The Compilation of Multilingual Concept Literacy Glossaries at the University of Cape Town: A Lexicographical Function Theoretical Approach*

Dion Nkomo, School of Languages, Rhodes University, Grahamstown, and Department of Afrikaans and Dutch, Stellenbosch University, Stellenbosch, South Africa (deeouf@yahoo.co.uk)

and

Mbulungeni Madiba, Multilingualism Education Project, Centre for Higher Education Development, University of Cape Town, Cape Town, South Africa (mbulungeni.madiba@uct.ac.za)

Abstract: This article proposes a lexicographical approach to the compilation of multilingual concept literacy glossaries which may play a very important role in supporting students at institutions of higher education. In order to support concept literacy, especially for students for whom English is not the native language, a number of universities in South Africa are compiling multilingual glossaries through which the use of languages other than English may be employed as auxiliary media. Terminologies in languages other than English are developed by translating English terms or coining new terms in these languages to exploit the native language competence of most students. The glossary project at the University of Cape Town (UCT) which was conceived under the auspices of the Multilingualism Education Project (MEP) is discussed. It is shown that the UCT glossaries are compiled using methods consistent with those employed in modern lexicography or proffered in lexicographical theory. The lexicographical function theory is specifically used to account for the glossaries and their production. It is suggested that modern lexicography can provide useful guidance for the production of glossaries, given that the earliest glossaries constitute the humble beginnings of lexicography.

Keywords: MULTILINGUAL CONCEPT LITERACY GLOSSARIES, GLOSSARY, DICTIONARY, TERMINOLOGY, TERMINOGRAPHY, LEXICOGRAPHY, SPECIALISED LEXICOGRAPHY, USERS, CORPUS, WORDSMITH

Opsomming: Die samestelling van veeltaligekonsep-geletterheidswoorde-

^{*} Part of this article was presented as a paper at the *International Conference on Language Policy, Planning and Support in Higher Education — Challenges of Multilingualism,* organised by the Language Centre of Stellenbosch University and held at the Spier Hotel, Stellenbosch, 17-20 November 2009.

lyste by die Universiteit van Kaapstad: 'n Leksikografiesefunksieteoretiese benadering. Hierdie artikel stel 'n leksikografiese benadering tot die samestelling van meertaligekonsep-geletterheidswoordelyste voor wat 'n baie belangrike rol kan speel by die ondersteuning van studente by instellings vir hoër onderwys. Om konsepgeletterdheid te ondersteun, veral vir studente vir wie Engels nie die moedertaal is nie, stel 'n aantal universiteite in Suid-Afrika meertalige woordelyste saam waardeur die gebruik van ander tale as Engels as hulpmedia aangewend kan word. Terminologieë in ander tale as Engels word ontwikkel deur Engelse terme te vertaal of nuwe terme in hierdie tale te skep om die moedertaalvaardigheid van die meeste studente te benut. Die woordelysprojek by die Universiteit van Kaapstad (UK) wat ontwerp is onder die beskerming van die Multilingualism Education Project (MEP) word bespreek. Daar word aangetoon dat die UK-woordelyste saamgestel word deur die aanwending van metodes ooreenstemmend met dié gebruik in die moderne leksikografie of voorgestel in die moderne leksikografiese teorie. Die leksikografiese funksieteorie word spesifiek gebruik om die woordelyste en hul totstandbrenging te verantwoord. Daar word aangevoer dat die moderne leksikografie nuttige leiding vir die totstandbrenging van woordelyste kan verskaf, gegee dat die vroegste woordelyste die nederige begin van die leksikografie uitmaak.

Sleutelwoorde: Veeltaligekonsep-geletterheidswoordelyste, woordelys, woordeboek, terminologie, terminografie, leksikografie, gespesialiseerde leksikografie, gebruikers, korpus, wordsmith

1. Introduction

Following South Africa's multilingual language policy in 1994, African language terminologists, who are all first language speakers of the various official languages have been employed to document African language terminology on a variety of subject fields (Alberts 2010: 610). This has resulted in the publication of technical dictionaries or glossaries under the auspices of the Department of Arts and Culture (DAC). Alberts (1999, 2010) provides an outline of the policies, procedures and principles that have guided such terminology work. Besides the terminology work conducted by the Terminology Coordination Section, i.e. the section that deals with terminology under the National Language Services of the DAC, several South African institutions of higher education have engaged in the compilation of multilingual concept literacy glossaries. These glossaries may also play a significant part in the development of terminology in South Africa (Alberts 2010: 616). This article argues for a lexicographical approach towards the compilation of such glossaries. It suggests that modern lexicography provides a theoretically sound and comprehensive scope that may facilitate the production of functional and userfriendly products. Firstly, common fundamental motivations between the compilation of the earliest glossaries and modern lexicographical practices are noted, thereby dismissing what have been purported to be major differences between glossaries on the one hand and dictionaries on the other. Secondly, the boundaries between lexicography, particularly specialised lexicography, and

terminography, have already been dismissed as fluid and practically irrelevant according to the lexicographical function theory (Bergenholtz and Tarp 1995, 2003; Bergenholtz and Nielsen 2006; Tarp 2000; Fuertes-Olivera and Arribas-Baño 2008). The lexicographical function theory is then critically applied to account for the multilingual concept literacy glossaries being compiled under the auspices of the Multilingualism Education Project (MEP) at the University of Cape Town (UCT). The glossaries seek to satisfy the pedagogical needs especially of first year learners at UCT, who are non-native speakers of English.

2.1 Disciplinary Claims for Subject-Field Glossaries

Which discipline is responsible for the production of glossaries? This may seem to be an irrelevant question. Probably it has not been raised before because glossary makers, i.e. glossarians, glossarists or glossators, as they are variously called (cf. Hartmann and James 1998), are more likely to be concerned about the final products, i.e. glossaries, rather than how their disciplines shape the outputs. However, the question is both theoretically and practically significant. It is rather more a question of quality than of quantity or availability, taking into cognisance the motivations behind the creation of the earliest glossaries and the developments having since taken place to inform the production of glossaries today. Once the influence of the various disciplines informing the production of glossaries is clear, the nature of the available glossaries may be better understood with a view to determining how they may be improved.

There are at least three established practical fields or disciplines to which the compilation of glossaries have been or may be attributed. These are *glossography*, *terminography/terminology* and *lexicography*. The theoretical and disciplinary status of each of these will be examined in order to determine how much influence each has had on the production of glossaries in the course of time. Furthermore, the relationships and lack thereof between them is also explored in order to demonstrate the viability of the lexicographical approach.

