdere vertaalekwivalente doeltaalsinonimie impliseer. Dit gaan dus om leksikale divergensie. Die vertaalekwivalente wat soms ter bewerking van een lemma aangebied word, lyk egter nie altyd na doeltaalsinonieme nie, maar eerder na ekwivalente ter bewerking van verskillende polisemiese waardes van die lemma. Die bewerking van die lemma *fetch* lyk na so 'n voorbeeld van semantiese divergensie:

fetch: stryklengte, aanloop, ontstaangebied, bereik, opwekafstand, haal

In 'n artikel soos hierdie is bykomende inskrywings nodig om die gebruiker met die nodige konteksleiding te help.

Sinonimie-aanduiding word nie net ten opsigte van die doeltaal verstrek deur die toepassing van leksikale divergensie nie, maar 'n bykomende prosedure van lemmatiese adressering bied ook, waar toepaslik, sinonieme vir die brontaalvorm. Hierdie brontaalsinonieme volg, in hakies, onmiddellik op die lemma en word, soos die lemma, ook in vet druk gegee. Deur hierdie tipografiese struktuurmerker word die doeltaalsinoniem nie net ter bewerking van die lemma aangebied nie maar word dit ook 'n primêre bewerkingseenheid en die slotkomponent van 'n dubbellemma. Soms word daar meer as een brontaalsinoniem gegee, maar in hierdie gevalle bied die toeligtende aantekeninge ongelukkig geen aanduiding van 'n voorkeurvorm of 'n keuse volgens gebruiksfrekwensie nie. Hierdie probleem word daardeur vererger dat al die brontaalsinonieme op hulle beurt lemmastatus kry met telkens 'n aanduiding van die ander lede van die sinoniemparadigma asook die vertaalekwivalente. Per gevestigde leksikografiese konvensie word die gebruiklikste lid van 'n brontaalsinoniemparadigma uitvoerig bewerk terwyl die ander lede as lemmata van verwysingsartikels verstrek word. Hierdie konvensie is 'n vorm van leksikografiese teksverdigting wat veral in 'n woordeboek met kundige gebruikers ruimtebesparing in die hand werk. Ook in OW sou dit met welslae toegepas kon gewees het.

OW maak ongelukkig nie 'n onderskeid tussen brontaalsinonieme en brontaalwisselvorme nie. Wisselvorme, bv. *faetal* en *fetal*, word op dieselfde manier as brontaalsinonieme aangebied.

In OW word dieselfde vertaalekwivalentparadigma soms in die bewerking van al die brontaalsinonieme aangebied, maar soms is daar 'n effense verskil. Die gebruiker kry egter geen leiding oor hoe om hierdie verskille te interpreteer nie. Vergelyk die geval *reference station* met *standard station* en *standard port* as sinonieme. In die bewerking van hierdie drie brontaalsinonieme word dieselfde vertaalekwivalentparadigma, nl. *verwysingspunt, verwysingstasie* telkens gegee. Dit is 'n stelselgetroue aanbieding wat egter nie in alle artikels gevolg word nie. In hulle optrede as lemmata kry elk van die brontaalsinonieme *bacterial film, diatom film, primary film* en *slime film* 'n lysting van al die ander lede van die sinoniemparadigma. Hieruit lei 'n mens af dat hulle absolute sinonieme is. Die vertaalekwivalentparadigma wat aangebied word, verskil egter van geval tot geval:

bacterial film primêre laag, primêre vlies

diatom film diatomeelagie, kristalwierlagie, slymlaag, primêre laag, primêre vlies

primary film primêre laag, primêre vlies, slymlaag

slime film slymlaag, primêre laag, primêre vlies.

Vir die gebruiker wat op soek is na die mees gepaste Afrikaanse ekwivalent is hierdie soort bewerking onbevredigend.

Die multidissiplinêre aard van die oseanologie en die gevolglike multiveldkarakter van hierdie woordeboek noodsaak die gebruik van leksikografiese etikette om konteksleiding te gee. Lemmata word geëtiketteer, bv. die lemma *lee wave* met sy etiket *(meteorology)* en die lemma *needle* waar die etiket *(e.g. ice)* konteksleiding verskaf. Etikette word ook in die vertaalekwivalentparadigma gebruik, byvoorbeeld waar al die vertaalekwivalente nie tot dieselfde veld of subveld behoort nie. Vergelyk in hierdie verband die artikel van die lemmateken *deflection*:

> **deflection:** afwyking *(rigting)*, uitwyking *(geologie)*, defleksie *(strale, algemeen)*

Hierdie tipe etikettering kom algemeen in die OW voor en dra veel daartoe by om die pragmatiese verankering van bron- en doeltaalvorme te verseker.

Die makrostruktuur van die OW maak nie net vir woordlemmata voorsiening nie, maar ook vir meerwoordige leksikale items wat as multileksikale lemmata opgeneem word. Vergelyk die opname van lemmata soos *finite difference method, mean sounding depth* en *reversals of the earth's magnetic field*. Hierdeur slaag die OW daarin om die oseanologiese leksikon in sy breë omvang weer te gee. Ook die vertaalekwivalente is dikwels meerwoordige leksikale items of selfs 'n bondige parafrase wat as surrogaatekwivalent optree. Die makrostruktuur bied naas hooflemmata ook sublemmata. Sublemmata word in 'n blokvorm as 'n lemmanes aangebied. Vergelyk die volgende voorbeeld:

> **production** ...
> > **production, androgen** ...
> > **production, bacterial** ...
> > **production, carbon** ...

Alhoewel die stelsel nie in die toeligtende aantekeninge verklaar word nie, behoort die gebruiker nie probleme hiermee te ervaar nie aangesien die stelsel konsekwent toegepas word. Wat leksikografies minder aanvaarbaar is, is die feit dat die sublemmata elders op hulle alfabetiese plek as hooflemmata opgeneem word: *androgen production, bacterial production*, e.s.m. Hierdie soort duplisering behoort vermy te word.

In OW is daar verskeie ander kleiner lastighede en stelselprobleme, maar die foute word volledig oorskadu deur die bydrae wat hierdie woordeboek maak. Die bydrae is enersyds die bymekaarbring van die Engelse terminologie, maar andersyds, en veral, die aangawe van die Afrikaanse vertaalekwivalente. In hierdie verband speel die OW 'n belangrike rol in die vestiging en die stan-

daardisering van die oseanologiese vakterminologie. OW is 'n wesenlike toe-voeging tot die Afrikaanse vakleksikografie en prof. Johann Lutjeharms ver-dien erkenning vir 'n werk van onbetwisbare gehalte.

R.H. Gouws
Department Afrikaans en Nederlands
Universiteit van Stellenbosch
Stellenbosch
Republiek van Suid-Afrika
(rhg@sun.ac.za)

428   Resensies / Reviews

Die Taalkommissie van die Suid-Afrikaanse Akademie vir Weten-skap en Kuns. *Afrikaanse Woordelys en Spelreëls*. Negende, verbeterde en omvattend herbewerkte uitgave. 2002, 592 pp. ISBN 1 86890 034 7. Kaapstad: Pharos Woordeboeke. Prijs: R170.

Herdrukken van wetenschappelijke publicaties of andere standaardwerken worden zelden uitvoerig besproken. Meestal moeten zij het stellen met een aankondiging in de vakliteratuur. Hoe anders is dat als er een herdruk ver-schijnt van een van de codices waarin de standaardtaal is beschreven of waarin de regels die een optimale communicatie binnen een taalgemeenschap moeten bevorderen worden uitgelegd. Grammatica's, woordenboeken en woordenlijs-ten behoren tot die laatste categorie.

In 2002 verscheen in opdracht van Die Suid-Afrikaanse Akademie vir We-tenskap en Kuns de negende uitgave van *Afrikaanse Woordelys en Spelreëls* (AWS). Het betreft een woordenlijst van ongeveer 28 000 trefwoorden. Dat aan-tal geeft al aan dat het niet gaat om een volledige inventaris van het hedendaags Afrikaans. Een agglutinerende taal als dat Afrikaans kent immers een woord-voorraad die meer dan 100 000 verschillende lexicale eenheden omvat.

In haar verantwoording bij deze druk geeft de voor de samenstelling ver-antwoordelijke 'Taalkommissie' aan wat zij precies gedaan heeft sedert de ver-schijning van de achtste druk in 1991. Er is veel aandacht geschonken aan de verfijning en vereenvoudiging van de spelregels. Om dat te bewerkstelligen heeft men gebruikers ingeschakeld bij de herformulering van de regels. Die regels zijn in eenvoudiger taal korter en explicieter beschreven en met veel goed gekozen voorbeelden. De belangrijkste vernieuwing is in dit opzicht evenwel dat de regels in een numerieke volgorde worden aangeboden. Dus, in hoofdstuk 3 dat handelt over afkortingen, vinden we als zevende regel (3.7): "Tradisioneel word punte by afkortings gebruik" en als 3.8 "By afkortings van die name van grade, diplomas en sertifikate is punte opsioneel." Verder is de woordenlijst uitgebreid met nieuwe lemma's en met meer afleidingen en sa-menstellingen. Maar, over de wijze van selectie daarvan geen woord.

Het is dit gegeven dat vragen oproept. Geen enkel lid van een spelling-commissie of een taaladviescommissie durft de stelling aan dat hij alle eenhe-den die spellingproblemen in een taal kunnen opleveren overziet of kent. Cor-pusonderzoek heeft taalkundigen buitengewoon bescheiden gemaakt over hun kennis van de moedertaal. De vraag naar de aard van de samenstelling van de AWS is hier dus gewettigd.

Het Afrikaans is een dynamische taal, lexemen komen en gaan. De nieuwe moeten een officiële spelling krijgen volgens de bestaande regels of naar analo-gie daarvan of door het aanpassen van bestaande regels. Maar hoe weet men wat de nieuwe woorden zijn die problematisch zijn qua spelling? Alleen al het betekenismoment 'nieuw' kan voor verschillende leden van een commissie of taalgemeenschap al arbitrair zijn. Wat de een 'nieuw' noemt, is voor de ander allang geïntegreerd in het taalgebruik en wordt niet meer als nieuw ervaren.

Vaak zijn nieuwvormingen eendagsvliegen of ad hocformaties en die moeten natuurlijk niet worden opgenomen. Maar, wanneer besluit je dat een woord zo levensvatbaar is in een taal dat de vorm ervan in het vigerende spellingsjasje gestoken kan worden? Deze vraag is niet eenvoudig te beantwoorden, maar dat ontslaat niemand van de noodzaak naar een controleerbaar antwoord te streven. Je moet als commissie een referentiekader hebben waaraan je alles kunt toetsen. Een inventaris of corpus dat de voorbeelden oplevert die de regels als het ware bewijzen, maar ook die de commissie uitdagen om de regels scherper of anders te formuleren. Kijk ik naar de samenstelling van de 'Taalkommissie', dan vermoed ik dat ongetwijfeld gebruik is gemaakt van het corpus van *Die Woordeboek van die Afrikaanse Taal*, maar waarom vermeldt men niet met trots dat men de materiaalverzameling daarvan als empirische basis heeft gebruikt voor zijn spellingberegeling en voor de expansie van de lijst? En als men het anders gedaan heeft, bijvoorbeeld op grond van de eigen intuïtie, dan moet men dat eveneens verantwoorden.

Het tweede onderdeel van de AWS wordt gevormd door 'Wenke vir gebruikers'. Grote waardering heb ik hier voor de wijze waarop men een antwoord heeft proberen te geven op de vraag: Hoe bewerkstelligen we dat de gebruiker niet onder het jargon van de regels bedolven raakt? Die vraag is essentieel omdat het overgrote deel van de leden van een taalgemeenschap gehinderd wordt door het jargon van de taalkundige professional. Begrippen als 'consonant, vocaal, predicatief' of 'affricaat' zijn een absolute belemmering om de gebruikers de spellingsregels te doen doorgronden. Ook binnen het Nederlandse taalgebied buigen wij ons over de vraag hoe vermeden kan worden dat het taalkundig jargon de potentiële gebruikers van de *Woordenlijst* negatief beïnvloedt. Wij zoeken onze oplossing in het aantrekken van een op de doelgroep gerichte tekstschrijver, maar de oplossing van de 'Taalkommissie' is stellig het overwegen waard omdat ze geen concessies doet aan de precisie en de nuance van de regels. Zij zoekt de handreiking aan de gebruiker in een afzonderlijke lijst waarin alle moeilijke termen uit de spellingregels zijn ondergebracht en gedefinieerd (blzz. 557 t/m 576). Het probleem wat zich hier evenwel onmiddellijk voordoet, is dat de gebruikers geleerd moet worden dat zij elders in de AWS meer informatie kunnen vinden over de abstracte begrippen uit de definities. Wat de gebruikers ook zullen moeten leren bij het hanteren van die lijst met lexicografisch jargon is geduld. In iedere definitie komen namelijk weer nieuwe jargonwoorden voor die hem verplichten om op andere plaatsen aanvullende informatie te zoeken. Overigens blijft het de vraag of regels zonder enig jargon ondubbelzinnig beschreven kunnen worden. De gekozen oplossing is hoe dan ook een werkbare met als aangename bijkomstigheid dat ze ook pedagogisch van aard is. Immers, door niet-taalkundig geschoolden in aanraking te laten komen met sommige eenvoudige taalkundige termen draagt men ertoe bij dat onbekend niet langer onbemind hoeft te blijven.

