
When Professor Murray and all his assistants and voluntary readers created the first edition of the Oxford English Dictionary it took 70 years and involved more than six million slips of paper and Murray even had the floor of his office strengthened to be able to sustain the tonnes of paper (Winchester: The Professor and the Madman 1999: 103, 109). The usefulness of large electronic corpora to lexicography cannot be questioned. This usefulness, however, is not limited to the world of lexicography, but as *Words and Phrases* by Michael Stubbs demonstrates also applies to the field of lexical semantics. The status of the lexeme as the traditional unit of meaning is severely questioned and connotative meaning is put forward as a greatly neglected subject in comparison with that of denotational meaning (though even at the time of Professor Murray it was acknowledged that meaning is more than static, denotational meaning. Dr Trench talks about “the twists and turns of meanings” and describes “the way almost every word slips in its silvery, fishlike way, weaving this way and that, adding subtleties of nuance to itself, and then perhaps shedding them as the public mood dictates” (Trench in Winchester 1999: 105)). *Words and Phrases* suggests that what is at present known about connotative meaning is only the top of the iceberg and the advent of corpus linguistics will help to reveal more.

To whet your appetites let me start by declaring that *Words and Phrases* is a very good read. It has a user-friendly structure and each chapter is followed by “Summary and Implications”, “Background and Further Reading” as well as “Topics for Further Study”. By adding these sections Stubbs is able to accommodate all levels of background knowledge and the book is readable and relevant for researchers, students and teachers alike. Central to the book is the study of meaning and how the meanings of words depend on their different uses. To this end corpus linguistics is introduced and applied and Stubbs convincingly demonstrates how large corpora can be searched for observable patterns which provide evidence of what words mean. Although Stubbs in no way rejects intuition as a valuable starting point and has nothing against the odd introspective example he rejects intuition and introspection as generally reliable tools: “In many areas of semantics and pragmatics, intuitions are strong and stable, across all native speakers, whether linguistically naive or trained and must be given the status of data” (p. 71). It is not necessary to demand corpus evidence of all lexical relations (e.g. everybody would probably agree that ‘naked nude’ is redundant). “However, there are many cases in the literature where the intuitions of native speakers are less certain, or where intuitions are demonstrably unreliable or just missing altogether” (p. 71). Sometimes intuitions about connotations are
very reliable, but at other times they are not. “It may also be that intuitions about the core meaning of a word are reliable, but that intuitions about its potential use in different situations are not” (p. 72).

The majority of examples in the book are so-called ‘attested language’, i.e. the author has not invented suitable examples, but relies almost entirely on authentic corpus-based examples (in contrast with Lyons – the author claims that in Lyons 1977 (more than 800 pages) there is not a simple authentic example). Stubbs argues that a theory of semantics should deal primarily with normal cases: what does typically occur, not what might occur under strange circumstances. Much linguistics is concerned with “What people can say” (traditional linguistics) and not with “what people do say” (corpus linguistics) (p. 61). The arguments and methods of the book are very much based on those of Sinclair (see e.g. 1991, 1996 and 1998) and the scholars who have subsequently based their work on his initial findings. The work of Sinclair was a radical development of Firth and Stubbs himself consequently describes his own approach as ‘neo-Firthian’.

Apart from four pages of very useful “Notes on Corpus Data and Software” in the beginning of the book, it consists of three parts; Part I Introduction, Part II Case Studies and Part III Implications which are again divided into ten chapters. Two approaches characterise the book as a whole: 1. meaning is use, and 2. corpus linguistics can be used as the main evidence for the uses and meanings of words and phrases.

**Part I**

Chapter 1 is concerned with the importance of phraseology in the study of meaning. It aims at showing the close relation between how words are used and what they mean. In contrast to the old semantic principle of compositionality (the assumption that the meaning of larger units is equal to the sum of the meanings of smaller units) Stubbs argues that combinations of words in phrases are a good candidate for the basic semantic unit of language in use. “Instead of regarding the meaning as being carried by the individual word, we could see things as follows: It is not the words which tell you the meaning of the phrase, but the phrase which tells you the meaning of the individual words in it” (p. 14). Chapter 2 discusses the basic terms and concepts which are required for a systematic study of meaning. Most of these concepts are known from traditional semantics and they include denotation, connotation, lexical field and sense relations, but chapter 2 also introduces important terms and concepts from corpus linguistics such as ‘node’, ‘span’ and ‘collocates’. These terms and concepts are explained so that they can be used in the following chapters, but chapter 2 also aims at showing how attested data collected in large corpora can be used to
illustrate and develop the concepts in question (especially connotational meaning). Together chapters 1 and 2 argue that it is not individual words which are the basic units of meaning, but longer phrases and collocations – the so-called extended lexical units, as connotations are often not attached to a single word, but are distributed across several.