In addition to the three fields or disciplines mentioned in the previous paragraph, subject-field glossaries have also been, and continue to be, produced by experts in a variety of other scientific, professional and academic disciplines. Some of the glossaries have been published as independent hard copies; others have been integrated or appended to textbooks or practical manuals while others are available online. Such glossaries will not be discussed at length here. Suffice it to say that their compilation largely relies on the specialised knowledge possessed by the experts in the specialised subjects. Despite their accuracy as far as specialised knowledge is concerned, some of the glossaries fall short when it comes to the linguistic and communicative needs of the target users.

2.1.1 Glossography

According to the definitions from the Free Online Dictionary, Thesaurus and Encyclopedia (accessed on 6 November 2009), glossography may be defined in sim-

ple terms as the compilation of glosses and glossaries. However, the term seems not to be popular any more as it does not appear in many contemporary English dictionaries, including dictionaries of linguistics. Even Sauer (2008), who has written a book chapter about glosses and glossaries, does not use the term. However, Hartmann and James (1998: 63) have lemmatised it and treated it in a relatively comprehensive and insightful way. They define *glossography* as:

A complex of activities concerned with the compilation of glossaries from the Middle Ages to the Renaissance. Starting with glosses or notes on hard words in Latin texts, vocabulary lists gradually developed into context-independent compilations, laying the foundations for lexicography, monolingual as well as bilingual and multilingual, alphabetic as well as thematic, lexical as well encyclopaedic.

This article of Hartmann and James confirms that glossography is concerned with the compilation of glosses and glossaries. However, it seems to suggest that this complex of activities was time-framed, i.e. "from the Middle Ages to the Renaissance". This then make the question regarding the disciplinary claim for the compilation of glossaries in the modern era relevant for determining the progress that has been made in practice. It is clear that it should no longer be glossography, which Hartmann and James (1998) describe retrospectively. More significantly, the article of Hartmann and James suggests that glossography provided the "foundations for lexicography". Then, given the progress that has been made by lexicography up to the present, to what extent has the production of glossaries incorporated modern developments in lexicographical theory and practice? This question, among others related to the similarities and differences between lexicography and the compilation of modern glossaries, will be addressed partly in Section 2.2.2 and thereafter in the remainder of the article.

2.1.2 Terminography (Terminology)

Terminography is another field which has produced subject-field glossaries through the documentation of terms (Sager 1996, Bowker and Pearson 2002). As a practical activity, terminography may be understood in the context of terminology. According to Sager (1996: 3), terminology has the following three senses:

- a vocabulary of a special subject field, i.e. terms;
- practices and methods used for the collection, description and presentation of terms;
- a theory, i.e. the set of premises, arguments and conclusions required for explaining the relationships between concepts and terms which are fundamental for a coherent activity.

Terminography is encapsulated in the second sense. Glossaries result from the collection, description and presentation of terms. In African languages, terminography work does not only involve the collection, description and presentation of terms. It also includes term-creation which is integral since equivalent terms are created for established terms in languages such as English. In South Africa, terminologists and terminographers have produced subject-field glossaries, e.g. the trilingual glossaries for subject fields such as Law, Psychology and Economics compiled by Pumlani Sibula at the Unit for IsiXhosa of the Language Centre of Stellenbosch University. However, terminological theories do not account fully for all the processes involved in the compilation of such glossaries. For instance, in the case of African languages, terminology development is not purely onomasiologically oriented as proffered in traditional terminological theories since in some cases term-creation begins with an English term rather than a concept. Furthermore, the issues of data presentation and accessibility in the final products may not be fully appreciated within the scope of existing terminological theories.

As a conscious practical activity and subject of intellectual enquiry, terminology may be traced back to the 1930s when Eugen Wüster, later to be regarded as the father of terminology, developed the General Theory of Terminology (GTT). Although the GTT was a milestone in the intellectual development of terminology as a field, it met severe criticism over time. Some of the issues continue to bedevil practical terminology as acknowledged in the *POINTER Project Final Report* (1996: 1):

A common problem of terminology work is that the importance and indeed the very nature of terminology is poorly understood. ... many people simply have no idea of what it is, while others, searching for an explanation of some sort, end up associating it with "thermal science" and hence radiators ... Related professions in the communications field, such as translation and technical writing, will often be aware of the word without having precise knowledge of what it entails.

The situation described here may be explained by interrogating the very foundations of terminology as a practical field. "Terminology ... was a science pioneered by subject experts" (Antia 2005). Wüster himself was an engineer and, as such, treated terms as engineered language, and not part of natural languages. Thus the initial focus of terminology was on concepts and their designations, terms, which were understood as units of knowledge, and not as lexical units. The relationship between concepts and terms was supposed to be fixed and univocal regardless of context. In quest of standardised communication, especially between experts, terms were strongly prescribed. Synonymy and variation were not tolerated. Terminography became extremely prescriptive to such an extent that it resulted it products which did not fully address cognitive and communicative needs, especially of laypersons.

However, successive scholars such as Sager (1984, 1996), Temmerman (2000) and Cabré (2000) have expanded the scope of understanding terminology. For example, Sager (1996) identifies three dimensions of terminology, namely the cognitive, the linguistic and the communicative. Temmerman (2000) and Cabré (2000) argue for a sociocognitive approach, with the latter developing an alternative terminological theory called the Communicative Theory of Terminology. Bowker and Pearson (2002) have incorporated developments from corpus linguistics in terminology development, particularly the compilation of glossaries. These later developments differ from the GTT in that they view terms as part of natural languages whose meanings may be dynamic, depending on their disciplines, communicative situations and contexts. They also acknowledge the fact that specialised communication does not only occur among experts but also between an expert and a semi-expert or an expert and a layperson, a user-categorisation which has been refined within the context of metalexicography (cf. Bergenholtz and Tarp 1995, 2003). In such situations, context plays an important role in the conceptualisation of the represented concept.

In the light of the foregoing, this article contends that the compilation of subject-field glossaries in South Africa needs to draw insights from modern theories and practices instead of being restricted by the tenets of the GTT. This has been shown elsewhere by scholars such as Madiba (2004, 2010) who demonstrates how corpus linguistics and its analytic tools may improve terminological work. However, this article argues that such developments in terminology may be incorporated into a lexicographical model and that such a model may improve the compilation and use of glossaries. In so doing, the article will dismiss some implied claims that lexicography lacks the dynamics facilitating the production of functional and user-friendly language tools (cf. Bowker and Pearson 2002; Moropa 2004).