Er is in de vorige alinea veel over gebruikers gesproken. En terecht. Maar, wie is nu eigenlijk de beoogde gebruikersgroep van de AWS. En heeft dat 'ge-

bruiken' een wetmatig of een vrijblijvend karakter? Als ik deductief te werk ga, dan weet ik, gezien de afzonderlijke lijst met gedefinieerd jargon, dat AWS niet bedoeld is voor de professionals. Moeten we concluderen dat AWS verplicht is voor ambtenaren en in het onderwijs? En dat andere groepen uit de samenleving de regels als een advies kunnen beschouwen? En hoe verhoudt zich dit dan weer tot een van de drie overwegingen die bepaalden welke woorden geselecteerd werden voor opname? Op bladzijde 14 en 15 lezen we dat afleidingen en samenstellingen opgenomen zijn om het resultaat te laten zien van bepaalde woordvormingsprocédés. Maar, tendeert zo'n overweging niet naar een dienstbetoon aan morfologen? Dus toch voor professionals? En ook de doortastende precisie waarmee op bladzijde 11 geformuleerd wordt dat alleen woorden in de lijst beschreven zijn, die door de eigentijdse taalgebruikers gewoonlijk als gangbaar en als standaardtaal beoordeeld worden, lost mijn onzekerheid over de doelgroep van de woordenlijst niet op. Of moet ik concluderen dat de lijst bestemd is voor iedereen die zich schriftelijk overeenkomstig vigerende spellingsregels en een beproefd spellingssysteem wil uitdrukken zonder dat er op een onjuist gebruik sancties bestaan?

Duidelijk is wel dat de 'Taalkommissie' met AWS een spellingbeleid wil voeren. Wie goed leest, leest op bladzijde 11 dat zij een prescriptieve norm beoogt, gebaseerd op descriptie van het Afrikaans dat door "daardie geslag taalgebruikers" als gangbaar wordt erkend en als standaardtaal wordt ervaren.

Met dit uitgangspunt heb ik geen enkele moeite, maar toch gooi ik hier de handdoek in de ring. Immers, wil men de geschreven taalproductie van de huidige generatie taalgebruikers adequaat be- en voorschrijven, dan heeft men behoefte aan uitkomsten van taalattitudeonderzoek en aan vele zegslieden die men kiest op grond van sociolinguïstische parameters. Dat brengt hoe dan ook het probleem met zich mee van de begrenzing van het begrip 'daardie geslag'. Is dat bijvoorbeeld het taalgebruik van één generatie of van drie? Ik prijs in dit verband het besluit van de 'Taalkommissie' dat zij bij haar werk "verskeie akademici, taalpraktisyns en andere deskundiges" (blz. 9) heeft ingeschakeld, Maar of zij daardoor (het zijn immers allen professionals), een geheel onbetwistbaar antwoord kan geven op de vraag wat gebruikelijk of gangbaar is in het Afrikaans en wat tot de Afrikaanse standaardtaal behoort? Enige wetenschappelijke twijfel of voorzichtigheid is hier hoogstwaarschijnlijk wel op haar plaats. Maar als niet alleen de professionals, maar ook allerlei niet-professionele controlegroepen zijn bevraagd en ik ten onrechte wat azijn op tafel zet, dan neem ik vol respect mijn hoed af en betreur ik het dat de 'Taalkommissie' deze werkelijk nieuwe methodologische benadering niet breder heeft uitgemeten.

Wensen worden in besprekingen altijd breder uitgemeten dan waardering. Maar, wensen leveren misschien een bijdrage aan een nog betere AWS. Het besluit om woorden die ontleend zijn aan andere talen uit Afrika op te nemen verdient een warm onthaal. De bladzijden 203-208 geven een uitstekend inzicht in wat er met deze categorie leenwoorden is gebeurd.

Ten opzichte van de achtste druk is het aantal vaktermen in de AWS drastisch verminderd. De reden die men daarvoor opgeeft, maakt op mij de indruk

van een drogreden. Omdat ze gemakkelijk terug te vinden zijn in technische woordenboeken en vaktaalwoordenboeken zijn ze gereduceerd. Woorden die tot de algemene woordenschat behoren zijn toch ook gemakkelijk terug te vinden in verklarende monolinguale woordenboeken? Ook hier zou het enige richtsnoer de gebruiker en diens gemak moeten zijn. Juist bastaardwoorden en internationalismen veroorzaken onzekerheid over hun spelling en doen gebruikers grijpen naar een boek dat die onzekerheid ongedaan maakt. Uit moeten wijken naar twee of meer bronnen om uitsluitsel te krijgen tasten het vertrouwen van de consument aan in het product dat hij juist aanschafte om zijn problemen met de spelling op te lossen.

De regels voor het spellen beslaan de pagina's 21 t/m 208. In 18 afzonderlijke paragrafen worden ze geformuleerd. Van afbreken, afkortingen, trema, hoofdletters, koppeltekens tot aan verkleiningsvormen en woorden van vreemde herkomst worden de regels beschreven. En die regels zijn niet op zand gebouwd. Natuurlijk rijzen er vragen, maar het geheel lijkt in absolute zin echt meer op een ets dan op een schets. Met de bij de regels gegeven voorbeelden is elke gebruiker in staat om analoge gevallen correct te spellen. Desondanks veroorloof ik mij ook hier een kritische kanttekening. Ik vind dat de 'Taalkommissie' niet prescriptief of normatief genoeg is geweest. Er blijft veel te veel optioneel; er blijft ook veel te veel variatie toegestaan. Van beide een voorbeeld. Regel 3.9 luidt: "By afkortings wat net uit hoofletters bestaan, is punte opsioneel." Ik vind dat de expert hier verder moet gaan omdat hij de ontwikkelingsgang van dit soort woorden kent. Er is bijvoorbeeld in het Nederlands allereerst *H.I.V.* Vervolgens dient zich *HIV* aan als letterwoord. Worden er samenstellingen gevormd met *H.I.V.* dan zien we de volgende fases: *H.I.V.-virus, HIV-virus, hiv-virus* en ten slotte *hivvirus.* Tenzij er sprake is van betekenisdifferentiatie tussen *H.I.V.* en *HIV*, vind ik dat een normerende instantie moet durven overgaan tot het voorschrijven van *HIV*. In aansluiting bij regel 3.14 "Koppeltekens word nie in afkortings gebruik nie, maar wel in samenstellings wat bestaan uit 'n afkorting plus 'n volle woord" vraag ik mij af of deze regel in voldoende mate voorziet in een antwoord op de vraag hoe in het Afrikaans *vigspasiënt* of *radarstasie* beregeld en gespeld worden. Gevallen van acroniemen in samenstellingen als de zo juist genoemde komen slechts terloops ter sprake in bijvoorbeeld de regels 12.9 en 12.11, maar deze kwestie had niet alleen op bladzijde 558 behandeld moeten worden, maar ook in de paragraaf die geheel aan "Afkortings" gewijd is (blz. 32 e.v.).

Te veel variatie zit er in bijvoorbeeld regel 14.6 die de volgende schrijfwijzen sanctioneert: *Bloemfonteinmuseum, Bloemfontein-Museum, Bloemfontein-museum* en *Bloemfontein Museum*. Natuurlijk weet ik ook wel dat iedere speller, hoe getrouw hij de voorschriften ook wil volgen, voor vijf procent een spellingsanarchist blijft, maar al te veel vrijheid in wijze van schrijven zal ertoe leiden dat de wal het schip keert. Ik betwijfel ook of het grote publiek die vrijheid wenst. Men zoek toch in AWS-achtige producten om te weten hoe het moet en niet om te zien hoe het allemaal kan? De gebruikers van woordenlijsten en

lexica zijn bereid het gezag van de lexicograaf te erkennen mits die zich baseert op valide regels en redeneringen. Ik meen evenwel met grote zekerheid te weten dat de huidige 'Taalkommissie' gebonden is aan besluiten van haar voorgangers die in de beregeling van de spelling van samengestelde eigennamen toegeeflijker waren. Waarschijnlijk vindt zij het een zwaktebod om de klok dienaangaande terug te draaien. Er is evenwel een Nederlands spreekwoord dat luidt: Liever ten halve gekeerd dan ten volle gedwaald. In de jaren dat ik in spellingcommissies gezeten heb, heb ik wat oorspronkelijk afgekeurd werd, tot norm zien verheffen en omgekeerd. Respect voor het werk van voorgangers hoeft in mijn ogen nooit een argument te zijn om spellingregels niet te verbeteren of om de door de gebruiker ongewenste dubbelzinnigheid op te heffen. Maar, toegegeven, ik formuleer hier vanuit een West-Europese wetenschappelijke ivoren toren, zonder mij al te nadrukkelijk te realiseren dat spellingsbeleid elders op de wereld genoodzaakt is te streven naar nog grotere compromissen.

Grote waardering heb ik ook voor de paragraaf waarin de commissie met 57 regels probeert het los en vast schrijven van woorden te beregelen (blz. 142-174). Bij de voorbereiding van de *Woordenlijst Nederlandse taal* voor 2015 zal de Nederlandse commissie, waarvan ik dan geen deel meer uitmaak, zich stellig door het Afrikaanse model laten inspireren.

Wat ik ten slotte in hevige mate betreur, is dat de numerieke code van de regels niet terugkeert in de woordenlijst zelf. Ongetwijfeld zullen financiële motieven deze vorm van dienstverlening naar Utopia hebben verwezen. Maar toch. De gebruiker die zoekt bij *sint Nikolaas* zou graag een verwijzing hebben naar 9.26, de regel die bepaalt hoe godsdienstige namen geschreven worden. Hoe de attributieve bepalingen bij zo'n religieuze naam geschreven worden in het Afrikaans blijft voorlopig een mysterie. Althans dat concludeer ik uit de informatie op bladzijde 73. Daar wordt niet gesproken over *sint*, noch over *Heilige Eucharistie, Heilige Gees, Heilige Hostie* of *Hemelse Vader*. Ronduit gebruikersonvriendelijk vind ik de verwijzing binnen de opmerkingen bij een spellingsregel naar andere regels. Als ik in 9.32 de spelling van de namen van talen wil vinden, dan wordt mijn incasseringsvermogen zo smal als een Amsterdams grachtenhuis als ik naar andere bladzijden verwezen wordt. Concreet: ik wil bij 9.32 alle informatie over de wijze waarop taalnamen geschreven worden en niet stuiten op een formulering als: "Vir verdere besonderhede oor die skryfwyse van taalname, kyk Reëls 14.13 tot 14.23." Redundantie neem ik graag op de koop toe.

Naast wat ik hier slechts oppervlakkig de revue heb laten passeren, bevat AWS nog veel meer. Er is een afzonderlijke lijst met afkortingen, een niet uitputtende lijst van representatieve voorbeelden van Afrikaanse woorden die tot de omgangstaal behoren, een lijst van buitenlandse aardrijkskundige namen met hun afleidingen en er is een alfabetische index die AWS nog beter toegankelijk maakt.

Taal en taalgebruik veranderen al naar gelang de maatschappij verandert. Ook de spelling blijft niet onaangetast door die veranderingen. Het is een wel-

daad voor de spelling van het Afrikaans dat er telkens weer leden voor de 'Taalkommissie' gevonden worden die willen piekeren over erudiete details van de spelling om een nog betere geschreven communicatie te bevorderen. Bovenstaande kanttekeningen zijn dan ook slechts bedoeld als een blijk van grote waardering voor het verrichte werk dat elegant uitnodigt tot een systematisch observeren van taalveranderingen en tot een discussie met lexicografen die in het zelfde schuitje zitten.

Piet van Sterkenburg
Instituut voor Nederlandse Lexicologie en (R)UL
Leiden
Nederland
(sterkenburg@inl.nl)

Rita Temmerman. *Towards New Ways of Terminology Description: The Sociocognitive Approach.* 2000, xv + 258 pp. ISBN 90 272 2326 2 (Eur.), 1 55619 772 1 (US) (Hb.). Amsterdam/Philadelphia: John Benjamins. Price: €75, US$75.

This book appeared as Volume 3 in the Series Terminology and Lexicography Research and Practice edited by Helmi Sonneveld and Sue-Ellen Wright. The author, Rita Temmerman, presently working at the Erasmus Hogeschool, Brussels and specialising in problems of terminology in various domains of the life sciences, presents a polemical, stimulating and innovative monograph which continues and deepens her previous research work. Her doctoral dissertation (Louvain 1998) focused on *Terminology Beyond Standardisation: Language and Categorisation in the Life Sciences.* The aim of the book under review, *Towards New Ways of Terminology Description: The Sociocognitive Approach,* is to elaborate a new theory, method and application of terminology research which seeks to overcome the obvious limitations of traditional terminology as chiefly represented by the Vienna School (Eugen Wüster, Helmut Felber, Infoterm and associated institutions).

