

# Revealing Language Patterns Through the Use of Technology

**Shivaneeranjula Ilankoon**  
English Language Teaching Unit

Many students struggle to cope with the demands of English medium instruction as institutes of higher education are moving over to English medium instruction from mother tongue education in secondary education. As the Faculty of Arts, University of Colombo, too has been exploring the possibility of such a transition, the study aimed at bringing to light the recurring patterns in academic texts that students would encounter, in order to provide well targeted input data that would facilitate learning.

The study was carried out using a Corpus linguistics approach, which is gaining popularity with the technological advances in the field of computers. This approach investigates “actual patterns of language” in both spoken and written texts (Reppen and Simpson, 2002, p. 92) through the compilation of a corpus. A corpus is defined as “[a] large collection of authentic texts that have been gathered in electronic form according to a specific set of criteria” (p. 9). Furthermore, the guiding principles of a particular corpus are driven by the research question/s for which the corpus is compiled. Representativeness and balance, content and size are factors that need to be predetermined in the design and compilation of a corpus. Furthermore, the investigation of language through computers has greater advantages today as text data can be effortlessly analyzed through corpus analysis software.

The study aimed to compile a corpus of written data from five different disciplines in the Faculty of Arts, University of Colombo, comprising reference reading material prescribed by the departments in order to investigate the relevance of teaching the two thousand high frequency words posited by Michael West and the “Academic Word List (AWL) of 570 words put forward by Coxhead (2000). These two word lists are expected to cover 80% and 10% of academic texts respectively. However, Hyland and Tse, (2007) question the reality of one “core list” that can be recommended for every field of study.

The objective of this study was to investigate the above criticism and verify the relevance of the AWL for teaching purposes. It also aimed at bringing to light the Keywords and collocations that occur in the fields of Sociology, Economics, History and International Relations, Political Science and Demography.

For analyzing the data the corpus analysis software “The Compleat Lexical Tutor” (CLT) was utilized. Three different types of word lists were obtained from the corpus using the above programme. The features ‘VocabProfiler’, ‘Keyword Extractor v.1’ and ‘N-Gram extractor’ of the CLT were utilized for this purpose.

It was revealed that out of the 570 Academic words, 548 words occurred in the compiled corpus demonstrating the value of exposing the students to the AWL. Around 300 Keywords were identified from each corpus, while many interesting collocations were identified in the five different fields of study. As words in the AWL, Keywords and collocations occurred in the compiled corpus with noteworthy frequency, the researcher recommends that classroom teaching should focus on the above aspects of vocabulary.