2.1.3 Lexicography

Among a variety of its products, dictionaries are the primary concern of lexicography, with practical lexicography focusing on dictionary production and metalexicography focusing on the development of theories to support lexicographical practice. Glossaries have not received significant attention from metalexicography and any suggestion that attributes the production of glossaries to lexicography may meet with strong opposition, especially from those engaged in the production of glossaries. There are at least two reasons for this:

Most of the available dictionaries have major limitations and barely satisfy the needs they are purported to serve. Consequently, glossaries have been seen as an alternative to dictionaries which unfortunately provide limited assistance (cf. Bowker and Pearson 2002; Moropa 2004).

There is a general lack of awareness of the progress metalexicography has made to address some major limitations of existing dictionaries and the resultant potential practical lexicography has accrued over the years. Consequently, poor quality dictionaries continue to be used as general references to lexicography. In the light of this, scholars tend to view glossaries and other language resources such as corpora as alternatives to dictionaries without noting that these resources are not necessarily incompatible with dictionaries (cf. Bowker and Pearson 2002; Moropa 2004).

Because of reasons such as the above, the production of glossaries is justifiable. Glossaries are meant to serve certain needs which dictionaries have failed or do not seem to have the capacity to address. However, it will be argued that the compilation of glossaries per se is not a sufficient remedy given that dictionaries developed as modifications of glossaries (Sauer 2008). What really matters is the adoption of an approach and principles that would ensure that the products are not simply better than the available dictionaries but also much better by far than the earliest glossaries which laid the foundations for dictionarymaking. A poor glossary would not be any better than a poor dictionary. Therefore it will be demonstrated that the production of useful glossaries cannot be inspired by having as a point of departure anti-lexicographical discourses which are not only based on superficial distinctions between a dictionary and a glossary but which also fail to demonstrate the advantages of the latter over the former. This will be done by highlighting a lexicographical approach which is not based on the compartmentalisation of knowledge, be it between glossography, terminography and lexicography itself.

2.2 Building Useful Glossaries: Breaking Unnecessary Disciplinary Barriers

There is a disciplinary and theoretical confusion which affects the practical production of glossaries for various purposes. This confusion emanates from the often unnecessarily sharp distinction between glossaries and dictionaries as products of distinct disciplinary practices, namely glossography or terminography and lexicography. In South Africa, some practitioners involved in the compilation of glossaries are regarded as translators simply because translation is one of the main methods of creating terms whereby equivalent terms in indigenous languages are created to meet punctual translation needs of and by translators (Madiba 2004). While the alleged distinction between glossography and lexicography has barely been investigated in an insightful way, the distinction between terminography and lexicography, especially specialised lexicography, has been effectively dismissed as of no practical value (cf. Bergenholtz and Tarp 1995, 2003; Bergenholtz and Nielsen 2006; Tarp 2000; Fuertes-Olivera 2010; Fuertes-Olivera and Arribas-Baño 2008). More attention will be given to the former distinction while a brief recapitulation of the cited

sources will be made on the distinction between the latter. In so doing, the article puts the compilation of glossaries within a historical perspective which, in turn, translates into a lexicographical perspective.

2.2.1 A Historical Perspective: From Glossaries to Dictionaries

A historical perspective makes it clear that glossography is the activity preceding lexicography, whose progress has added the dimension of "a field of study" to the meaning of the suffix -graphy, an Anglicisation of the French -graphie, inherited from the Latin -graphia, which is a direct transliteration of the Greek verb which means "to write" (Free Online Dictionary, Thesaurus and Encyclopedia, accessed on 6 November 2009). The dictionary article corresponding to the lemma glossography, taken from Hartmann and James (1998: 63) makes it clear that the practice started with glosses or notes on hard words in Latin texts. Vocabulary lists then "gradually developed into context-independent compilations, laying the foundations for lexicography, monolingual as well as bilingual and multilingual, alphabetic as well as thematic, lexical as well encyclopaedic". Glossography may therefore be regarded as the humble beginnings of lexicography.

In spite of these reflections, it cannot be ascertained whether the earliest glossarians, glossarists or glossators called themselves such, but it is unlikely given that they were mainly educators who considered it helpful to explain to their learners some of the hard words used by text producers. It is doubtful whether they would also have called their activity 'glossography', as it was not their main engagement. What is seen is the development of a practice whereby text producers use difficult words of foreign origins (Latin and French) and educators make efforts to explain them, first within texts and later outside their contexts as independent compilations. This is similar to the modern practices mentioned earlier where specialised subject experts identify difficult concepts and explain them in glossaries which are either published as independent texts or integrated into textbooks.

The real functions of glosses and glossaries are summarised by Sauer (2008: 19):

The main purpose of the glosses, as well as glossaries, thus must have a didactic one: interlinear glosses facilitate the understanding and possibly also the learning of the glossed Latin text. Glossaries help the acquisition of the Latin vocabulary (and probably also of the English vocabulary). Thus many of the glosses and glossaries must have been used for teaching purposes in schools, especially in monastic and cathedral schools.

As may be noted, it was a quest to solve communication problems and facilitate knowledge acquisition and dissemination that motivated the earliest glossaries. The problems could have emanated from the fact that the texts contained foreign language elements and also that they contained specialised lan-

guage vocabulary items. The second cause of problems, namely terminology, is also highlighted by Sauer (2008: 19) in the context of professional fields such as medical practice:

Specific glossaries (class glossaries) such as the plant name glossaries were ... intended for use of physicians and healers, to help them identify the plants in Latin (and later also in Anglo-Norman) and to prepare the proper medicines.

The historicisation of lexicography shows not only that glossaries were fore-runners of dictionaries, as explained by Hartmann and James (1998) and Sauer (2008), but also that dictionaries have or ought to have exactly the same functions as glossaries (cf. Tarp 2008). However, as already stated, dictionaries have been criticised in favour of glossaries on the premise that they do not provide adequate assistance (cf. Bowker and Pearson 2002; Moropa 2004). For example, before arriving at such a conclusion, Bowker and Pearson (2002: 139) ask some questions:

How often do you have to consult more than one dictionary in order to find the information you are seeking? How often do you find that the information you are seeking is simply not there? How often do you choose a word in the dictionary without really knowing whether it is the right one? While dictionaries are an essential part of any language student's toolkit, they rarely provide the answer to all your questions.