In Temmerman's view, traditional terminology (for which she uses the metalingual expression *Terminology* with a capital letter) has actually impeded the creation, definition and placement of terms in a conceptual system. This argument is repeated in various contexts and nearly becomes a guiding motif throughout the book. In Chapter 1, entitled "From Principles of Standardisation to a Scientific Study of Terminology" she discusses the main principles, and also the shortcomings, of Terminology:

(1)   the onomasiological perspective
        (which implies how postulated concepts derived from the outside world are designated by terms)

(2)   the proposition that concepts are clear-cut
        (so that fuzzy areas, and overlaps should be excluded from the terminological system which has a well-designed structure of hyperonyms and hyponyms)

(3)   the allocation of terminological definitions to concepts
        (with the distinction between the intensional definition, the extensional definition, and the part-whole definition; preference, however, is given to the meaningful intensional definition of a concept)

(4)   the claim of univocity
        (this implies that each concept should be designated by only one term and one term should refer to one concept only, which in fact excludes cases of synonymy and polysemy)

(5)   the prevalence of synchrony
(the evolution of terms and word formation processes are not the concern of traditional Terminology; the shaping of new terms in technical texts, and specialist discourse in time and space are beyond its scope of interest).

From this general state of the art the author draws the conclusion that "the interest in terminological research was hindered by the interest of standardisation" (p. 15), that univocity of terms is an unrealistic claim, but rather polysemy and synonymy are typical phenomena. She blames traditional Terminology for "objectivism which is the basis of Western thinking" in that this doctrine (or "dogma") describes only isolated objects of the material world and ignores "the human capacity *to understand* and *to imagine*" phenomena, their meaning, function and interrelation (p. 16). Temmerman's critical assessment is unmitigated: "Wüster and his followers turned Terminology into a number of dogmatic principles for terminology description with language engineering, planning and standardisation, as a consequence of product standardisation, as the underlying socio-economic motivations" (p. 17). In her view, the ISO documents in Austria are the logical consequence of this pragmatic handling of terminology.

In the course of this chapter the author describes and assesses "Different schools of Terminology" which have a scientific foundation. Thus, European traditional terminologists believe that the terminology model starts with concepts whose meaning should be clearly delineated — an idea which is opposed by Canadian terminologists. Moreover, traditional terminologists are convinced that terms designate logically stipulated concepts and have no connotations, but can be placed in a system which is structured by "logical and ontological relationships" (p. 20). In contrast to Saussurian structuralist semantics, western terminologists hold the opinion that terms can be studied in separation from text, discourse, social and cognitive processes relating to the growth and restructuring (even a paradigm shift) of a scientific or technological domain and its special language. Here again, Canadian and Scandinavian terminologists hold different opinions (R. Kocourek, G. Rondeau; H. Picht, Ch. Laurén). And Temmerman quotes more international researchers who have levelled criticism at the Viennese School, such as Juan C. Sager (UMIST, Great Britain, 1986), Peter Weissenhofer (1995, Vienna); Britta Zawada and Piet Swanepoel (1994, South Africa), M. Teresa Cabré (1995), Ingrid Meyer (1992, Canada), the Canadian representatives of the new approach of socioterminology (F. Gaudin 1993, J.-C. Boulanger 1995), and the Japanese researcher Kyo Kageura (1995). These critical authors and their objections to the five principles of traditional (Viennese) Terminology are compared in an instructive synopsis (p. 35f.).

In Chapter 2 the author sets out to develop "New Propositions for Terminology" and to furnish alternative arguments for "the four key issues conceptualisation/categorisation, naming or lexicalisation, metaphorical models, and

the diachronic study of categories" (p. 39). She discusses the interplay of language use, the specialist's creative thinking and understanding of domain-specific phenomena, and the dissemination of information in texts where new terms may spring up, but are perfectly understood and become part of a terminological system.

The author draws the necessary empirical data for her corpus-driven alternative approach from a field with which she is deeply familiar: the life sciences. This complex area of research encompasses several domains. These include biology, genetics, molecular genetics, genetic engineering, biochemistry, recombinant DNA technology, biotechnology, and medicine (p. 234). As the author elaborates in a later sub-chapter, biotechnology is an umbrella category.

In Chapter 3 the author demonstrates her new methodology, "From Traditional Definitions of Concepts to Templates of Units of Understanding". At the outset, she makes an important statement: "Categories are all the units of understanding which are impossible to describe according to the principles of traditional Terminology. There is prototype structure, both within and between categories (intracategorially and intercategorially)" (p. 73). By choosing prototypical examples from the life sciences she shows that definitions of the entity *intron*, the activity *blotting* and the collective category *biotechnology* do not originate as stipulations from outside for isolated concepts, but rather develop from descriptions in the environment of a specialist text. As units of understanding they have a prototype structure, may be subject to variation and can be explained as categories of understanding based on cognitive models. For the definition the history of a unit of understanding should also be taken into consideration, for which *biotechnology* may serve as a suitable example. On the whole, this chapter proves the author's competent knowledge as a subject specialist in intricate matters which cannot be elucidated by the linguist alone.

Chapter 4 discusses crucial problems of "Univocity And Polysemy". The author formulates two propositions (p. 132):

(1)   When a new phenomenon is discovered which can be clearly delineated and defined there is at first a natural development towards univocity (…). The search for a better and more profound understanding of the unit of understanding and the possible meaning extension mechanisms of language (like metaphorising) are likely to lead to polysemisation of the term which was initially assigned to the unit of understanding.

(2)   Univocity as the ideal for unambiguous and therefore efficient reasoning and communication is unrealistic for prototypically structured categories (…). This has at least two aspects to it. Firstly, polysemy is functional in specialised language. It is the consequence of a search for more and more profound knowledge and understanding. (…) Secondly, (near-)synonymy develops for reasons of accuracy in expression which is linked to the perspective taken by the sender of the message.

The author quotes examples which illustrate that polysemy *(cloning)* and synonymy *(Southern blotting/Southern transfer/Southern hybridisation)* (p. 151) are functional.

Chapter 5 on "The Impact of Metaphorical Models on Categorisation and Naming", actually the pivotal part of Temmerman's book, provides most important insights and substantial data relating to the productive capacity of metaphor in term creation in a particular textual environment and the specialist's understanding. As known from semasiology and semantics of the 1960s (S. Ullmann, M. Halle), metaphors have a prominent function in reflecting the physical and spiritual world by the human mind; they are a panchronic linguistic process, operate at all times in all languages and cultures, and are a universal semantic feature. Temmerman gives a brief overview of the theories of metaphors since the 1980s, with emphasis on G. Lakoff and M. Johnson (1980), and then treats the attitude of traditional Terminology towards metaphor, which, on the whole, is not encouraging in the light of term formation which follows the cognitive principle of analogy. The author, however, applies Lakoff's "Metaphoric Idealised Cognitive Model" (m-ICM) to a number of prominent genetic terms. These include metaphorical namings for information like *GENETIC MATERIAL (DNA) IS A LANGUAGE; THE TOTALITY OF THE GENETIC MATERIAL (THE GENOME) OF AN ORGANISM IS AN ATLAS OF MAPS; GENETIC MATERIAL (DNA) IS THE SOFTWARE OF THE CELL; GENETIC MATERIAL IS A FILM-TAPE* (p. 165f.). These sub-domains of the Idealised Cognitive Model are elaborated individually and substantiated by relevant text quotations which even yield more metaphorical expressions. These share semantic similarities in a particular cognitive field, e.g. *the messenger RNA … carries the message telling that … deciphering the DNA code … written in 'words' of three letters* (p. 186), *'codons' … that code for amino acid* (p. 189). Moreover, metaphorical meanings in the language of genetics have been attached to the (linguistic) terms *transcription, translation, duplication/duplicate; proofreading mechanism* (pp. 190-192). Other metaphors have been drawn from the vocabulary of geography: *a detailed map of human DNA; genetic linkage maps; genetic distances; contig maps/contigs* (i.e. "a set of contiguous overlapping cloned fragments that have been positioned relative to one another", p. 198); *gene hunters*, etc. A typical metaphorical sentence reads: "Like any explorer, *gene hunters* survey as they go, compiling *'maps'* of chromosomes on which they record easily identifiable *landmarks* and the locations of specific genes" (p. 199). (This source may be a textbook or a systematic introduction to Genetic Analysis.)

Temmerman aptly distinguishes between "creative and didactic metaphor" and draws parallels with their use in researchers' articles where "neolexicalisations are the result of conscious or subconscious analogical thinking", specialist manuals which are aimed at users who are not completely lay persons, but have a certain scientific background, and finally, popularising texts intended for a non-specialist readership of interested lay persons. Thus, the preference of certain metaphors is also influenced by the text type and the level

of specialism in discourse. This most diversified chapter concludes with a brief history of "Technological factors (which) influence the development of the life sciences" (p. 214ff.), such as "Computation and the Genome Project" and the "Sequencing technique". The author considers possibilities of metaphorical models to be applied in terminography.

Chapter 6, "Towards New Ways of Terminology Description", in fact draws general conclusions from the selected corpus analysis provided by the previous chapter.

As an appendix, the comprehensive "Bibliography" comprises a section on "Dictionaries and Encyclopedias", another one on "References in the Field of Biotechnology" and an independent part on "Theoretical References" to linguistics. The book concludes with a combined subject and name index.

The monograph is logically structured in the elaboration of arguments about the previous state of the art in Terminology and the author's alternative procedure in the following chapters. Temmerman is closely familiar with the principles of academic writing as recommended by American editorial boards of learned journals for anglophone publications by non-native international contributors. As she is fully aware of the writer's responsibility for the text, she consistently uses advance organisers in each chapter or subchapter, and also intermediate summaries. Special mention deserve the numerous figures and tables which provide a well-balanced synopsis of theoretical arguments, semantic features of terms or of concepts under discussion, or of facts in a particular field of the life sciences. This kind of condensed information in a clear and convincing layout serves as welcome signposts for the reader, and the tables enhance the book's intelligibility and readability.

Another feature of Temmerman's way of academic writing is her constant repetitions of key arguments. There are many cross-references between individual chapters to support thematic coherence, and helpful explanations of previously introduced abbreviations, which are resumed in later chapters and explained in parentheses. It is clearly for this didactic purpose that the author has incorporated elements of "positive" (favourable, supportive) redundancy, a textual and stylistic approach which is fully justified. Moreover, every chapter is headed by a motto which consists of one or more quotations from a literary or an academic writer and has a more or less direct bearing on the content of the following text. As analysts of academic writing (such as John Swales and Tony Dudley-Evans) have shown, references to literary texts are not uncommon in specialist texts on subjects of science and technology, especially for the purpose of didactisation or popularisation. And here again, the principle of analogy plays an important part.

Academic writing is also known for making use of ready-made Latin expressions which may be terms, clichés or stereotyped phrases. *Pars pro toto* is definitely a linguistic term and designates the semantic (and cognitive) relation "the part for the whole". On the other hand, it names the well-known figure of speech, *synecdoche*. The reverse semantic relation is also possible; but by no

means is it designated *toto pro pars* as Temmerman wrongly says (p. 145). In Latin, as a highly inflectional language, the correct formulation can only be *totum pro parte.*

In conclusion, one wonders why the author in her criticism of traditional Terminology has not mentioned the unstable status of nomenclatures and the total absence of phraseology. Both lexical inventories are integrated parts of specialist vocabularies. Items of nomenclature in the Viennese branch of Terminology only count as a special case of terms; they designate univocal items in a scientific classification or taxonomy. Temmerman refers to nomenclature only in a footnote on p. 127 relating to Juan C. Sager (1990).

On the other hand, phraseological units as terms do not play any part in the trend-setting Viennese Terminology. But they are a key issue for the Danish terminologist Heribert Picht (cf. his publications since 1989 and 1996).

Another moot point in Terminology and its standardisation activities, besides the disapproval of metaphors as terms, are the recommendations to avoid proper names as term constituents. The argument is that proper names do not express intrinsic features which are a clue to the meaning of a term, but rather have relational features (by referring to an inventor or discoverer). The examples of *Southern blotting, Southern transfer* and *Southern hybridisation* in genetics and even the extensions by 'wrong' name analogy, *Northern blotting, Western blotting, South-western blotting, North-western blotting,* quoted by Temmerman (p. 82, 151) indicate that proper names in term formation or the creation of units of understanding have a considerable explanatory power, last but not least for the history of such a designation.

To sum up, this new approach to terminology description offered by the author breaks new ground and is a thought-provoking alternative in domains of the social and human sciences where terms are evasive to any clear-cut notional systems of hyperonyms and hyponyms.

Rosemarie Gläser
Am Staffelstein 6
D-01328 Dresden
Germany

L.J. van der Veen et S. Bodinga-bwa-Bodinga. *Gedandedi sa Geviya/ Dictionnaire Geviya–Français*. 2002, 569 pp. ISBN 90-429-1105-0 (Peeters Louvain)/ISBN 2-87723-621-8 (Peeters France). Louvain/Paris/Sterling, Virginia: Peeters. Price: €60.

