This paper aims to encourage teachers and researchers to explore the nature of university student writing through the BAWE corpus.

It is important first to understand the development, composition and classification of assignment texts in the BAWE corpus. The 6.5 million word corpus includes 2858 proficient texts from over thirty disciplines and four levels of study distributed evenly over four disciplinary groups. The texts, which were collected from four universities in England, have been classified into thirteen genre families: Exercises, Explanations, Critiques, Essays, Literature Surveys, Methodology Recounts, Research Reports, Proposals, Case Studies, Design Specifications, Problem Questions, Narrative Recounts and Empathy Writing.

A keyword comparison of the BAWE corpus with the BNC corpus of general English yields a potentially useful list of ‘academic’ vocabulary, including terms such as data, therefore, theory, due, analysis, whose use can be explored further by discipline, genre, level of study etc. as well as through identification of frequent forms and collocates.

Surprisingly, perhaps, the lemma I is used widely and is a top key word. It is approximately nine times more frequent in the BAWE corpus (925.1 ARF per million words) than in the BNC (108.8 ARF per million words). Such corpus findings require further investigation before they begin to make sense in terms of what we know about academic writing. This can be done through the filtering, searching for collocates, concordancing and sorting functions of SketchEngine. This is illustrated through an account of the language of argumentation in Essays across the disciplines from multiple perspectives.

About the Speaker:

Dr. Sheena Gardner is Reader in Educational Linguistics and member of the MOSAIC Centre for Research on Multilingualism at the University of Birmingham, UK. Her research focuses on language in educational contexts from a broadly systemic functional perspective. She has published on young multilingual learner classroom interaction; and the development of the BAWE corpus of student academic writing.

The BAWE corpus is freely available through SketchEngine and from the Oxford Text Archive. See [www.coventry.ac.uk/bawe](http://www.coventry.ac.uk/bawe) for details.