

© Copyright Michael Stubbs 2015.

This is an abstract of an article in:

In N Groom et al eds. *Corpora, Grammar, Text and Discourse*. Benjamins.

THE TEXTUAL FUNCTIONS OF LEXIS

Michael Stubbs

This article reviews studies which show how lexis contributes to textual organization. It concludes that there are many interesting case studies, but that some hypotheses could be formulated more explicitly, and that findings could be better integrated into a functional theory of lexis.

Since the 1960s and 1970s, corpus studies have provided substantial quantitative information about the frequency and distribution of words, n-grams and collocations across large computer-readable text collections, and therefore valuable evidence of lexico-grammatical units of meaning in the language system as a whole and in particular text-types (e.g. academic articles and informal conversation). Nevertheless, such studies have limitations.

Corpora are objects which are produced artificially by linguists. Some corpora contain whole texts, others contain only text fragments, but in either case the texts have been selected by the linguist, usually according to a sociolinguistic theory (e.g. of social or stylistic variation) in order to investigate certain aspects of the language (e.g. to contrast different spoken or written genres). It is texts which are produced naturally by language users, and it is therefore texts which should be the real object of study. But when quantitative methods are applied to corpora, this typically removes fragments of language (e.g. words or concordance lines) from their original texts, and displays them in alphabetical or frequency order. This provides evidence of formal patterns in the system (e.g. in specific genres), but does not explain the function of the patterns in individual texts.

At least part of this explanation must relate to textual organization, and, since around the year 2000, many studies have shown that quantitative corpus methods can identify mechanisms of textual cohesion. The article reviews some of the main findings from this work, and then provides a short case study which illustrates that quantitative data on the distribution of words within and across texts (e.g. their frequency of co-occurrence with other words, or their frequency in different text-types) can be interpreted as indirect evidence of lexical function.

As a short illustration of the detailed descriptive work which is still required, the article provides a case study of the textual functions of one lemma (ALLEVIATE).
