Creating search facilities for BASE and BAWE: the design of a web-based interface

Name: Professor Hilary Nesi
Position: Professor
Email: h.nesi@coventry.ac.uk
Address: Faculty of Business, Environment and Society, Coventry University, Priory Street, Coventry, CV1 5FB.
Tel: 024 7688 8984
Fax: 024 7688 8679

This paper introduces and demonstrates the beta versions of our web-based interfaces for the British Academic Spoken English (BASE) and the British Academic Written English (BAWE) corpora. It also discusses some key issues surrounding their creation and development. The interfaces employ the functionality of Sketch Engine, a powerful corpus query system developed by Adam Kilgarriff and Pavel Rychly, and provide collocation, pattern and frequency information, searchable by parameters such as academic discipline and genre. The fact that they will be freely accessible should encourage use of the corpus data by users from within and outside the corpus linguistics research community, and we would like to attract English language students, EAP tutors and materials developers to the site. This leads to tensions between the conflicting requirements of usability and meaningfulness, however. We started with a query language that was powerful but opaque, but we wonder whether attempts to simplify the search process (for example by creating drop-down menus of search categories) might encourage more superficial, perhaps pointless, queries, and lead users away from open-ended investigations. Our challenge is to provide sufficient support and information to enable non-corpus linguists to pose queries that will lead to meaningful answers. The final interface designs will be arrived at through consultation and trialling with all kinds of potential end-users, so we will welcome feedback at the conference.

Paper for Corpus Linguistics Special Interest Group (SIG) track