

‘...now it is matter is the scaffolding’: Semantic structure and corpora

Professor Marc Alexander, University of Glasgow

This talk discusses the question of what it means to analyse linguistic data – in this example, historical textual data – from both the traditional standpoint of words and also from the perspective of semantic categories. We frequently explore lexis through ever more complex computational means, while our understanding of meaning in these databases has lagged behind the technical advancements of clustering words together into concepts.

In this talk, I will provide a case study using two resources I work with (the Historical Thesaurus of English and the Hansard Corpus), to apply semantic knowledge at a higher scale. In particular, words in semantic categories which are at the top of the frequency list (when ignoring the so-called ‘closed class’ of words) act differently to those elsewhere in terms of how they structure discourse. (These include adverbs which represent the traditional relationships of place, time, circumstance, manner, degree, and cause; common discourse concerns which are widely lexicalised and when clustered show the containing concept’s high frequency; and categories whose relative internal distribution expresses information about the text as a whole. I will argue that these semantic domains carry meaning not in terms of the aboutness of a text, but rather are indexes of discursive frames that scaffold and shape texts.