

## Theoretical linguistics and corpus linguistics

Corpus analyses have had a considerable influence on the ways linguists think about language data and the way in which they view language in use. In this paper we explore the relation between theoretical linguistics and corpus linguistics. The main contributions of corpus analyses are: (i) highlighting the pervasiveness and the range of functions of collocations and other syntagmatic units with a lexical component; (ii) the provision of frequency information; and (iii) quantifying the extent of variability in language. We will explore the theoretical consequences of these findings, exploring the role of frequency information in particular.

The principal theoretical contribution of corpus linguistics within what we might loosely call the British tradition has been the elaboration of phraseology as an important component of grammar. Associated with the detailing of the work at the phraseological level has been the highlighting of the extent of the formulaic or collocational component of language. One important strand of this work can be traced from Firth (1957) to Sinclair (1991) and Hunston and Francis (2000). In *Verb Patterns*, Francis et al. (1996) describe a variety of phrasal patterns such as V N N and V AMOUNT and relate those patterns to lexical items grouped according to semantic categories. This approach, named Pattern Grammar (Hunston and Francis 2000), is the most extensively elaborated grammatical framework that follows this British tradition. (See also Mauranen and Sinclair 2006.)

The American tradition is characterized by the adoption of various forms of structuralism, which have generally not favoured a corpus approach, although even some generative theories have recently been influenced to some degree by corpus findings. Perhaps the most natural link between theoretical linguistics and corpus linguistics lies in what is called usage-based theories to language (Barlow and Kemmer 2000).

### References

- Barlow, M. and S. Kemmer (eds), 2000. *Usage-Based Models of Language*. Stanford: CSLI
- Firth, J.R. 1957. *Papers in linguistics* (1934-1951). London: Oxford University Press
- Francis, G. S. Hunston and E. Manning. 1996. *Grammar Patterns 1: Verbs*. London: HarperCollins
- Hunston, S. & G. Francis. 2000. *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam: Benjamins.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J and A. Mauranen. 2006. *Linear Unit Grammar: Integrating speech and writing*. Amsterdam: Benjamins