Introduction

La parution aux Éditions Peeters de l'ouvrage de référence lexicographique intitulé *Gedandedi sa Geviya/Dictionnaire Geviya–Français* (en abrégé DGF dans les reste de l'article) a été accueilli avec joie et fierté aussi bien par les locuteurs geviya[1] en particulier que par les Gabonais en général. Ce dictionnaire, résultat d'une longue et fructueuse collaboration scientifique entre Lolke van der Veen et Sébastien Bodinga-bwa-Bodinga, marque le début d'une nouvelle ère dans la rédaction des dictionnaires au Gabon. En effet, la majorité des ouvrages de référence lexicographique disponibles dans les langues gabonaises sont des dictionnaires bilingues rédigés par les missionnaires catholiques et protestants ou par les administrateurs coloniaux dans la période allant de 1800 à 1960, date de l'accession du Gabon à la souveraineté internationale (Nyangone Assam et Mavoungou 2000, Mavoungou 2001).

Notre objectif est de mieux cerner l'intérêt du DGF et la richesse de son contenu, en retraçant d'abord le travail prodigieux accompli par Van der Veen et Bodinga-bwa-Bodinga, et en relatant ensuite la grande amitié entre ces deux chercheurs.

Dans le présent article, nous avons également adopté implicitement la méthodologie de Wiegand et Kučera (1981, 1982) en ce qui concerne l'élaboration d'un dictionnaire. Ladite méthodologie repose sur les centres d'intérêts suivants: les auteurs et leurs objectifs, la base du dictionnaire, la macro- et microstructure du DGF, et le traitement lexicographique des termes de spécialité.

1.   Les auteurs et leurs objectifs

Avant d'aborder le sujet même du DGF, quelques mots sur les auteurs principaux, ainsi que sur la maison d'édition. Le DGF est le fruit du travail prodigieux accompli par deux chercheurs: Lolke van der Veen et Sébastien Bodinga-bwa-Bodinga. C'est au terme de plusieurs années de labeur assidu que Bodinga-bwa-Bodinga — ancien sécrétaire du gouverneur de Mouila (Gabon) à la retraite et lexicographe autodidacte — est allé voir en 1987 Van der Veen pour lui confier un document dactylographié de 1 287 pages traitant du lexique geviya. Van der Veen, linguiste de formation, va au prix d'un travail d'analyse et de vérification de plusieurs années avec un assistant de langue donner au manuscrit la forme linguistique et lexicographique du dictionnaire qui verra finalement le jour en 2002. Il est évident que derrière ces deux noms, il y a toute une équipe, et nous espérons ne pas leur faire trop d'injustice en ne mentionnant pas les noms de tous leurs collaborateurs.

Les Éditions Peeters, qui ont publié le DGF, sont une autorité reconnue dans le domaine de la linguistique africaine, de la lexicographie et de la lexicologie. Dans le cadre de la collection "Langues et Littérature de l'Afrique Noire"[2], les Éditions Peeters ont publié aussi bien des ouvrages proprement linguistiques ou lexicographiques (dictionnaires, grammaires, recherches théoriques ou appliquées, etc.) que littéraires (recueils de contes, proverbes, devinettes, etc.). La tradition de cette maison est déjà une garantie de qualité. Toutefois, ces faits n'ont pas mis le dictionnaire à l'abri des critiques. En Europe comme ailleurs on a critiqué sa publication, compte-tenu du nombre restreint des locuteurs natifs du geviya. Les plus extrémistes ont dit du geviya qu'il n'était plus parlé que par une dizaine de locuteurs, tandis que selon d'autres on pouvait encore dénombrer une cinquantaine de locuteurs natifs. L'on a également fait un certain nombre de réserves quant au système de transcription (alphabet) utilisé par les auteurs. Nous reviendrons plus en détail sur ces points.

Parmi les objectifs des auteurs, citons que d'un point de vue strictement scientifique, le DGF a pour objectif de conserver une partie du patrimoine ethnolinguistique bantou de l'Afrique Centrale et de revaloriser une langue en principe condamnée à disparaître. Ces différents objectifs donnent également une indication du public cible de ce dictionnaire. Il s'adresse avant tout aux jeunes Eviya qui font l'apprentissage d'une langue qui souffre aujourd'hui cruellement d'un très faible nombre de bons locuteurs, mais aussi à des locuteurs adultes Eviya soucieux d'approfondir leur connaissance de la langue geviya. Le choix de ce public est aussi un indice de la vocation pédagogique de ce dictionnaire d'apprentissage. Sa vocation scientifique consiste en ce qu'il pourrait susciter l'intérêt d'un public assez large composé d'universitaires et de chercheurs, ainsi que de toute autre personne soucieuse de transmettre aux générations futures les traditions et cultures orales du Gabon et partant celles du continent africain.

2.   La base du dictionnaire

La base d'un dictionnaire comprend l'ensemble des sources qui ont servi à sa planification, son élaboration et sa publication. Ces sources sont de trois ordres, à savoir: les sources primaires, secondaires et tertiaires. La constitution des sources primaires ou le corpus lexicographique d'un dictionnaire consiste à l'enregistrement d'une grande variété d'occurrences orales en provenance d'autant de domaines de recherche différents que possible. Le corpus lexicographique qui avait alimenté au début le DGF était essentiellement constitué de textes oraux ou occurrences orales du geviya assemblés sur plusieurs années à partir de diverses sources. Ces textes oraux vont servir à la mise en place d'une base de données textuelles aujourd'hui accessible sur Internet. Le DGF étant le premier dictionnaire de la langue geviya, les dictionnaires monolingues ou bilingues du français publiés pendant la période de son élaboration sont les sources secondaires[3] de l'ouvrage. Les sources tertiaires incluent toutes les monogra-

phies, communications et grammaires linguistiques utilisées pour la constitution de la base du dictionnaire. Elles portent d'une part sur diverses publications à caractère ethnolinguistique comme les expressions proverbiales, l'histoire des Eviya et les lexiques spécialisés de la flore et des affections pathologiques en geviya, et d'autre part sur un certain nombre de descriptions de la phonologie segmentale et tonale, de la morphologie et de la syntaxe du geviya. À ce propos, nous pouvons citer les contributions suivantes: Raponda-Walker et Sillans (1961), Raponda-Walker (1967), Bodinga-bwa-Bodinga (1969), Blanchon (1988) et Van der Veen (1991, 1992, 1999).

3.   La macrostructure du DGF

La nomenclature du DGF présente la langue geviya telle qu'elle est utilisée par les communautés linguistiques vivant entre Fougamou et Sindara dans un village situé sur la rive droite de la Ngounié (affluent important du fleuve Ogooué qui a d'ailleurs donné son nom à la province toute entière) à proximité des chutes de Samba-Nagoshi. Les éléments macrostructurels du dictionnaire comprennent des mots utilisés dans le discours quotidien, les noms géographiques (ou toponymes), les noms de plantes et d'animaux, les anthroponymes et les expressions idiomatiques (idiomes et proverbes). C'est en vérité un ensemble représentatif des occurrences lexicales du geviya. Il n'y a aucun doute que les lexicographes ont voulu que leur dictionnaire soit un "véritable réservoir de connaissances" (McArthur 1986).

La macrostructure du DGF contient environ 6 200 articles classés par ordre alphabétique sous la lettre initiale du thème de chaque lemme. Les unités de traitement sont précédées d'une part de leurs préfixes[4] respectifs entre parenthèses, et d'autre part sont suivies d'une indication de leurs schèmes tonals respectifs accompagnée d'une spécification de la catégorie grammaticale (partie du discours) ainsi que du numéro de classe auxquels appartiennent les lemmes. On ne peut prétendre à une analyse de la macrostructure d'un dictionnaire d'une langue africaine sans faire état des problèmes liés à la tradition du mot versus tradition du thème. Comme nous l'avons souligné plus haut, le DGF suit largement la tradition du thème. Les raisons avancées par les auteurs pour le choix d'un tel système sont assez classiques. La tradition du thème est souvent présentée comme étant plus scientifique et plus adaptée aux besoins des usagers que la tradition du mot, et ceci pour deux raisons essentielles: cette méthode permet de regrouper les items lexicaux par affinités sémantiques, tout en évitant certains secteurs de la nomenclature d'être surchargés pendant que d'autres sont presque vides. Cela est particulièrement vrai pour, entre autres, certaines catégories grammaticales telles que les verbes, les adjectifs et les possessifs. Les deux exemples textuels suivants extraits du DGF représentant le traitement lexicographique des substantifs et des verbes illustrent bien l'organisation interne des articles du dictionnaire selon la tradition du thème, de la méthode ou du système choisi par les auteurs.

> (mo-)**gadi** H *n 1/2* •**1**°épouse. *Tsìkà mógádì ámὲ wá mósòngá nà lὲlὲ.* Laisse ma première épouse tranquille. *Mògàdì á ngà.* Femme d'autrui. Voir mo-gɛtɔ. •**2**° (fig.) femelle (d'un animal) *Mògàdì á ngòèà.* La femelle du potamochère. *Mògàdì á ὲdɔmbὲ.* Brébis. *Mògàdì á gèvàlàngò.* Cane.

Exemple textuel 1: article (mo-)**gadi** (DGF 2002: 136-137)

> (e-)**kokega** H *v tr* arrimer, bien fixer, bien entasser. *Dyàkó-kègè éómá ὲtsɔ è ndὲ gó gègàrà. Entasser bien toutes les affaires qui sont dans la cantine.*

Exemple textuel 2: article (e-)**kokega** (DGF 2002: 213)

Comme il apparaît clairement, les lemmes sont présentés avec à la fois une indication tonale (ton haut en abrégé H) et des parties du discours (nom et verbe transitif en abrégé *n.* et *v. tr.*). Du point de vue de la lemmatisation, toutes les catégories grammaticales confondues ont reçu le statut de lemme, à savoir: substantifs, adjectifs, déterminants, verbes, adverbes, locutions adverbiales, idéophones, onomatopées, interjections, conjonctions, prépositions, etc. D'un point de vue traditionnel, les dictionnaires traitent séparément des noms communs et des noms propres. *Le Petit Larousse Illustré* (1998) est l'exemple parfait de cette pratique lexicographique. La première section de ce dictionnaire monolingue encyclopédique du français inclut les items lexicaux du vocabulaire standard du français, tandis que la deuxième section donne un exposé des noms de personnes, de lieux, de bâtiments, et des oeuvres d'art célèbres à travers le monde. De même, les rédacteurs du DGF ont choisi de ne point séparer les anthroponymes et les noms communs. Tous apparaissent dans la nomenclature à leur propre place alphabétique. Les exemples textuels suivants illustrent le traitement lexicographique des noms propres (anthroponymes et toponymes) dans le DGF.

> (mo-)**dabu** B (*n 3*) nom propre donné aux femmes.

Exemple textuel 3: article (mo-)**dabu** (DGF 2002: 93)

> (mo-)**gesagano** B (*n 3*) toponyme, nom d'un des anciens villages eviya sur la rive droite de la Ngounié

Exemple textuel 4: article (mo-)**gesagano** (DGF 2002: 147)

Quant au système d'écriture utilisé dans le DGF, les rédacteurs ont transcrit les faits de langue geviya en utilisant les symboles de l'Alphabet Scientifique des

Langues du Gabon⁵ (ASG), basé à la fois sur les symboles de l'Alphabet International de Phonétique (API, révisé en 1993 et actualisé en 1996), et l'Alphabet Africa publié par l'International African Institute (IAI) aussitôt après sa création en 1926. La valeur des symboles ou caractères spéciaux de l'API ou de l'Alphabet Africa bien qu'évidente pour les spécialistes, ne l'est pas pour le grand public. Leur utilisation dans un dictionnaire doit par conséquent répondre aux besoins et aux aptitudes de référence des utilisateurs-cibles. Dans les notes introductives du DGF (2002: 4), les auteurs soulignent qu'"on peut estimer le nombre de bon locuteurs à une quarantaine seulement. Ce chiffre extrêmement bas indique sans ambiguïté que le geviya est une langue dont la disparition est sur le point de s'achever". L'utilisation des symboles de l'ASG pour la transcription écrite des occurrences orales du geviya se justifie par la sérieuse menace de sa disparition. Pour éviter que la langue ne disparaisse sans laisser de traces, il fallait utiliser des caractères spéciaux capables de reproduire correctement le système phonétique du geviya. Dans l'analyse critique et l'évaluation des dictionnaires disponibles dans les langues gabonaises, il a été démontré que le point faible principal de ces travaux lexicographiques était l'absence des tons dans la transcription écrite des occurrences orales (cf. Nyangone Assam et Mavoungou 2000). De même, le DGF, en tenant compte du phénomène des tons, rend justice à cette composante de base de l'analyse phonologique de la langue.