Intriguing as these questions are and accurate as the conclusion is regarding many dictionaries, Bowker and Pearson fail, however, to construe a plausible argument for the preference of glossaries. That dictionaries contain more information than glossaries (Sauer 2008: 21), which is generally used as the main distinction, may not be their limitation except when such information serves no function or remains inaccessible. The progress made in lexicography, from the earliest glossary making up to the present, will also invalidate the claims that the limitations of dictionaries compared to glossaries is their inability to provide a specific type of information. For example, grammatical information, usage examples and sense discrimination are among other information types which can be found in modern dictionaries. This fact is ignored in Bowker and Pearson (2002). Sauer (2008: 22) correctly observes that "the difference between a glossary and a dictionary is one of degree and ... more by convention that the early collections are usually called glossaries and the later collections ... dictionaries". To this it suffices to add that the argument that no single dictionary will ever provide the information required by a student (Bowker and Pearson 2002: 140), also applies to some glossaries and does not make them any better than dictionaries in general. One product will be better than the other depending on how it has been conceived in view of its functions, not simply because one is a glossary and the other is a dictionary.

From the foregoing, it may not be wrong to suggest that in its reverse

mode, lexicography would go back to embody glossography. The earliest glossaries may then be regarded as the earliest lexicographical products. The question as to why dictionaries as they are known today developed from glossaries may be easily answered by pointing to their capacity to accommodate more data. Then why do glossaries continue to be produced alongside dictionaries? The answer is that the quality of dictionaries did not increase proportionally with the increase in the quantity of information they provide. That the quality of assistance provided by glossaries is generally better than that provided by dictionaries has so far not been adequately demonstrated. As such, the production of glossaries may still benefit from lexicographical theories and methods primarily focused on dictionaries.

2.2.2 A Lexicographical Perspective

The historical perspective which suggests that lexicography encapsulates the compilation of glossaries, or glossography, prompts the lexicographical approach advanced here. It is said that dictionaries, not in their current form but as glossaries, were first compiled over thousands of years (McArthur 1986; Tarp 2008). However, it was not until the second half of the 20th century that lexicography started to constitute itself as an academic discipline and professional practice. Prior to that, dictionary making was generally a part-time activity, starting with educators compiling glossaries for their students and then subject-field experts compiling subject-field glossaries, with linguists claiming the rest through their involvement in 'general language lexicography'. This became the greatest source of confusion regarding the disciplinary status of lexicography, with linguists generally regarding lexicography as part of so-called 'applied linguistics'.

That lexicography is not part of linguistics was first challenged by Wiegand (1984) who argued that it was a professional activity and a scientific practice whose object was not language but the production of dictionaries. He also argued that dictionaries were utility products with each having its own genuine purpose, thereby making the user an important variable in the whole enterprise. His ideas were adopted by practising and academic lexicographers at the Centre for Lexicography at the Aarhus School of Business, Denmark, who collectively developed what they now call the lexicographical function theory (Bergenholtz and Tarp 1995, 2003; Tarp 2000, 2002, 2008). However, the proponents of the lexicographical function theory did not uncritically adopt Wiegand's theories, especially his general theory of lexicography. Tarp (2000, 2002, 2008), for example, has criticised Wiegand for having a linguistic bias, before strongly arguing that lexicography is in fact an independent scientific discipline. This called for a development of lexicographical theories which would improve the lexicographical practice and its products. Linguistic training is no longer regarded as a prerequisite qualification and it has been shown that the relevance of its theories and findings varies from one dictionary to another, with LSP lexicography requiring more than linguistic skills.

The lexicographical function theory and its application in lexicographical practice and dictionary criticism have been described in more detail by its proponents in several publications which include Bergenholtz and Tarp (1995, 2003), Bergenholtz and Nielsen (2006), and Tarp (2000, 2002, 2008). Its main argument is that every dictionary should be produced with a specific user in mind, devoting attention to questions such as the following (cf. Bergenholtz and Tarp 2003: 173):

- Who are the intended users of the prospective dictionary?
- What is their native language and how are they competent in it?
- What is their competence in a particular foreign language?
- What are the characteristics of the intended users in terms of their competence regarding specific languages and subject knowledge?
- In what situations are they likely to experience problems which may be solved by referring to a dictionary?
- What problems are they likely to encounter in the specific situations which may be solved by referring to a dictionary?
- What information needs may address the specific problems?
- What types of data need to be included in a dictionary which may provide users with certain information types required to solve specific problems encountered by specific users in specific situations?
- What lexicographical devices should be employed to ensure an unimpeded and successful access to data while reducing information costs?

Focusing on the glossary project of the University of Cape Town, Section 3 applies some ideas from the lexicographical function theory to demonstrate that the production of multilingual concept literacy glossaries may benefit from some lexicographical insights.

3. UCT's Multilingual Concept Literacy Glossaries

The UCT Multilingual Concept Literacy Glossaries project was initiated in 2007 as part of the implementation of the University's Language Policy (1999 revised in 2003) and the Language Plan adopted in 2003. The Language Plan requires that multilingual concept literacy glossaries be developed to support students for whom English is not their first language. Thus, the glossaries are aimed at concept literacy in the different content-learning areas, with the pilot project focusing on

developing glossaries for Statistics, Economics and Law. All these glossaries are based on the Special Language corpora constructed for this purpose. Drawing from the lexicographical function theory, the remainder of the article characterises the target users of the glossaries, defines the user situations, enumerates the envisaged assistance that the glossaries were conceived to provide to the target users and shows how all this informed the compilation of the glossaries.

3.1 User Characterisation

The glossaries being produced are targeted at first year students of various specialised subject fields at the University. The following screen shot shows a list of the disciplines that the project seeks to cover in the long term.

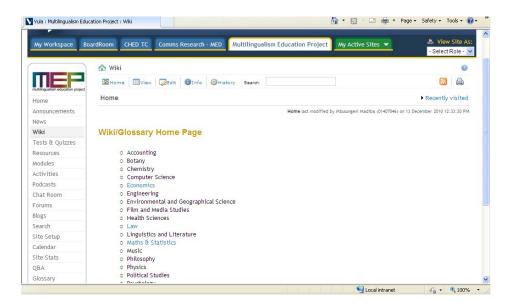


Figure 1: Subject fields for which glossaries are created at UCT

At first year stage, students have low levels of competence in the respective fields and low proficiency in English which is the primary language of instruction at UCT. Some of these students have high levels of competence in their home languages depending on the location of the schools in which they matriculated. In addition to language competences and competence in the specific subjects, another characteristic of the students that should be considered is their cultural backgrounds, which may have a bearing on their competence in subjects such as Law.

