

Abstract

Lexical bundles in academic prose were originally identified in a 5.3 million word corpus of consisting of 75 book extracts and 33 research articles from 15 academic disciplines (Biber et al 1999; Biber et al 2004). This talk argues that the fragmentary nature of Biber's corpus data – collected as a record of register – raises questions about the pragmatic functions, such as discourse organising or attitudinal stance, assigned to these lexical bundles. For example, would a lexical bundle which organises the discourse of introduction sections be found more often in a corpus containing short full texts than in a similar-sized corpus containing long or fragmented texts? Or would the size of the corpus be sufficient for the heterogeneous nature of the data to have no effect on the frequency of occurrence or discourse function of the lexical bundles found?

In order to determine the effect of corpus composition on the occurrence of contiguous word combinations similar to lexical bundles, this paper introduces a methodology for comparing “absolute” and “pragmatic” frequencies of these word combinations. It presents results from two versions of a corpus of research articles from five journals in eight academic disciplines. One version contains a million tokens from each of the five journals in each discipline, so that occurrences of lexical bundle candidates can be compared across similar numbers of words in each discipline. The other, parallel, version of the corpus contains 500 texts from each of the five journals in each discipline, so that occurrences of word combinations can be compared across similar numbers of texts in each discipline. The implications of the results for research on lexical bundles are discussed.

References

- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan (1999) (eds) Longman Grammar of Spoken and Written English. London: Longman.
- Biber, D., Conrad, S., and Cortes, V. (2004) If you look at ...: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics* 25(3): 371–405

¹ *e-mail*: d.j.oakey@bham.ac.uk