4.   La microstructure du DGF

Au niveau microstructurel, les articles varient en longueur entre explications brèves d'unités lexicales et traitement plus complet des items lexicaux. La microstructure inclut plusieurs données. L'indication des parties du discours est immédiatement et systématiquement suivie par une des traductions du lemme. Pour illustrer l'emploi d'un lemme spécifique, les rédacteurs utilisent de nombreux exemples en italique (chaque exemple en geviya est immédiatement suivi de son correspondant en français). Des renvois guident le lecteur vers des articles sémantiquement liés, tel l'article du lemme (m-/my-)**opa**:

> (m-/my-)**opa** BH *n 3/4* •**1°** lutte. *Màkídàndò nà Mògèkè wà kógùrànà môpà*. MAKIDANDO et MOGEKE *luttent corps à corps*. •**2°** (fig.) talisman de lutte. Voir e- nimba.

Exemple textuel 5: article (m-/my-)**opa** (DGF 2002: 350)

Dans l'article ci-dessus, le lemme apparaît en caractères gras tandis que les exemples d'emploi sont en italiques. À la suite du traitement des sens 1 et 2 du lemme, l'usager a droit à une entrée médiostructurelle ou un renvoi: *Voir e-nimba*. Les entrées *Voir* et *e-nimba* constituent deux segments différents de texte. Le premier est désigné sous le nom de marqueur de référence, car son rôle est

d'indiquer la relation de référence. Le deuxième segment de texte est connu sous la dénomination d'entrée marquant l'adresse de référence. Par conséquent, dans l'article ci-dessus le signe-lemme (m-/my-)**opa** est la position de référence, et l'entrée *e-nimba* est l'adresse de référence.

5.   Le traitement lexicographique des termes de spécialité

L'examen macro- et microstruturel du DGF permet également de mettre en évidence une autre caractéristique intéressante de ce dictionnaire, notamment du traitement des termes de spécialité qui appartiennent à la flore et aux affections pathologiques, comme le démontre le traitement de l'article du lemme (o-)**goma**:

> (o-)**goma** B *n 11/10a* variété de grand arbre, irvingiacée, Eveuss à petites feuilles, Klainedoxa gabonensis Pierre, var. Microphylla Pellegr. Bois très dur. Usage méd.: furonculose (graines en lavements ou frictions), douleurs rhumatismales précédant le pian (écorce). *Mòn ɔ̀ngɔ̀ wà géɔ̀dɛ̀dí sá mótè ó vègú tè ògómà! Quelle dureté que celle de l'arbre appelé "ogoma"!*

Exemple textuel 6: article (o-)**goma** (DGF 2002: 156-157)

Les rédacteurs du DGF commencent le traitement lexicographique par une indication du signe-lemme suivie des entrées suivantes: indication tonale (B), partie du discours (*n*), et indication du genre (*11/10a*) auquel le lemme appartient. À la suite de ces différentes entrées, l'usager a droit à une paraphrase de la signification du lemme (o-)**goma**, suivie de plusieurs traductions. Les entrées *Klainedoxa gabonensis* et *Microphylla* sont des données techniques ou sophistiquées: elles représentent les noms scientifiques de l'arbre décrit. Ensuite, dans le cas spécifique de l'article introduit par le marqueur structurel *Usage méd.* (usage médicinal), une explication est donnée de la description et des utilisations de l'arbre en question. Ces utilisations mettent généralement en exergue quelques aspects culturels (vision du monde, savoir médico-magique, etc.) de la vie des Eviya. L'utilisateur peut obtenir plusieurs d'informations utiles de ces entrées. Par rapport aux sources de ces données portant sur la pharmacopée traditionnelle en milieu Eviya, le lecteur lira avec profit les travaux de Raponda-Walker et Sillans (1961), Bodinga-bwa-Bodinga et Van der Veen (1993, 1995).

En résumé, le traitement des termes de spécialité a été fait d'une manière tout à fait satisfaisante en ce qui concerne le contexte culturel, puisque l'utilisateur a accès aux données lexicographiques concernant la flore et la faune, ainsi que l'habitat naturel des populations Eviya.

6.   Conclusion

Dans l'analyse critique du DGF ci-dessus, nous avons fait ressortir un certain nombre de réserves quant à ce dictionnaire. Ainsi avons-nous critiqué son public cible ainsi que la place accordée aux caractères spéciaux dans le DGF. Bien que plausibles, ces reproches sont de moindre importance, si on tient compte de la visée scientifique et pédagogique du DGF. En effet, devant la sérieuse menace de disparition du geviya il était impératif de fixer la langue dans un dictionnaire en utilisant un système d'écriture qui soit à même de reproduire correctement sa phonologie segmentale et suprasegmentale en particulier.

L'examen des pré-textes du DGF fournissent aux usagers assez d'informations concernant les sources primaires, secondaires et tertiaires de l'ouvrage. D'un point de vue macrostructurel, il ne fait aucun doute que les lexicographes aient utilisé un volumineux corpus lexicographique de la langue, ce qui fait du dictionnaire une source pertinente et un véritable réservoir de connaissances. Les éléments macrostructurels ont été sélectionnés parmi divers champs de connaissance (la tradition orale, la religion, la sylviculture, et la pharmacopée traditionnelle entre d'autres). La microstructure contient plusieurs entrées utiles: traductions, exemples d'emploi, renvois, etc. Les items lexicaux de spécialité sont traités d'une manière satisfaisante en ce sens qu'ils contiennent la vision du monde propre aux populations Eviya. Les descriptions médicinales et médico-magiques des plantes sont très importantes car elles permettent aux usagers du dictionnaire de prendre conscience de la richesse de la culture Eviya.

Notes

1.   Le geviya est la langue parlée par les Eviya, localisés dans la province de la Ngounié (Gabon). D'un point de vue linguistique, le geviya (B30 Guthrie 1948, 1953, 1969–1971) fait partie du groupe Tsogo (B30) de la branche linguistique bantoue qui inclut le getsogo, le gepinzi(pinzi), l'okande, le gehimba(ka), le gevove et l'ebongwe. Des recherches linguistiques ont montré qu'au sein de ce groupe, le geviya présente le plus d'affinités avec le getsogo et le gevove. En outre, d'autres études ont révélé la forte influence que le geviya a subi et continue de subir de la part du gisira (B40), langue parlée dans le voisinage immédiat des Eviya.

2.   Cette collection est une iniative de l'Institut National des Langues et Civilisations Orientales (INALCO).

3.   Tous les dictionnaires consultés durant la phase de rédaction (Wiegand et Kučera 1981: 100ff).

4.   "Chaque préfixe nominal possède en principe une variante morphologique spécifique devant un thème à initiale vocalique. On trouve devant un thème nominal à initiale vocalique la forme pleine du préfixe, lorsque ce thème est dérivé d'un verbe. Ceci est illustré par les exemples suivants: Le marqueur nominal de classe 11 est /o-/. Ex. /o-lémè/ 'langue' (en tant qu'organe). Devant une voyelle, ce préfixe se semivocalise en w-. Ex. /o-áá/ 'ongle' devient

[wàà]. Mais on trouve aussi /o-átédí/ 'façon de s'habiller' se réalisant [òàtèdì] sans semivo-calisation" (DGF 2002: 18-19).

5. L'Alphabet Scientifique des langues du Gabon (ASG, voir Carpentier de Changy et Voltz 1990) fut créé lors du Séminaire des experts sur l'établissement d'un Alphabet Scientifique des langues du Gabon organisé par le LUTO (Laboratoire Universitaire de la Tradition Orale) à Libreville. Ce séminaire s'est déroulé du 20 au 24 février 1989, avec la participation des délégations en provenance de trois pays: le Gabon, la Belgique (représentée par l'Université Libre de Bruxelles ainsi que par le Musée Royal de Tervuren), la France (représentée par l'Université Lumière Lyon 2) ainsi que par quelques experts en provenance de deux institutions internationales, le CICIBA (Centre International des Civilisations Bantu) et l'ACCT (aujourd'hui Agence Internationale de la Francophonie).

Références

Dictionnaires et encyclopédies

Petit Larousse Illustré. 1998. Paris: Larousse.

Raponda-Walker, A. et R. Sillans. 1961. *Les plantes utiles du Gabon: essai d'inventaire et de concordance des noms vernaculaires et scientifiques des plantes spontanées et introduites du Gabon*. Paris: Lechevalier.

Autre sources

Alphabet International de Phonétique (API, révisé en 1993 et actualisé en 1996). 2001. Publié par le *Journal of the International Phonetic Association*. Cambridge: Cambridge University Press.

Blanchon, J.A. 1988. Une langue mixte en voie de disparition: le geviya. *Pholia, Revue du Laboratoire de Phonétique et Linguistique Africaine* 3: 53-69.

Bodinga-bwa-Bodinga, S. 1969. *Traditions orales de la race eviya*. Paris: T.M.T.

Bodinga-bwa-Bodinga, S. et L.J. van der Veen. 1993. Plantes utiles des Evia. *Pholia, Revue du Laboratoire de Phonétique et Linguistique Africaine* 8: 26-65.

Bodinga-bwa-Bodinga, S. et L.J. van der Veen. 1995. *Les proverbes evia et le monde animal, la communauté evia à travers ses expressions proverbiales*. Paris: Harmattan.

Carpentier de Changy, H. et M. Voltz. 1990. Alphabet Scientifique des Langues du Gabon (ASG) (Scientific Alphabet of Gabonese Languages, SAG). *Revue Gabonaise des Sciences de l'Homme* 2: 113-115.

Guthrie, Malcolm. 1948. *The Classification of the Bantu Languages*. London: Oxford University Press.

Guthrie, Malcolm. 1953. *The Bantu Languages of Western Equatorial Africa*. London: Oxford University Press.

Guthrie, Malcolm. 1969–1971. *Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages*. 5 tomes. Farnborough: Gregg International.

Mavoungou, P.A. 2001. Macro- and Microstructural Issues in Mazuna Lexicography. *Lexikos* 11: 122-138.

McArthur, T. 1986. *Worlds of Reference*. Cambridge: Cambridge University Press.

Nyangone Assam, B. et P.A. Mavoungou. 2000. Lexicography in Gabon: A Survey. *Lexikos* 10: 252-274.

Raponda-Walker, A. 1967. *Contes gabonais*. Paris: Présence Africaine.

Van der Veen, L.J. 1991. *Étude comparée des parlers du groupe Okani — B30 (Gabon)*. Thèse de doctorat (nouveau régime). Université Lumière-Lyon 2.

Van der Veen, L.J. 1992. Le système tonal du ge-via (Gabon). *Journal of West African Languages* 22(2): 17-41.

Van der Veen, L.J. 1999. *Les Bantous eviya (Gabon-B30): langue et société traditionnelle*. Notes de synthèse en vue de l'obtention de l'Habillitation à Diriger des Recherches en Sciences du Langage, soutenue le 28 janvier 1999 à l'Université Lumière-Lyon 2.

Wiegand, H.E. et A. Kučera. 1981. Brockhaus-Wahrig: Deutsches Wörterbuch auf dem Prüfstand der praktischen Lexikologie. I. Teil: 1. Band (A–BT); 2. Band (BV–FZ). *Kopenhagener Beiträge zur germanistischen Linguistik* 18: 94-217.

Wiegand, H.E. et A. Kučera. 1982. Brockhaus-Wahrig: Deutsches Wörterbuch auf dem Prüfstand der praktischen Lexikologie. II. Teil: 1. Band (A–BT); 2. Band (BU–FZ); 3. Band (G–JZ). Wiegand, H.E. 1982. *Studien zur neuhochdeutchschen Lexikographie II. Germanistische Linguistik* 3–6/80: 285-373.

P.A. Mavoungou
Département des Sciences du Langage
Université Omar Bongo
Libreville
République Gabonaise
et
Département d'Afrikaans et du Néerlandais
Université de Stellenbosch
Stellenbosch
République d'Afrique du Sud
(moudika2@yahoo.fr)

F.I.J. van Rensburg (Redakteur). *Afrikaans — lewende taal van miljoene.* 2004, xv + 190 pp. ISBN 0 627 02579 X. Pretoria: Van Schaik. Prys: R89.95.

Hierdie publikasie is die resultaat van 'n projek begin deur die Vrydaggroep, 'n aantal kenners wat hulle landswyd beywer vir studies oor Afrikaans binne die veeltalige konteks van Suid-Afrika. Die redakteur was die projekleier. Die titel is afkomstig uit 'n hoofartikel met die Afrikaanse opskrif "Reg vir die toekoms" wat twee dae ná die verkiesing van 29 April 1994 in die *London Times* verskyn het. Daarin word Afrikaans beskryf as die "lewende taal van miljoene".

Die boek sluit aan by 'n lang reeks soortgelyke bundels oor Afrikaans wat oor die afgelope 45 jaar verskyn het: vanaf die meer bewonderende publikasies van Du Buisson (1959), Labuscagne (1959) en Kok (1974) tot die meer realistiese versamelings van Prinsloo en Van Rensburg (1984), Du Plessis en Du Plessis (1987), Webb (1992) en Giliomee en Schlemmer (2001). Die stukke in hierdie bundel neig soms na die bewonderende en soms na die meer realistiese.

Die bundel wat bestaan uit 'n twintigtal artikels geskryf deur 23 kenners op hul onderskeie gebiede, word voorafgegaan deur 'n verantwoording deur die redakteur F.I.J. van Rensburg: "Die behoefte aan 'n publikasie soos dié het voortgekom vanuit 'n besef van die omvang van die geknelde situasie waarin Afrikaans hom op die oomblik bevind as gevolg van die drastiese verkleining van sy gebruiksfeer" (bl. ix).

