

The Kernerman Dictionary

Three applications for the Kernerman Dictionary Research Grants have been accepted for 2002: The Assessment Committee of Afrilex awarded a grant to Karen Hendriks, from the University of Stellenbosch, for her research on the particular needs of a multilingual society and their effects over the structure and nature of bilingual dictionaries, and a grant to P.M. Lubisi, from the University of Pretoria for his project to build a major corpus for siSwati, one of the official Bantu languages of South Africa. The Asialex committee awarded a grant to Jacqueline Lam Kam-mei and her team from Hong Kong University of Science and Technology for their project to create a glossary of academic vocabulary.

A Glossary of Essential Academic Vocabulary

Jacqueline Lam Kam-mei with Sue Chang and Gregory James

Research has shown that word-lists have an important role to play in acquiring a language other than the mother tongue. At the university level, students need to be equipped with a sufficient vocabulary to follow lectures, comprehend academic texts, participate in seminar discussions, and write coherent research papers. Learners need to be familiar with at least 95% of the running words in a text if they hope to achieve these ends. Failure to acquire the core academic vocabulary for their field will adversely affect their work. A well-selected word-list based on academic texts used in tertiary institutions is therefore essential to help new entrants to acquire as many vocabulary items as possible in a relatively short time, so that they can cope with the heavy university student load once they have enrolled.

Based on the afore-mentioned rationale, a team of researchers, directed by Gregory James, at the Language Centre of the Hong Kong University of Science & Technology, have compiled five properly sampled electronic datasets of texts extracted from first-year textbooks used by university students in Hong Kong since 1991. These corpora, which have now been completed, comprise c.1,000,000 words in each of the following disciplines: (i) Computer Science, (ii) Business Studies & Economics, (iii) Biology, Biochemistry & Chemistry, (iv) Engineering and (v) Humanities & Social Science.

The proposed project represents a coherent and long-term strategy aimed at providing undergraduates and potential university students with an academic word-list which will enhance their knowledge of academic vocabulary, thus improving their ability to cope with a three-year university programme.

The project aims, firstly, to identify a list of frequently occurring academic words appearing across the five corpora mentioned. This list (hereafter the HKUST Wordlist) will contain those content words most useful for academic reading comprehension, discussions and paper writing. The second aim is to develop a glossary, based on the HKUST wordlist, that will include key features such as definitions, examples in context and examples of usage informed by relevant corpora and pronunciation. Other features will be considered and included subsequently as a result of feedback from user-perspective surveys.

All tertiary-level students, but first-year undergraduates in particular, in Hong Kong would benefit from this project through a systematically planned vocabulary learning scheme based on the HKUST Wordlist. Another group benefiting

from this project would be students in Secondary Forms 6 and 7 intending to pursue further studies at tertiary level in the Science and Technology streams. Through using the proposed HKUST Wordlist, students could better equip themselves to make the often difficult transition from general or secondary academic reading/writing to tertiary academic reading/writing. More importantly, the proposed HKUST Wordlist will serve as useful reference material for language materials designers, when seeking to produce appropriate language learning and teaching materials to enhance their students' language proficiency.

The project has three main phases:

Phase One:

1. Identifying and generating word-lists including single words, compound words and distinctive phrases (bi-grams, tri-grams, etc.) in each corpus;
2. Identifying words and n-grams that appear in all five corpora and generating a unified word-list across five disciplines;
3. Devising and implementing measures to confirm the reliability of the word-list compiled.

Phase Two:

4. Designing and compiling the prototype HKUST Wordlist in both paper and electronic form by
 - 4.1 determining headwords;
 - 4.2 writing definitions;
 - 4.3 illustrating definitions through appropriate phrases and sentences from the corpora;
 - 4.4 exemplifying word usage;
 - 4.5 including phonetic symbols for pronunciation help;
5. Designing and engaging in user-perspective surveys;
6. Collating and analysing survey feedback.

Phase Three:

7. Designing and compiling the HKUST Wordlist;
8. Developing training materials and vocabulary acquisition workshops and seminars, based on the HKUST Wordlist, to ensure students learn the words.

Once the framework for the HKUST Wordlist has been developed, it is envisaged that word-lists for learners at different levels and disciplines, for example, more discipline-specific lists for undergraduates, could also be compiled for use.