The focus on first year non-native English speaking students is very much

consistent with the requirement of the lexicographical function theory that every lexicographical product should be conceived with a specific user in mind. This requirement may be problematic when applied uncritically in developing countries where lexicographical resources and resources required to support lexicographical projects are very limited. In this regard, other potential users of the glossaries were not neglected at the conceptualisation stage. English native-speaking first year students may also use the UCT glossaries, seeing that general language competence in English does not automatically translate to LSP competence in the various subject fields. Furthermore, students at higher levels of university education may still consult the glossaries when they experience some of the typical challenges encountered by first year students for whom the glossaries are specifically produced. Even outside university, translators have also been considered as potential users of the UCT multilingual glossaries, while those who have participated in the project found it potentially useful. What is critical though is that first year non-native English speaking students are deliberately prioritised, while not neglecting other potential users.

3.2 User Situations

The situations in which the multilingual glossaries are envisaged to help students relate to their learning of the key concepts in the various subject fields, coupled with the mastering of the respective LSP varieties and registers. The learning situations may include lectures, tutorials, the writing of assignments and revision for examinations. In terms of the lexicographical function theory, these situations may be classified into communicative and cognitive functions (cf. Tarp 2008: 84). The glossaries were conceived to address some of the problems learners may encounter in such learning situations.

3.3 Problems

In the learning situations described in Section 3.5, students first encounter communication-oriented problems, primarily the reception of oral texts (lectures) as well as written texts (textbooks, course readers and assignment questions). They also experience text production problems like oral contributions and questions in lectures and tutorials, as well as text production in the form of writing assignments and examinations. Ultimately, this prevents them from grasping the key concepts and reaching the required performance levels in the specific subjects.

3.4 User Needs

Students have various needs to address both communicative and cognitive

problems in their specific learning situations. They need appropriate terminology and information regarding meaning, register and examples of its usage. This is necessary to enable them to receive and produce academic texts in the various subject fields. Those students whose home language is not English need translation equivalents of the terminology as well as translated definitions in order to address text reception problems. In addition to this, all students need special and encyclopaedic information about the concepts so that they may master the respective subject field.

3.5 Functions of the Glossaries

The MEP concept literacy glossaries seek to serve both cognitive and communicative functions.

The major cognitive functions are:

- Assisting university students at first year level with specialised knowledge of threshold concepts in various subject areas;
- Assisting university students at first year level with English LSP in various subject areas;
- Assisting students with special subject-field information in the form of key concepts and conceptual relations within the subject fields.

The communicative functions include:

- Assisting students with text reception in English;
- Assisting students with text production in English;
- Assisting lecturers and tutors with text production in indigenous languages;
- Assisting lecturers, tutors, learners and translators with the translation of texts from English to indigenous languages;
- Assisting students with text reception in English;
- Assisting students with text reception via their indigenous languages as auxiliary medium;
- Assisting students with text production in English.

In view of the above functions, the genuine purpose shared by the MEP glossaries is to facilitate concept literacy in English, the main language of education at UCT, and in indigenous languages as auxiliary media. The glossaries are thus conceived as multilingual pedagogical resources which facilitate communication and learning in the context of special subject fields. These

aspects guided the planning and compilation of the glossaries, for, following Tarp (2004: 312), "no data is included because of tradition or the practice of existing dictionaries". This is a central tenet of the lexicographical function theory.

The Compilation Process 3.6

The compilation of the multilingual concept literacy glossaries at UCT was mainly based on special language corpora. However, lecturers, tutors as well as translators and terminologists who are native speakers of the official languages of South Africa other than English were involved for their different kinds of expertise. The next subsections describe the special language corpora, the compilation of wordlists, the use of corpus data in defining and the online presentation of the glossaries.

3.6.1 Special Language Corpora

The special language corpora are used as the main conceptual bases for the glossaries. The corpora are based on generally accepted criteria for designing special language corpora (cf. Bowker and Pearson 2002; Madiba 2004). These criteria include size, text types, publication status, text origin, constitution of the texts, authorship, external and internal criteria. With regard to the first criterion of size, special language corpora tend to be much smaller than general language corpora as a result of their content and the compilation process (Lawson 2001: 293). Several studies recommend at least 30 000 words for a small special language corpus. Although the size of some of the UCT corpora is small, most important is the purposes for which they have been designed (Meyer and Mackintosh 1996).

The second criterion of text selection is important since not all texts are useful for special language corpora. Text selection involves choice of the language and type of the texts. The UCT corpora comprise English texts only since there are no texts available in the indigenous African languages at higher education level. Only written texts were selected and such texts are based on either module texts or full texts. To achieve a good balance in each corpus, texts are selected according to external as well as internal criteria. The external criteria are essential to maintain a good balance in the corpora. These criteria include the domains of individual texts, text genres, publication status and age, text origins, constitution, and persons who produce texts (authors). With regard to the domains of individual texts, specialised texts are considered, i.e. texts covering different genres developed for special purposes and specialised subject fields. In order to allow for the analysis of terms used in specialised subject domains and the compilation of specialised glossaries for these domains, it is important that domains be clearly identified from the start. Once the boundaries of a domain have been determined, it was attempted to establish the truest possible representation of the delimited domain in the corpus. The chosen body of texts should cover all aspects of the domain, including subdomains and related domains, as equally as possible. As a rule of thumb, the focus should be on one domain at a time (cf. Madiba 2004).

The third criteria concerns text genres. A text genre is a text type with its own linguistic and pragmatic conventions (Meyer and Mackintosh 1996). Since different domains produce different texts, various texts occurring in a single domain were considered. These texts were divided into two general types, namely, *instructional texts* (texts used for pedagogical purposes such as books and tutorials) and *advanced texts* (abstracts of theses/dissertations and full theses/dissertations where necessary). With these genres the corpora should be able to achieve linguistic as well as pragmatic balance.

The fourth criterion concerns the selection of texts on the basis of their publication status or age. The UCT corpora comprise published and unpublished instructional materials. If unpublished, the authenticity of such texts is checked. With regard to age, the main focus is on current texts as the purpose is to establish synchronic terminological corpora focusing on the present state of terminology used in different domains.

Lastly, the constitution of the texts and their internal criteria were considered. The internal criteria have to deal with formal linguistic characteristics such as factuality and technicality. In compiling the corpora, it was attempted to include texts that are as factual as possible and thus to enrich the corpus linguistically and conceptually. The texts should provide an extensive range of terms generated by the discourse community, accompanied by a maximum number of usages of concepts in different contexts. The technicality of the texts should range from high to low. Highly technical texts usually involve communication between subject specialists whereas less technical texts relate to communication between non-specialists (e.g. tutorials). The different degrees of technicality is useful for the identification of terms and concepts used in different domains, that is, specialised usage and everyday usage (cf. Madiba 2004).