Die eerste ongeveer twee-derdes van die boek bestaan uit veertien artikels wat 'n oorsig gee van die stand van Afrikaans op verskillende gebiede: die uitsaaiwese, die onderwys, die pers, die regsdenke, die teologie, die natuurwetenskappe en tegniek, die reklamewese, e.s.m. Sommige van hierdie artikels wat Afrikaans binne die wyer aktualiteit van die teenswoordige Suid-Afrika of in sy huidige verhouding met die ander amptelike Suid-Afrikaanse tale plaas, is meer feitelik as ander. Die laaste derde van die boek bevat ses artikels wat "feite versus persepsies in verband met Afrikaans" behandel. Hierdie artikels wat sekere foutiewe aannames betreffende Afrikaans met die werklike feite probeer weerlê, is van die insiggewendste in die boek.

Alhoewel die boek basies vir taalliefhebbers en taaldeskundiges bedoel is, is dit ook vir taalywaraars en taalpolitici van belang. Indirek sou Afrikaanse leksikograwe en terminoloë en die uitgewers van hul woordeboeke heelwat inligting uit dié artikels kon haal ten opsigte waarvan hulle hul kan oriënteer in die veranderde situasie waarin Afrikaans hom tans bevind. Die gebruik van Afrikaans is oor die afgelope tien jaar op alle vlakke van die openbare terrein afgeskaal. Wat kan die bydrae en rol van die Afrikaanse leksikograaf en terminoloog en die uitgewers van hul woordeboeke in hierdie veranderde omgewing wees? Watter positiewe en negatiewe faktore konfronteer hulle in hierdie omgewing? En in hoeverre kan hulle, deur ook 'n bydrae aan die ander tale van Suid-Afrika en selfs Afrika te lewer, die voortbestaan en uitbreiding van Afrikaans op leksikografiese en terminologiese gebied verseker? Dit is vrae wat die Buro van die Woordeboek van die Afrikaanse Taal hom reeds vyftien jaar

gelede ten doel gestel het om te beantwoord (sien die artikel van Frikkie Lombard, bll. 30-34), en die antwoorde wat daaruit voortgekom het, te gebruik om homself so positief moontlik te posisioneer om by te dra tot Afrikaans se voortbestaan en uitbreiding.

'n Mens hoop dat die aktualiteit waarin Afrikaans hom bevind en wat deur die lees van hierdie artikels tot stand kom, as 'n aanmoediging eerder as 'n ontmoediging sal dien om die Afrikaanse leksikografie en terminologie te bly ontwikkel binne die groter konteks van die ander Suid-Afrikaanse tale.

Bibliografie

Du Buisson, M.S. (Red.). 1959. *Die Wonder van Afrikaans: Bydraes oor die ontstaan en groei van Afrikaans tot volwaardige wêreldtaal*. Johannesburg: Voortrekkerpers.

Du Plessis, Hans en Theo du Plessis (Reds.). 1987. *Afrikaans en taalpolitiek: 15 opstelle*. Pretoria: HAUM Opvoedkundige Uitgewery.

Giliomee, Hermann en Lawrence Schlemmer (Reds.). 2001. *Kruispad: Die toekoms van Afrikaans as openbare taal*. Kaapstad: Tafelberg.

Kok, B. e.a. (Reds.). 1974. *Afrikaans — ons pêrel van groot waarde*. Johannesburg: Federasie van Afrikaanse Kultuurvereniginge.

Labuscagne, G.J. (Red.). 1959. *Feesbundel: Aangebied ter geleentheid van die viering van "die wonder van Afrikaans"*. Johannesburg: Dagbreek Boekhandel.

Prinsloo, K.P. en M.C.J. van Rensburg (Reds.). 1984. *Afrikaans: Stand, taak, toekoms*. Pretoria: HAUM Opvoedkundige Uitgewery.

Webb, V.N. (Red.). 1992. *Afrikaans ná apartheid*. Pretoria: J.L. van Schaik.

J.C.M.D. du Plessis
Buro van die Woordeboek van die Afrikaanse Taal
Stellenbosch
Republiek van Suid-Afrika
(wat@sun.ac.za)

Hessel Visser. *Naro Dictionary: Naro–English, English–Naro.* Fourth edition. 2001, viii + 240 pp. ISBN 99912-938-5-x. D'kar: Kuru Development Trust.

Although this dictionary was not published by a well-established publishing house, its quality and usefulness rival those of seasoned publishers in the field of Khoesan studies, such as the dictionaries by Dickens (1994), Traill (1994) and Haacke (2002). This fourth edition of Visser's *Naro Dictionary* is a medium-sized dictionary of 240 pages, including 16 pages of appendices. The work is the result of more than 10 years of lexical compilation which resulted in several editions of the dictionary, each of which was an expanded version of the earlier one. Thus, from a miniature version, it has grown into a sizeable high-quality work, in spite of the compiler's modest observation in his introduction that it is a mere "working document" and that it is still a "preliminary publication".

The dictionary begins with initial information for the users, such as the status of the present dictionary in relation to the earlier versions, the structure of the entire dictionary, the representation of the various sounds and the marking of the prosodic features such as tone, nasalization and vowel depression. According to the compiler, the work on the dictionary was part of a broader project, the Naro Language Project, which included the preparation of primers, literacy materials, magazines, a Bible translation and other linguistic documents.

The dictionary itself is divided into two main parts. The Naro–English part has about 5 500 entries, arranged in alphabetical order. Each entry begins with a headword in bold, followed by the corresponding phonetic transcription, in IPA symbols, in square brackets and the tone pattern in parentheses. The English equivalents are given either as single words or as paraphrases. In some cases, the variant forms, dialectal variation, corresponding part of speech, example phrases and idiomatic use of words are also given. In the case of fauna and flora, the equivalent scientific terms are often provided. Also in the case of loan-words, the language of origin is usually specified.

The second part of the dictionary consists of about 9 000 entries in which the English headwords, in bold and in alphabetical order, are followed by the equivalent Naro words or phrases in the conventional Naro orthography, the IPA transcription in square brackets, and the tone patterns in parentheses. No other information is provided. According to the compiler, there are more entries in this part than in the Naro–English part, because of the fact that many of the Naro words have several English equivalents. Thus it became necessary to establish more English headwords for the fewer Naro lexical items.

The 16-page appendices contain lists of specific lexical items referring to animals, birds, plants, pans and numerals as well as proper names used for men and women.

From the above, it is clear that this publication meets most of the requirements of a good reference dictionary, containing essential features such as head-

words arranged in alphabetical order, the pronunciation pattern of each word, the relevant meanings, and examples of word usage or idiomatic use. Moreover, the addition of the English–Naro part is highly appropriate as it allows those looking for Naro words, particularly local non-speakers and foreign learners and researchers, to benefit from the dictionary. Apart from these aspects, the dictionary could be said to be unique in the following ways:

First, it will be helpful to both local people and international scholars. Naro speakers and other local people who are now used to the roman-based orthographic convention (after very justifiably rejecting the use of the "sticks" in the IPA transcription of the clicks) are able to find the English equivalents for the Naro words. On the other hand, foreign linguists and other scholars will be able to search for Naro words and learn their pronunciation through the international phonetic convention, which is also provided in the dictionary. The strictly alphabetical order of the spelling system, rather than the pronunciation modes, will make the dictionary more practically useful.

Second, the dictionary is very user-friendly because of the many useful explanations in the introduction, the systematic organization and the use of simple and clear definitions, with occasional word-for-word or literal translations. In fact, the compiler succeeded to a very large extent in dealing with phenomena which involve intercultural conceptualizations such as *nqõó* "world" or *dqo̱m̀* "worship".

Third, the compiler has made a commendable effort to be specific in defining problematic entries, such as fauna and flora. Where many other dictionaries would have been content just to mention "a kind of tree" or "a kind of bird", the compiler has taken the trouble to find the scientific name of the plant or animal or to give a detailed description of the object involved.

Fourth, the dictionary was prepared with considerable accuracy and thoroughness. It is clear that the compiler has taken much care in recording the relevant information for each of the entries. Moreover, his close contact and association with representative speakers of Naro enabled him to check each entry.

However, a work of this complexity can never be perfect, since there are always a number of unnoticed inconsistencies, oversights and other shortcomings which may only be detected by inquiring users. This dictionary is no exception. The following shortcomings are worth pointing out.

First, some of the orthographic conventions used to represent the sounds in Naro may be confusing to the dictionary users. For example, the sound spelt as *kg* is given the phonetic value [kx] (a voiceless velar affricate) in the phonetic transcription. However, this spelling is also used in Setswana, the lingua franca of Botswana, with the phonetic value of [qʰ] (aspirated voiceless uvular plosive). The dictionary users, who are used to the Setswana orthography would tend to pronounce it as in Setswana. Moreover, the sound spelt as *gh* is given the phonetic value of [g] (voiced velar plosive), while the sound spelt as *g* is given the phonetic value of [x] (voiceless velar fricative). Although this may

have been done to ensure that the respective sound is pronounced as in Setswana and Afrikaans (another major language in the region), the spelling of these two sounds would be highly confusing to those used to other conventions, particularly where graphemes like *gh* are expected to represent fricative sounds.

Second, as the compiler himself states, tones have been indicated in some cases but not in others. Moreover, tonal notation was "guessed" in several incidences. This reveals the compiler's modesty and frankness, but also indicates how some of the tonal representations in the dictionary may be inadequate or misleading. On the other hand, in some cases where tone is not marked, users would find it difficult to tell whether the absence of tone means mid-tone or uncertainty about the tone pattern. This is the case of *gataga* [xataxa] "like, likewise, …".

Third, there is some inconsistency in the indication of the origin of loan-words. Although the origin of such words are given in most cases, there are cases where they are not. These include words such as *tshibi* "Sunday" (from Setswana *tshipi*) or *thebe* "one hundredth of a pula" (from Setswana *thebe*). Moreover, although the origin of most loan-words is indicated, the intermediate languages through which these words have entered Naro have not been given. This is the case with words such as *baebele* "Bible" and *buka* "book", which although originating from English have entered Naro through the intermediary of Setswana, where they are also spelt the same way.

Fourth, although the compiler provides a very helpful sketch of the marking system of person, gender and number in Naro on page 238, this information would go unnoticed by casual users of the dictionary. Such information ought to have appeared as part of the introduction, as it is vital in orienting users to the basic morphology of the language. In fact, it should have been systematically and clearly written to indicate the processes of derivation, inflection and case marking in the language.

Fifth, although the parts of speech of the headwords have been indicated in most cases, they are absent in a number of the entries. Thus the definition of a word like *ábàgu-qʼoo* "how someone is packing his goods" does not reveal whether the headword is an adverb, an ideophone or a noun. Moreover, since the parts of speech are not indicated in the English–Naro part, it is difficult to know whether the English headword "tame" (*kgóbó*) is used in its verbal or adjectival form, or the word "taste" (*xám̀*) in its verbal or nominal form, unless cross-checking with the Naro–English part.

Finally, since one of the primary aims of the dictionary is to assist the Naro-speaking community and the other people of Botswana to become familiar with Naro words, it would have been more helpful to have the equivalents in Setswana as well, since this is the lingua franca of the country and the language in which many Naro speakers have both spoken and written competence. Most Naro speakers do not speak English. A trilingual dictionary in which Setswana is one of the languages would have been ideal. Of course, one is aware of the technical intricacies in preparing a trilingual dictionary, but the trouble would be worth it. Perhaps this could be considered as a future project.

In spite of the above shortcomings, Visser's dictionary is an important contribution not only to the existing dictionaries of the Khoesan languages spoken in Botswana, but also to the number of books on the Naro language, whose orthography was only established several years ago. The dictionary will be useful not only to the speakers of Naro, the primary target users, but also to the general Botswana public as well as a good range of international scholars such as linguists, ethnographers, historians, missionaries, tourists and other language enthusiasts. In fact, the dictionary provides a wealth of information about the Naro people, their culture and their world view. For example, one is fascinated by the many terms denoting directions like west, east, north and south. This is presumably because of the importance of spatial directions and locations to the Khoesan communities. However, one is also surprized to learn that Naro speakers have terms for modern concepts such as "tourist" (*caate-kg'ao*), "zinc" (*nxam*) and "world" (*nqõó*). Most of all, one is astonished by the ingenuity of Naro speakers in creating terms for new concepts, such as "video" (*chõó-dxãwa*), which literally means "white person's devil".

Visser's dictionary is therefore of considerable value as a tool of reference. At the same time, it constitutes an important depository for a language which is classified as one of the endangered languages of the world.

Bibliography

Dickens, Patrick. 1994. *English–Ju/'hoan, Ju/'hoan–English Dictionary*. Quellen zur Khoisan-Forschung 8. Cologne: Ruediger Koeppe Verlag.

Haacke, Wilfrid H.G. and Eliphas Eiseb. 2002. *A Khoekhoegowab Dictionary with an English–Khoekhoegowab Index*. Windhoek: Gamsberg-Macmillian.