Research Grants

Certain Aspects of Bilingual Dictionaries in Multicultural and Multilingual Societies

Karen Hendriks

South Africans live in a culturally and linguistically diverse society with eleven official languages. The country has a complex political history, which cannot be severed from sensitive issues concerning language and communication. The process of acknowledging the truth and working towards reconciliation has been on its way for several years, yet the people of South Africa still have a long way to go in order to be able to genuinely celebrate their diversity rather than be threatened by it. I believe that lexicography, and specifically, good bilingual dictionaries, can make a vital contribution to the process of reconciliation, of affirmative action, and of building a nation out of the diverse speech communities and cultures in South Africa.

Bilingual dictionaries play an extremely important role in a multilingual society; they can be perceived as the key instruments in the communication between different communities. Efficient and active communication between different cultural groups is essential in the South African environment. Culture-specific lexical items, and the way they are treated in bilingual dictionaries, can have a great influence on this process. In the past, South African dictionaries have displayed a strong cultural bias in the treatment of these items. Today, lexicographers have to face the reality of the lexicographic wrongs of the past, and correct this in dictionaries for the present-day user.

My study will focus on the way the particular needs of a multilingual society imply certain adaptations to the structure and nature of a bilingual dictionary. In a multilingual and multicultural context the lexicographer may not rely on the intuition of the user, and the needs of different user groups have to be accommodated. The lexicographer has to keep in mind that the user of a general-purpose bilingual dictionary may well also be a learner of the target language.

Furthermore, I plan to examine the admission and treatment of culture-specific lexical items in bilingual dictionaries intended for a multilingual environment. Considering the vitality of enhancing clear and effective communication between speakers of different languages, the misrepresentation of culture-specific lexical items in bilingual dictionaries could seriously impede communication rather than enhance it. It is of great importance that lexicographers have adequate guidance in the treatment of these items.

South Africa is not the only country in the world with a multilingual and culturally diverse society. Users worldwide could benefit from a comprehensive study of the way dictionaries can contribute to and enhance communication between different cultural groups. I intend to examine the situation concerning this issue in South Africa and then to make suggestions and provide guidelines for the treatment of culture-specific lexical items in bilingual dictionaries that would apply to any multilingual environment.

Corpus Building for the SiSwati Dictionary Unit

P.M. Lubisi

SiSwati is one of the 11 official languages in the Republic of South Africa, spoken by about one and a half million people. It is, however, one of the least developed and underprivileged of these languages. There is not a single monolingual dictionary in siSwati, and the terminology development of this language is in its infancy. The first matriculants in siSwati were in 1987, and there is still only a small number of published materials, most of which are based on the school curriculum. SiSwati is offered at only two universities in South Africa, and no teacher training college is offering it.

Although the government, through the Pan South African Language Board, is in the process of establishing lexicography units, languages like siSwati will still be disadvantaged because of the lack of human capacity and resources. The siSwati lexicography unit is housed in a technicon that does not offer the language as a course, and this unit will receive little academic assistance. Lexicography is foreign in many technicons, including Pretoria Technicon, let alone corpus building. SiSwati has no professional lexicographers, and there are hardly any students who hold senior degrees in this language. Hence, it is developing at a snail's pace. Assistance is afoot in the collection of printed data, yet no attempt has been made to collect oral data from the siSwati radio stations, so a grant of this nature can play a pivotal role in the building of such a corpus. Moreover, it will allow enough time to do thorough research in this regard.

The aim of this project is to carry out a thorough research on corpus building for siSwati, by way of collecting data, especially oral data and information stored on tapes at the only two radio stations, namely Ligwalala FM in South Africa, and the Swaziland Broadcasting Corporation. There is a mammoth task to be done to retrieve this information as the unit does not at present have the necessary software to deal with this matter. The first stage of the project will concentrate on the oral literature, and will be conducted in the rural part of Mpumalanga lowveld, and in the four districts in Swaziland. A good tape recorder will be needed to record the oral data. The researcher will be assisted by veterans like Professor Daan Prinsloo from the University of Pretoria, who have already made inroads in this field. The university's department of African languages has accepted this project, and is actively participating in the building process and offers full access to all its resources and expertise. In addition, a major publishing house (Macmillan) has agreed to permit the use of their siSwati material. The collection of oral data is expected to gain full momentum by the receipt of the grant. The study is invaluable for the development and compilation of siSwati dictionaries in that it will be the first of its kind in this language. The gained know-how will be used also to assist other units that are still struggling to build their corpus.