Regarding the criteria discussed in the foregoing, it may be noted that consideration was given to the texts students encounter in their respective subject areas. The texts have direct consequences for the learning situations as well as the problems and needs experienced by the learners. The UCT Glossary of Economics will be used to illustrate aspects of the previous discussion. This glossary is based on a small corpus consisting of prescribed books, study guides and tutorials in Economics. The size of the corpus is about 70 000 running tokens. Although the corpus is small in size, it contains relevant key concepts and their contexts because of the module approach used in collecting texts. As Sinclair (2001) rightly pointed out, the advantage of a small corpus is that it is unnecessary to wait until the corpus is complete, but instead, work can start from the first day.

3.6.2 Extraction of Terms and Compilation of the Basic Wordlist

The extraction of terms was carried out by using WordSmith Tools. WordSmith Tools does automatic term extraction on the basis of statistical analyses. However, it requires that the texts be prepared in text format first before term extraction is carried out. The following are examples of terms extracted from the Economics corpus by using WordSmith Tools:

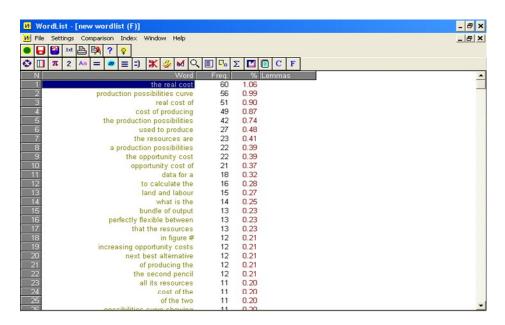


Figure 2: A frequency wordlist generated from the Economics corpus using WordSmith Tools

A final wordlist was compiled not solely on the basis of the number of occurrences or frequencies. Having gone through the basic list, Economics lecturers and tutors were consulted since some common words would normally have a high frequency of occurrence in specialised texts while some terms would be of a relatively low frequency. After the compilation of the final wordlist, concordances were generated using WordSmith Tools and Multiconcord to identify the meanings or senses of the terms in different contexts. Figure 3 is an illustration of the concordances of the term *opportunity cost* generated by using WordSmith Tools

The concordances in the examples provide different contexts of the term *opportunity cost*. In WordSmith tools, these contexts can be expanded as in the screen shot shown in Figure 4 to provide a better understanding of the concept:

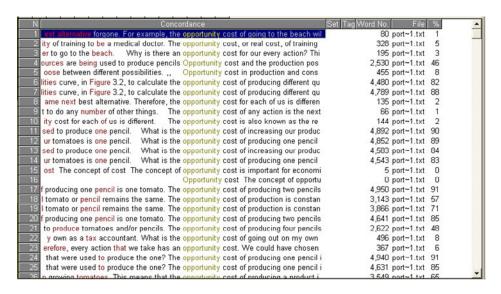


Figure 3: Examples of concordances extracted from the Economics corpus using WordSmith Tools

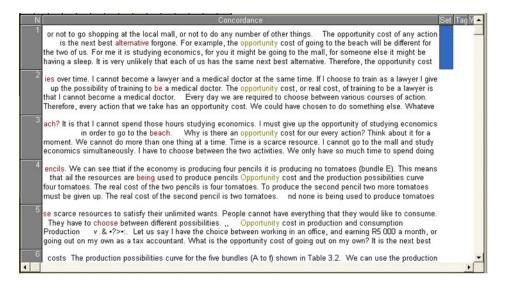


Figure 4: Examples of expanded concordances extracted from the Economics corpus using WordSmith Tools

The contextual examples are saved as a Word document and then converted into tables for better analysis as shown in Table 1.

EXAMPLES IN CONTEXT	SENSE
1. <s>The concept of opportunity cost is important for</s>	Basic meaning
economists. <s>Whenever we choose to do something we also</s>	(cost in terms of
choose not to do something else. <s>For example, if I choose</s>	choices)
to go to the beach today it means that I have also chosen not to	
study economics, or not to go shopping at the local mall, or	
not to do any number of other things. (choice)	
2. <s>The opportunity cost of any action is <u>the next best</u></s>	Basic meaning
<u>alternative forgone</u> . <s>For example, the opportunity cost of</s>	(cost in terms of
going to the beach will be different for the two of us. <s>For</s>	choices)
me it is studying economics, for you it might be going to the	
mall, for someone else it might be having a sleep. <s>It is very</s>	
unlikely that each of us has the same next best alternative.	
<s>Therefore, the opportunity cost for each of us is different.</s>	
3. <s>The opportunity cost is also known as the real cost.</s>	Additional meaning
<s>What is the real cost to me of going to the beach? <s>It is</s></s>	(real cost)
that I cannot spend those hours studying economics. <s>I</s>	,
must give up the opportunity of studying economics in order	
to go to the beach.	
4. <s>Why is there an opportunity cost for our every ac-</s>	Cost in terms of
tion? <s>Think about it for a moment. <s>We cannot do more</s></s>	time
than one thing at a time. <s>Time is a scarce resource. <s>I</s></s>	
cannot go to the mall and study economics simultaneously.	
<s>I have to choose between the two activities. <s>We only</s></s>	
have so much time to spend doing various activities during	
any hour, day, month, year, or lifetime. <s>The quantity of</s>	
time available to us is limited. <s>We cannot do everything</s>	
that we would like to do.	
5. <s>Let us say I have the choice between working in an</s>	Cost in terms of
office, and earning R5 000 a month, or going out on my own	money
as a tax accountant. <s>What is the opportunity cost of going</s>	money
out on my own? <s>It is the next best alternative. <s>If I did</s></s>	
not go out on my own I could be earning R5 000 a month.	
Strictly speaking, the real cost is the output that I could	
buy with the R5 000.	
6. <s>What is the opportunity cost to me if I spend R100</s>	Cost in terms of
on a CD? <s>It is the next best alternative. <s>This may be a</s></s>	production
shirt. <s>Therefore the real cost to me of consuming the CD is</s>	Production
that I cannot consume the shirt.	
7. > <s>Societies choose the output that they wish to con-</s>	National context
sume. <s>Whenever society decides to consume a particular</s>	(Cost of production
good it decides not to consume the next best alternative. <s>All</s>	and consumption)
societies have to consider the opportunity costs of any pro-	and consumption)
duction or consumption decisions that they take.	
duction of consumption decisions that they take.	