Traill, Antony. 1994. *!Xóõ Dictionary*. Quellen zur Khoisan-Forschung 9. Cologne: Ruediger Koeppe Verlag.

Willet, Shelagh et al. (Eds.). 2002. *The Khoe and San: An Annotated Bibliography. Volume One.* Gaborone: Lightbooks Publishers.

Herman M. Batibo
Department of African Languages and Literature
University of Botswana
Gaborone
Botswana
(batibohm@mopipi.ub.bw)

Publikasieaankondigings /
Publication Announcements

Willem Botha (Redakteur). *'n Man wat beur. Huldigingsbundel vir Dirk van Schalk-wyk.* 2003, xii + 384 pp. ISBN 0-9584401-4-X. Stellenbosch: Buro van die WAT. Prys: R150. (Resensie in hierdie uitgawe.)

Julie Coleman. *A History of Cant and Slang Dictionaries. Volume 1: 1567–1785.* 2004, xii + 259 pp. ISBN 0 19 925471 0 (Hb.). Oxford: Oxford University Press. Price: £45.

Ton den Boon (Samensteller). *Taal van het jaar drie: Kroniek van het Nederlands in 2003.* 2003, 123 pp. Utrecht/Antwerpen: Van Dale Lexicografie.

J.D.U. Geldenhuys en E.E. Viljoen-Smook. *New Business Dictionary/Nuwe Sake-woordeboek.* Fourth, revised and expanded edition/Vierde, hersiene en uit-gebreide uitgawe. 2003, xxi + 315 pp. ISBN 1 86890 038 X. Cape Town/Kaapstad: Pharos Dictionaries/Pharos Woordeboeke. Prys: R140. (Resen-sie in hierdie uitgawe.)

Manfred Görlach. *English Words Abroad.* 2003, xii + 189 pp. ISBN 90 272 2331 9 / ISBN 1 58811 435 X. Amsterdam/Philadelphia: John Benjamins. Price: €70, US$70 (Hb.). (Review in this issue.)

Volker Harm und Michael Schlaefer (Herausgeber). *Wörterbuchstadt Göttingen.* 2003, 140 pp. Göttingen.

R.R.K. Hartmann (Editor). *Lexicography: Critical Concepts. Volume I: Dictionaries, Compilers, Critics and Users.* xxvii + 412 pp. ISBN 0-415-25366-7. *Volume II: Reference Works across Time, Space and Languages.* xi + 383 pp. ISBN 0-415-25367-5. *Volume III: Lexicography, Metalexicography and Reference Science.* xi + 483 pp. ISBN 0-415-25368-3. ISBN 0-415-25365-9 (Set). London/New York: Routledge. Price: £425.

Thomas Herbst, Gunter Lorenz, Brigitta Mittmann und Martin Schnell (Her-ausgeber). *Lexikografie, ihre Basis- und Nachbarwissenschaften: (Englische) Wörterbücher zwischen "common sense" und angewandter Theorie.* 2004, VI + 207 pp. ISBN 3-484-39118-9. Lexicographica. Series Maior 118. Tübingen: Max Niemeyer. Price: €72.

456 Publikasieaankondigings / Publication Announcements

Werner Hüllen. *A History of Roget's* Thesaurus*: Origins, Development, and Design*. 2004, ix + 410 pp. ISBN 0-19-925472-9. Oxford: Oxford University Press. Price: £75.

Kyo Kageura. *The Dynamics of Terminology: A Descriptive Theory of Term Formation and Terminological Growth.* 2002, viii + 323 pp. ISBN 90 272 2328 9 (Eur.), 1 58811 314 0 (US) (Hb.). Amsterdam/Philadelphia: John Benjamins. Price: €92/US$92. (Review in this issue.)

Willy Kirkeby (Editor). *English–Swahili Dictionary*. 2000, 1069 pp. ISBN 9976888820 (Africa), 8299457327 (Non-African countries) (Pb.). Dar es Salaam: Kakepela Publishing. (Review article in this issue.)

Hans-Dieter Kreuder. *Metasprachliche Lexikographie: Untersuchungen zur Kodifizierung der linguistischen Terminologie*. 2003, X + 271 pp. Lexicographica. Series Maior 114. ISBN 3-484-39114-6. Tübingen: Max Niemeyer. Price: €78. (Review in this issue.)

Geoffrey Leech, Paul Rayson and Andrew Wilson. *Word Frequencies in Written and Spoken English*. 2001, xv + 304 pp. ISBN 0 582 32007 0. Harlow: Pearson Education. Price: £45.99.

Johann R.E. Lutjeharms. *Oseanologiewoordeboek Engels–Afrikaans/Dictionary of Oceanology English–Afrikaans*. 2004, xvi + 334 pp. ISBN 186890041X. Kaapstad/Cape Town: Pharos Woordeboeke/Dictionaries. Prys: R175. (Resensie in hierdie uitgawe.)

Sentrum vir Tekstegnologie. *Afrikaanse Speltoetser 2.0 en Woordafbreker*. 2003. Potchefstroom: Potchefstroomse Universiteit vir C.H.O.

W.F.J. Steenkamp en/and H.J.J. Reynders (Hoofredakteurs/Chief Editors). *Ekonomiese en Bedryfswoordeboek: Engels–Afrikaans/Dictionary of Economics and Business: English–Afrikaans.* Tweede, hersiene en vermeerderde uitgawe/Second, revised and enlarged edition. 2004, xii + 1199 pp. ISBN 1868900312. Kaapstad/Cape Town: Pharos Woordeboeke/Dictionaries. Prys/Price: R400.

Gabriele Stein. *Developing Your English Vocabulary: A Systematic New Approach*. 2002, VIII + 272 pp. Stauffenburg Linguistik 26. ISBN 3-86057-727-1. Tübingen: Stauffenburg Verlag. Price: €33.

Die Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. *Afrikaanse Woordelys en Spelreëls*. Negende, verbeterde en omvattend herbewerkte uitgawe. 2002, 592 pp. ISBN 1 86890 034 7. Kaapstad: Pharos Woordeboeke. Prys: R170. (Resensie in hierdie uitgawe.)

Rita Temmerman. *Towards New Ways of Terminology Description: The Sociocognitive Approach*. 2000, xv + 258 pp. ISBN 90 272 2326 2 (Eur.), 1 55619 772 1 (US) (Hb.). Amsterdam/Philadelphia: John Benjamins. Price: €75, US$75. (Review in this issue.)

Jenny Thumb. *Dictionary Look-up Strategies and the Bilingualised Learner's Dictionary: A Think-aloud Study*. 2004, XII + 127 pp. ISBN 3-484-39117-0. Lexicographica. Series Maior 117. Tübingen: Max Niemeyer. Price €48.

L.J. van der Veen et S. Bodinga-bwa-Bodinga. *Gedandedi sa Geviya/Dictionnaire Geviya–Français*. 2002, 569 pp. ISBN 90-429-1105-0 (Peeters Louvain)/ISBN 2-87723-621-8 (Peeters France). Louvain/Paris/Sterling, Virginia: Peeters. Price: €60. (Review in this issue.)

F.I.J. van Rensburg (Redakteur). *Afrikaans — lewende taal van miljoene*. 2004, xv + 190 pp. ISBN 0 627 02579 X. Pretoria: Van Schaik. Prys: R89.95. (Resensie in hierdie uitgawe.)

Piet van Sterkenburg (Editor). *A Practical Guide to Lexicography*. 2003, XI + 459 pp. Terminology and Lexicography Research and Practice 6. ISBN 90 272 2329 7 (Eur.), 1 58811 380 9 (US) (Hb.)/ISBN 90 272 2330 0 (Eur.), 1 58811 381 7 (US) (Pb.). Amsterdam/Philadelpia: John Benjamins. Price: €145, US$145 (Hb.)/€65, US$65 (Pb.).

Hessel Visser. *Naro Dictionary: Naro–English, English–Naro*. Fourth edition. 2001, viii + 240 pp. ISBN 99912-938-5-x. D'kar: Kuru Development Trust. (Review in this issue.)

VOORSKRIFTE AAN SKRYWERS

(Tree asseblief met die Buro van die WAT in verbinding (wat@wat.sun.ac.za) vir 'n uitvoeriger weergawe van hierdie instruksies of besoek ons webblad: <http://www.sun.ac.za/wat>)

A. REDAKSIONELE BELEID

1. Aard en inhoud van artikels
Artikels kan handel oor die suiwer leksikografie of oor implikasies wat aanverwante terreine, bv. linguistiek, algemene taalwetenskap, rekenaarwetenskap en bestuurskunde vir die leksikografie het.

Bydraes kan onder enigeen van die volgende rubrieke geklassifiseer word:

(1) **Artikels:** Grondige oorspronklike wetenskaplike navorsing wat gedoen en die resultate wat verkry is, of bestaande navorsingsresultate en ander feite wat op 'n oorspronklike wyse oorsigtelik, interpreterend, vergelykend of krities evaluerend aangebied word.

(2) **Resensieartikels:** Navorsingsartikels wat in die vorm van 'n kritiese resensie van een of meer gepubliseerde wetenskaplike bronne aangebied word.

Bydraes in kategorieë (1) en (2) word aan streng anonieme keuring deur onafhanklike akademiese vakgenote onderwerp ten einde die internasionale navorsingsgehalte daarvan te verseker.

(3) **Resensies:** 'n Ontleding en kritiese evaluering van gepubliseerde wetenskaplike bronne en produkte, soos boeke en rekenaarprogramme.

(4) **Projekte:** Besprekings van leksikografiese projekte.

(5) **Leksikonotas:** Enige artikel wat praktykgerigte inligting, voorstelle, probleme, vrae, kommentaar en oplossings betreffende die leksikografie bevat.

(6) **Leksikovaria:** Enigeen van 'n groot verskeidenheid artikels, aankondigings en nuusvrystellings van leksikografiese verenigings wat veral vir die praktiserende leksikograaf van waarde sal wees.

(7) **Verslae:** Verslae van konferensies en werksessies.

Bydraes in kategorieë (3)-(7) moet almal aan die eise van akademiese geskrifte voldoen en word met die oog hierop deur die redaksie gekeur.

2. Wetenskaplike standaard en keuringsprosedure
Lexikos is deur die Departement van Onderwys van die Suid-Afrikaanse Regering as 'n gesubsidieerde d.w.s. inkomstegenererende navorsingstydskrif goedgekeur.

Artikels sal op grond van die volgende aspekte beoordeel word: taal en styl; saaklikheid en verstaanbaarheid; probleemstelling, beredenering en gevolgtrekking; verwysing na die belangrikste en jongste literatuur; wesenlike bydrae tot die spesifieke vakgebied.

3. Taal van bydraes
Afrikaans, Duits, Engels, Frans of Nederlands.

4. Kopiereg
Nóg die Buro van die WAT nóg die African Association for Lexicography (AFRILEX) aanvaar enige aanspreeklikheid vir eise wat uit meewerkende skrywers se gebruik van materiaal uit ander bronne mag spruit.

Outeursreg op alle materiaal wat in *Lexikos* gepu

bliseer is, berus by die Beheerraad van die Woordeboek van die Afrikaanse Taal. Dit staan skrywers egter vry om hulle materiaal elders te gebruik mits *Lexikos* (AFRILEX-reeks) erken word as die oorspronklike publikasiebron.

5. Oorspronklikheid
Slegs oorspronklike werk sal vir opname oorweeg word. Skrywers dra die volle verantwoordelikheid vir die oorspronklikheid en feitelike inhoud van hulle publikasies.

6. Gratis oordrukke en eksemplare
Skrywers ontvang vyf gratis oordrukke van elke artikel of resensieartikel van hulle wat gepubliseer is asook een gratis eksemplaar van die uitgawe waarin sodanige artikel(s) verskyn het. Skrywers van suiwer evaluerende resensies en van bydraes tot die rubrieke Leksikonotas, Leksikovaria, Projekte en Verslae ontvang vyf gratis oordrukke van hulle bydraes. In laasgenoemde vier kategorieë kan die redaksie egter, afhangend van die aard en omvang van die bydraes, besluit om ook 'n eksemplaar van die betrokke uitgawe aan 'n skrywer toe te ken.

7. Uitnodiging en redaksionele adres
Alle belangstellende skrywers is welkom om bydraes vir opname in *Lexikos* te lewer en aan die volgende adres te stuur:

Die Redakteur: LEXIKOS
Buro van die WAT
Posbus 245
7599 STELLENBOSCH
Republiek van Suid-Afrika

B. VOORBEREIDING VAN MANUSKRIP

Die manuskrip van artikels moet aan die volgende redaksionele vereistes voldoen:

1. Lengte en formaat van artikels
Bydraes moet verkieslik nie 20 getikte A4-bladsye met teks in dubbelspasiëring en ruim kantlyne (ongeveer 2,5 cm) oorskry nie. Manuskrip moet verkieslik in elektroniese formaat as ASCII-teks, as volledig geformateerde Microsoft Word (DOS of Windows) lêers of as WordPerfect (DOS of Windows) lêers op rekenaarskyf (360 KB tot 1.44 MB) voorgelê word. 'n Rekenaardrukstuk van die artikel moet die skyf vergesel. Elke artikel moet voorsien wees van 'n Engelse opsomming van tussen 150 en 250 woorde, sowel as tussen 10 en 30 Engelse sleutelwoorde.