 Table 1: Concordances for opportunity cost

From these concordances, explanations can be developed by analysing the different contexts in which the term opportunity cost is used. Sometimes it is also possible to find a full definition of the term in the very concordances that are being analysed, concordances 1, 2 and 3, for example. From these concordances it can be observed that opportunity cost involves: (a) a choice between two alternatives, (b) a choice of the next best alternatives, and (c) the real cost. As shown above, an analysis of each of the concordances gives rise to different senses or meanings. The different contexts of the term opportunity cost provide the students with multiple exposures to the concept. Such multiple exposures to the term are essential for a student to understand its meaning or different senses. For students to understand the concepts, they need to be introduced to the concepts in their simplified form at the beginning, and then progress to learning more complex meanings of the concept. They also need to learn how to apply it in a range of contexts as shown in the examples above. With the aid of these examples, students are introduced to the basic meaning of the concept opportunity cost in the first concordance and then progress to its deeper meaning in the concordances that follow. In this way, definitions based on concordances are more elaborate and helpful to students than traditional definitions, especially from dictionaries not based on corpora. Providing students with mere definitions to memorize, results in a superficial understanding of the concept. According to O'Hara and Pritchard (2009: 11), "students must have both definitional and contextual information about words, as well as repeated exposures and opportunities to learn and review them". This supports the reception, production and cognitive functions of the glossaries.

3.6.3 Multilingual Concept Literacy Glossaries on Hypermedia (Vula)

The glossaries that are being developed at UCT are uploaded on Vula, the University Online Learning Environment developed by the Centre for Education Technology and powered by Sakai. This networked Online Learning Environment provides students with easy access to the glossaries and other online courses. It is observed that the use of the online environment exploits some of the developments utilised in the compilation of electronic dictionaries, a case of lexicographical development in the context of technological advances in the 21st century. The MEP Online Learning Environment on Vula is shown in Figure 5 followed by the Online Glossaries Site in Figure 6.

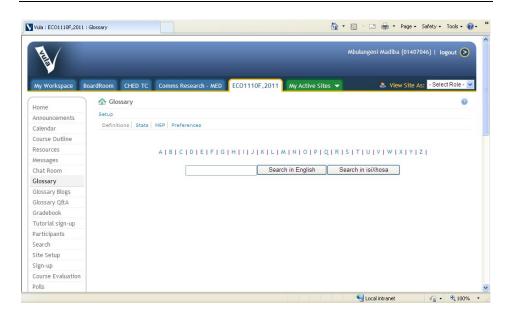


Figure 5: Economics Online Glossaries Hypermedia on Vula

From this screen shot, it may be observed that the Vula Hypermedia provides many search-route options, rendering the glossaries poly-accessible. Concepts can be accessed by keying the term into the search function or browsing the concepts in the selected languages. Language preference can be set from the start. In Figure 5, English and isiXhosa are shown as the preferred languages. However, any language can be selected as the preferred language. The search function is very effective in identifying conceptual relations as shown in the following examples of the concept *deficit*.

English Venda

Balance-of-payments deficit
Budget deficit
Capital account deficit
Conventional deficit
Cyclical deficit
Deficit
General government deficit
International balance-of-payments deficit
Payments deficit
Primary deficit
Trade deficit

Ndinganyiso kha mbadelo ya Ṭhahelelo Ṭhahelelo ya mugaganyagwama Ṭhahelelo kha akhauthu ya khephithala Ṭhahelelo ya mugaganyagwama Ṭhahelelo yo vhangwaho nga risesheni Ṭhahelelo ya muvhuso yo angaredzwaho Ṭhahelelo ya zwiṭundwa na zwivhambadzelwannda Mbadelo dzo padaho Difisithi ya phuraimari Phadambambadzelaseli

If the user does not want to search a specific concept, browsing can be used as in the traditional alphabetical glossary. In the following example, browsing is done on the letter A. This results in two wordlists, one for the source language and another for the target language.

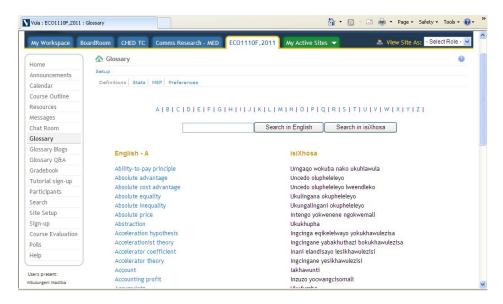


Figure 6: Economics Online Glossaries Hypermedia on Vula

Once the concept is identified, a further search can be made on its definition as shown in Figure 7.

The definition appears in both the source and the target languages. The list of other languages appears at the bottom of the screen shot. The glossary site also provides a space for the user to give feedback of the translation and the equivalent. The status of the translation equivalent is also shown as 'unassigned', 'draft', 'approved' and 'assigned'. This function is meant to facilitate the standardisation of the translated terms in the African languages based on popular usage and acceptability by the target users. While some of the terms were created by translators, the power to standardise the terms partly rests with the users, students and other professionals in the disciplines. Because terms and translation equivalents are not prescribed, synonyms and variants are provided, contrary to the stipulations of the GTT. The function of assigning status to an entry offers the users the best of both description and prescription, a practice which is called *proscription* in lexicography (Bergenholtz 2003).

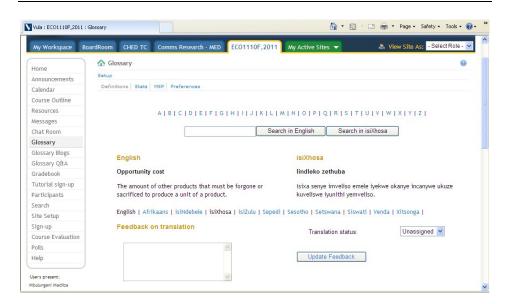


Figure 7: MEP Online Learning Environment on Vula

The definitions can further be linked with tutorials, pictures, graphics and podcasts. The glossary site also includes other functions such as modules, blogs, chat rooms and forums. Thus, unlike the traditional glossaries, the Vula Glossaries provide a more interactive environment. This makes the glossaries useful for users in reception, production and cognitive situations whereby punctual and extended searches are supported. This effectuates user-friendliness, a notion that has been discussed comprehensively in lexicography, especially with regard to the accessibility of lexicographical data.

4. Conclusion

It is still too early to assess the success of the UCT multilingual concept literacy glossaries. Further research will be conducted in this regard. However, their interactive nature and associated elements such as pictures, podcasts and blogs make the glossaries valuable resources for supporting non-native English speaking students learning concepts and their linguistic representations in their first languages. More critically, it is the user perspective which gives the glossaries more potential than existing glossaries and dictionaries. The user perspective is supported by the adoption of a corpus-based approach to the enterprise. Ahmad et al. (1992: 148) argue that the approach is "sound in as far as it is corpus-based and user-informed". Having identified the target users, their relevant characteristics and the situations in which their needs arise, the adopted corpus-based approach is characterised by a purposive selection of

texts and collaboration with both lecturers and students. All these ideas and principles have been adopted in modern lexicography. This therefore shows that the production of glossaries and other terminographical products may benefit from the advances made in lexicography.

The Compilation of Multilingual Concept Literacy Glossaries at UCT

References

- Ahmad, K., H. Fulford and M. Rogers. 1992. The Elaboration of Special Language Terms: The Role of Contextual Examples, Representative Samples and Normative Requirements. Tommola, H. et al. (Eds.). 1992. EURALEX '92 Proceedings. Papers Submitted to the 5th EURALEX International Congress, Tampere, Finland, 4-9 August 1992: 139-149. Tampere: Department of Translation Studies, University of Tampere.
- Alberts, M. 1999. Terminology in South Africa. Lexikos 9: 18-35.
- Alberts, M. 2010. National Language and Terminology Policies A South African Perspective. Lexikos 20: 599-620.
- **Antia, B.** 2005. Lexicography versus Terminology: Some Practical Reasons for Distinction. Presentation at the *International Training Workshop on Basic Principles of Terminology Management*, organised by the Department of Arts and Culture, Republic of South Africa. Johannesburg, 26-30 September, 2005.
- **Bergenholtz**, **H.** 2003. User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography. *Lexikos* 13: 65-80.
- Bergenholtz, H. and S. Nielsen. 2006. Subject-field Components as Integrated Parts of LSP Dictionaries. *Terminology* 12(2): 281-303.
- **Bergenholtz H. and S. Tarp.** 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes. Journal of Linguistics* 31: 171-196.
- Bergenholtz, H. and S. Tarp. (Eds.). 1995. Manual of Specialised Lexicography. Amsterdam: John Benjamins.
- Bowker, L. and J. Pearson. 2002. Working with Specialized Language: A Practical Guide to Using Corpora. London: Routledge.
- **Cabré, M.T.** 2000. Elements for a Theory of Terminology: Towards an Alternative Paradigm. *Terminology* 6(1): 35-57.
- **Free Online Dictionary, Thesaurus and Encyclopedia**. http://www.thefreedictionary.com. Accessed on 6 November 2009.
- **Fuertes-Olivera, P.A and A. Arribas-Baño.** 2008. *Pedagogical Specialised Lexicography. The Representation of Meaning in English and Spanish Business Dictionaries*. Amsterdam: John Benjamins.
- Fuertes-Olivera, P.A. (Ed.). 2010. Specialised Dictionaries for Learners. Berlin: De Gruyter.
- Hartmann, R.R.K. and G. James. 1998. Dictionary of Lexicography. London/New York: Routledge.
- Lawson, A. 2001. Collecting, Aligning and Analyzing Parallel Corpora. Ghadessy, M., A. Henry and R.L. Roseberry (Eds.). 2001. Small Corpus Studies and ELT: Theory and Practice: 279-310. Amsterdam: John Benjamins.
- **Madiba, M.** 2004. Parallel Corpora as Tools for Developing the Indigenous Languages of South Africa with Special Reference to Venda. *Language Matters* 35(1): 133-147.
- Madiba, M. 2010. Fast-tracking Concept Learning to English as an Additional Language (EAL) Students through Corpus-based Multilingual Glossaries. *Alternation* 17 (1): 225-248.

- **McArthur, T.** 1986. Worlds of Reference. Lexicography, Learning and Language from the Clay Tablets to the Computer. Cambridge: Cambridge University Press.
- Meyer, I. and K. Mackintosh. 1996. The Corpus from a Terminographer's Viewpoint. *International Journal of Corpus Linguistics* 1(2): 257-285.
- Moropa, K. 2004. A Parallel Corpus as a Terminology Resource for Xhosa: A Study of Strategies Used to Translate Financial Texts. *Language Matters* 35(1): 162-178.
- O'Hara, S. and R. Pritchard. 2009. Teaching Vocabulary with Hypermedia. Boston: Pearson.
- POINTER Project Final Report. 1996. Collated by the POINTER Workpackage 7 team (K.Ahmad, R. Bonthrone, G. Engel, A. Fotopoulou, D. Fry, C. Galinski, J. Humbley, N. Kalfon, M. Rogers, C. Roulin, K. Schmalenbach and E. Tanke) under the co-ordination of D. Fry. http://www.computing.surrey.ac.uk/ai/pointer/report/section1. Accessed on 12 April 2011
- Sager, J.A. 1984. Terminology and the Technical Dictionary. Hartmann, R.R.K. (Ed.). 1984. *LEXeter* '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983: 315-326. Tübingen: Max Niemeyer.
- Sager, J.A. 1996. A Practical Course in Terminology Processing. Amsterdam: John Benjamins.
- Sauer, H. 2008. Glosses, Glossaries and Dictionaries in the Medieval Period. A.P.Cowie (Ed.). 2009: 17-40. *The Oxford History of English Lexicography. Volume I: General-Purpose Dictionaries*. Oxford: Clarendon Press.
- Sinclair, J.M. 2001. Preface. Ghadessy, M., A. Henry and R.L. Roseberry (Eds.). 2001. *Small Corpus Studies and ELT: Theory and Practice*: vii-xv. Amsterdam: JohnBenjamins.
- Tarp, S. 2000. Theoretical Challenges to Practical Specialised Lexicography. Lexikos 10: 189-208.
- Tarp, S. 2002. Basic Elements of Lexicographic Theory. Emejulu, J.D. (Ed.). 2002. Éléments de Lexicographie Gabonaise. Tome II: 7-20. New York: Jimacs-Hillman.
- **Tarp, S.** 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.
- **Tarp, S.** 2008. Lexicography in the Borderland between Knowledge and Non-Knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography. Tübingen: Max Niemeyer.
- **Temmerman, R.** 2000. Towards New Ways of Terminology Description: The Sociocognitive Approach. Amsterdam: John Benjamins.
- University of Cape Town. 1999 (revised 2003). Language Policy. Cape Town: University of Cape Town.
- University of Cape Town. 2003. Language Plan Towards a Language Plan for the University of Cape Town: 2005-2010. Cape Town: University of Cape Town.
- Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Ed.). 1984. LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983: 13-30. Tübingen: Max Niemeyer.