2. Grafika
Een stel duidelike oorspronklike illustrasies, tabelle, grafieke, diagramme, of kwaliteitsafdrukke daarvan, moet voorgelê word. Die plasing van grafika binne die teks moet duidelik aangedui word.

3. Bibliografiese gegewens en verwysings binne die teks
Kyk na onlangse nommers van *Lexikos* vir meer inligting.

INSTRUCTIONS TO AUTHORS

(For a more detailed version of these instructions, please contact the Bureau of the WAT (wat@wat.sun.ac.za) or refer to our website: <http://www.sun.ac.za/wat>)

A. EDITORIAL POLICY

1. Type and content of articles
Articles may deal with pure lexicography or with the implications that related fields such as linguistics, general linguistics, computer science and management have for lexicography.

Contributions may be classified in any one of the following categories:

(1) **Articles:** Fundamentally original scientific research that has been done and the results that have been obtained, or reflecting existing research results and other facts in an original, synoptic, interpretative, comparative or critically evaluative manner.
(2) **Review articles:** Research articles presented in the form of a critical review of one or more published scientific sources.

Contributions in categories (1) and (2) are subjected to strict anonymous evaluation by independent academic peers in order to ensure the international research quality thereof.

(3) **Reviews:** An analysis and critical evaluation of published scientific sources and products, such as books and computer software.
(4) **Projects:** Discussions of lexicographical projects.
(5) **Lexiconotes:** Any article containing practice-oriented information, suggestions, problems, questions, commentary and solutions regarding lexicography.
(6) **Lexicovaria:** Any of a large variety of articles, announcements and press releases by lexicographic societies which are of particular value to the practising lexicographer.
(7) **Reports:** Reports on conferences and workshops.

Contributions in categories (3)-(7) must all meet the requirements of academic writing and are evaluated by the editors with this in mind.

2. Academic standard and evaluation procedure
The Department of Education of the South African Government has approved *Lexikos* as a subsidized, i.e. income-generating research journal.

Articles will be evaluated on the following aspects: language and style; conciseness and comprehensibility; problem formulation, reasoning and conclusion; references to the most important and most recent literature; substantial contribution to the specific discipline.

3. Language of contributions
Afrikaans, Dutch, English, French or German.

4. Copyright
Neither the Bureau of the WAT nor the African Association for Lexicography (AFRILEX) accepts any responsibility for claims which may arise from contributing authors' use of material from other sources.

Copyright of all material published in *Lexikos* will be vested in the Board of Control of the Woordeboek van die Afrikaanse Taal. Authors are free however to use their material elsewhere provided that *Lexikos* (AFRILEX Series) is acknowledged as the original publication source.

5. Originality
Only original contributions will be considered for publication. Authors bear full responsibility for the originality and factual content of their contributions.

6. Free offprints and copies
Authors will receive five free offprints of each of their articles or review articles published, as well as one complimentary copy of the issue containing such article(s). Authors of purely evaluative reviews and of contributions to the categories Lexiconotes, Lexicovaria, Projects, and Reports receive five free offprints of their contributions. In the case of the latter four categories, the editors may, however, depending on the nature and scope of the contributions, decide to grant the author a copy of the issue concerned.

7. Invitation and editorial address
All interested authors are invited to submit contributions for publication in *Lexikos* to:

The Editor: LEXIKOS
Bureau of the WAT
P.O. Box 245
7599 STELLENBOSCH
Republic of South Africa

B. PREPARATION OF MANUSCRIPTS

Manuscripts of articles must meet the following editorial requirements:

1. Length and format
Contributions should not exceed more than 20 typewritten A4 pages with double spacing and ample margins (about 2,5 cms). Manuscript should preferably be in electronic form on a (360 KB to 1.44 MB) floppy disk as either ASCII text, fully-formatted Microsoft Word (DOS or Windows) or WordPerfect (DOS or Windows) files. A computer printout of the article should accompany the disk. Each article must be accompanied by an English abstract of 150 to 250 words, and between 10 and 30 English keywords.

2. Graphics
One set of clear original drawings, tables, graphs, diagrams or quality prints thereof must be submitted. The locations of graphics must be clearly indicated in the text.

3. Bibliographical details and references in the text
Examine recent issues of *Lexikos* for details.

HINWEISE UND RICHTLINIEN FÜR AUTOREN

(Nehmen Sie bitte mit dem Büro des WAT Kontakt auf (wat@wat.sun.ac.za) für eine ausführlichere Wiedergabe dieser Hinweise oder besuchen Sie usere Webseite: <http://www.sun.ac.za/wat>)

A. REDAKTIONELLE ZIELSETZUNGEN

1. Art und Inhalt der Artikel

Es können Artikel aufgenommen werden, die sich mit Themen der Lexikographie befassen oder mit Zusammenhängen, die zwischen der Lexikographie und benachbarten Fachgebieten wie z.B. Linguistik, allgemeiner Sprachwissenschaft, Lexikologie, Computerwissenschaft und Management bestehen.

Die Beiträge sollten einer der folgenden Kategorien entsprechen:

(1) **Artikel**, die grundlegend über neue Forschungsansätze und deren Ergebnisse berichten, oder die bestehende Forschungsergebnisse und andere Informationen selbständig, interpretativ, vergleichend oder kritisch bewertend wiedergeben.

(2) **Rezensionsartikel**, die in der Form eines Forschungsartikels eine oder mehrere veröffentlichten wissenschaftlichen Quellen kritisch rezensieren.

Beiträge in Kategorien (1) und (2) werden streng anonym von unabhängigen wissenschaftlichen Experten begutachtet, um ein internationales fachliches Niveau in *Lexikos* zu gewährleisten.

(3) **Rezensionen**, die veröffentlichte wissenschaftliche Quellen und Produkte, wie z.B. Bücher und Software, analysieren und kritisch bewerten.

(4) Lexikographische **Projekte**, die vorgestellt werden.

(5) **Notizen zum Lexikon**, die praxisbezogene Informationen, Vorschläge, Probleme, Fragen, Kommentare und Lösungen hinsichtlich der Lexikographie enthalten.

(6) **Lexikovaria**, die unterschiedliche Beiträge, Ankündigungen und Pressemitteilungen lexikographischer Vereinigungen, die dem praktischen Lexikographen wichtig sein können, einschließen.

(7) **Berichte** über Konferenzen und Workshops.

Beiträge in Kategorien (3)-(7) müssen im akademischen Stil abgefaßt werden. Sie werden von der Redaktion unter diesem Gesichtspunkt beurteilt.

2. Wissenschaftliche Standards und das Beurteilungsverfahren

Das Erziehungsministerium der südafrikanischen Regierung hat *Lexikos* als eine subventionierte, d.h. einkommenerzeugende Forschungszeitschrift anerkannt.

Artikel werden auf Grund der folgenden Gesichtspunkte bewertet: Sprache und Stil; Sachlichkeit und Verständlichkeit; Problembeschreibung, Argumentation und Schlußfolgerung; Hinweise auf die neueste und wichtigste Literatur; wesentlicher Beitrag zum besonderen Fachgebiet.

3. Sprache der Beiträge

Afrikaans, Deutsch, Englisch, Französisch oder Niederländisch.

4. Das Urheberrecht

Weder das Büro des WAT noch die African Association for Lexicography (AFRILEX) übernehmen Verantwortung für Ansprüche, die daraus entstehen könnten, daß Autoren Material aus anderen Quellen benutzt haben.

Das Urheberrecht aller in *Lexikos* publizierten Artikel wird dem Aufsichtsrat unseres Büros übertragen. Es steht Autoren jedoch frei, ihren Beitrag anderweitig zu verwenden, vorausgesetzt, *Lexikos* (AFRILEX-Serie) wird als Originalquelle genannt.

5. Originalität

Nur Originalbeiträge werden begutachtet. Autoren tragen die volle Verantwortung für die Originalität und den sachlichen Inhalt ihrer Beiträge.

6. Sonderdrucke und Freiexemplare

Autoren erhalten fünf Sonderdrucke ihrer veröffentlichten Artikel oder Rezensionsartikel gratis sowie ein Freiexemplar der betreffenden Ausgabe. Rezensenten und Autoren von Beiträgen zu den Kategorien Notizen zum Lexikon, Lexikovaria, Projekte und Berichte erhalten fünf Sonderdrucke ihrer Beiträge gratis. Die Redaktion kann sich jedoch, abhängig von der Art und dem Umfang der Beiträge der letztgenannten vier Kategorien, vorbehalten, dem Autor ein Freiexemplar der Ausgabe zu überlassen.

7. Einladung und redaktionelle Adresse

Alle Autoren, die interessiert sind, Beiträge für *Lexikos* zu liefern, sind herzlich willkommen. Sie werden gebeten, ihre Artikel an die folgende Adresse zu schicken:

Der Redakteur: LEXIKOS
Buro van die WAT
Postfach 245
7599 STELLENBOSCH
Republik Südafrika

B. VORBEREITUNG DES MANUSKRIPTS

Ein Artikelmanuskript muß den folgenden redaktionellen Anforderungen entsprechen:

1. Umfang und Format

Beiträge sollen nicht länger als 20 getippte A4-Seiten in zweizeiligem Abstand und mit Randabständen von ca. 2,5 cm sein. Das Manuskript sollte möglichst als elektronischer Text auf einer (360 KB bis 1.44 MB) Diskette vorgelegt werden, entweder im ASCII-Format, oder in formatiertem Microsoft Word (DOS oder Windows) bzw. WordPerfect (DOS oder Windows). Ein Ausdruck des vollständig formatierten Artikels soll mit der Diskette eingereicht werden. Jedem Artikel ist eine Zusammenfassung im Umfang von 150-250 Wörtern beizufügen. Ferner sollen etwa 10-30 inhaltskennzeichnende Stichwörter zu jedem Artikel angegeben werden.

2. Abbildungen

Ein reproduktionsfähiger Satz der originalen Abbildungen, Illustrationen, Tabellen, Graphiken und Diagramme oder Qualitätsabdrucke muß vorgelegt werden. Der Text selber sollte klare Hinweise auf die Position der Abbildungen enthalten.

3. Bibliographische Einzelheiten und Hinweise im Text

Zu Einzelheiten des bibliographischen Systems sind neuere Ausgaben von *Lexikos* einzusehen.

INSTRUCTIONS AUX AUTEURS

(Pour une version plus détaillée de ces instructions, contacter le Bureau du WAT (wat@sun.ac.za)
ou consulter notre website: <http://www.sun.ac.za/wat>)

A. POLITIQUE ÉDITORIALE

1. Caractéristiques et contenu des articles
Les articles seront consacrés à la lexicographie pure, ou aux rapports entre la lexicographie et les disciplines voisines telles que la linguistique, la linguistique générale, l'informatique et le management.

Les contributions pourront appartenir à l'une des catégories suivantes:

(1) **Articles:** Recherches scientifiques originales, avec leurs résultats; ou présentations originales, synoptiques, interprétatives, comparatives, évaluatives et critiques des résultats de recherches en cours;

(2) **Articles bilans:** Articles de recherche présentés sous forme de bilan critique de travaux scientifiques déjà publiés.

Les contributions appartenant aux catégories (1) et (2) seront soumises de manière anonyme à des experts spécialistes indépendants afin d'en assurer la qualité scientifique au niveau international.

(3) **Recensions:** Analyses et évaluations critiques de travaux de recherche et de productions scientifiques, telles que livres ou logiciels;
(4) **Projets:** Présentations de projets lexicographiques;
(5) **'Lexiconotes':** Textes contenant des informations pratiques, ou des suggestions, des problèmes, des questions, des commentaires et des solutions concernant des activités lexicographiques;
(6) **'Lexicovaria':** Articles, annonces, communiqués de presse émanant de centres de lexicographie et qui revêtent un intérêt particulier pour les lexicographes;
(7) **Rapports:** Rapports sur des colloques et ateliers.

Les contributions dans les catégories (3) à (7) devront répondre aux exigences de qualité des publications scientifiques et seront évaluées dans cette optique.

2. Critères et procédures d'évaluation
La revue *Lexikos* est reconnue et subventionnée par le Ministère de l'Éducation du gouvernement Sud-Africain comme revue devant générer des revenus.

Les articles seront évalués selon les critères suivants: langue et style, concision et clarté, formulation de la problématique, raisonnement et conclusion, référence aux travaux les plus importants et les plus récents, contribution substantielle à la discipline.

3. Langue des contributions
Afrikaans, allemand, anglais, français ou néerlandais.

4. Copyright
Le Bureau du WAT ou l'*African Association for Lexicography* (AFRILEX) décline toute responsabilité en cas de réclamations motivées par l'utilisation d'autres sources par les auteurs.

Les droits d'auteurs des documents publiés dans *Lexikos* appartiennent au Board of Control du *Woordeboek van die Afrikaanse Taal* (*WAT*). Cependant, les au

teurs 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*.