Part II
Chapter 3 is mandatory reading for the novice of corpus studies as it introduces and discusses the concepts, data and methods of corpus linguistics (which are also important to the case studies of the following chapters) and describes in more detail how patterns can be discovered in corpus data. With the tools of corpus studies it is now clear that phraseology plays a major part in our language use and thus in our understanding of the concept of meaning:

“Corpus studies show that what typically occurs in language use is only a small percentage of what seems possible within the language system. A large amount of language use consists of words occurring in conventional combinations. Such collocations are not an idiosyncratic and peripheral phenomenon, but a central characteristic of language in use. Native speakers’ unconscious knowledge of collocations is an essential component of their idiomatic and fluent language use and an important part of their communicative competence” (p. 73 – my emphasis).

Chapter 4 presents the findings of a case study of the phraseology of English and shows that the freedom to combine words in text is much more restricted than often realised. In chapter 5 Stubbs argues that “Corpora can be used to document the norms of language use, which are the background against which individual texts are interpreted” (p. 100). Corpus studies e.g. show that the Saussurean slot-and-filler model is inadequate and that syntagmatic organisation (i.e. the norms) is much stronger than often realised – “there are rarely, if ever, free paradigmatic choices of lexis” (p. 119). Predictable combinations constitute a large percentage of normal language use. The chapter furthermore demonstrates how extended lexical units contribute to textual cohesion. The main aim of chapter 6 is to show (on the basis of a case study of a short story) how computer-assisted methods can be used to study lexical patterns in texts and the validity of such studies is discussed.

As was argued in chapter 5 individual instances can only be evaluated on the background of usage norms and an important part of these is culture. On the basis of cultural keywords chapter 7 shows that collocations, even of very common words, often carry cultural connotations: they are a significant compo-
Chapter 8 illustrates how corpus linguistics may provide new data on etymology and word meaning.

Part III

Chapters 9 and 10 discuss some wider implications of corpus studies for linguistic theory. Chapter 9 centres on one of the major findings of corpus semantics up till now; The fact that “many more words and phrases have evaluative functions than is usually recorded in dictionaries” (p. 197). Within traditional lexical semantics denotation (and truth conditions) is considered the central part of meaning and connotation is a more peripheral phenomenon which in the words of Leech (1981 in Stubbs: 197) is described as “unstable” and “indeterminate”. Lexical semantics – even in its traditional form – offers problems aplenty so in retrospect it seems only natural that denotational meaning should be the starting point of serious semantic studies, but as many semanticists have realised, an account of denotational meaning merely offers us half the truth about meaning – and perhaps not even the most significant part. Evaluative aspects of meaning seem less accessible to intuition than non-evaluative meaning and this may very well be the reason why it has not hitherto been a focal point of interest to traditional (armchair) semanticists. Naturally, connotation is by no means a new phenomenon, but it has always suffered under predicates such as “idiosyncratic” and “slippery” and is only briefly mentioned even in major textbooks. Connotation is still a very challenging subject, but corpus linguistics offers the possibility of studying connotation in a more systematic way and research so far shows that a great number of connotations are not idiosyncratic in the least. On the contrary they are widely shared in a discourse community. In chapter 9 Stubbs furthermore briefly introduces a new and interesting area of research, namely that of identifying networks of inter-collocating words – i.e. a kind of connotational lexical fields. I am convinced that this is an area where we will see plenty of new and exciting results in the future. Chapter 10, the final chapter, deals with the principles and problems of corpus linguistics. The author claims that corpus linguistics is inherently sociolinguistic, diachronic and quantitative. The importance of frequency or proportions has not been sufficiently acknowledged in the past or, one might add, has been ignored because we did not have tools which were powerful enough to determine these factors. Chapter 10 states and discusses six very relevant and highly interesting objections to corpus studies and Stubbs goes on to discuss the limitations of the dualisms prevalent in Saussurian (langue/parole) and Chomskyan (competence/performance) linguistics. Stubbs argues in favour of a pluralist position and maintains that “Corpus study may eventually provide a way of avoiding unfortunate dualisms and of integrating different kinds of cognitive and behavioural data.” (p. 244)
It is only since the mid-1980s that computer-assisted methods have been able to provide evidence about word meaning by searching across large text collections. Now that we have the tools a whole new world of study opens to researchers within a large number of fields with a wealth of new information. Instead of merely relying on what can be said (which is of course very relevant to some kinds of study) we can now focus on what is actually being said and how often and in what situational context. Though on the face of it it may seem a simple matter of calculation, frequency is a complex matter especially when dealing with word combinations and patterns of meaning and cannot be ignored. “It is as if chemists knew about the different structure of iron and gold, but had no idea that iron is pretty common and gold very rare; or as if geographers knew how to compare countries in all kinds of ways, but had never noticed that Canada is bigger than Luxembourg” (Kennedy 1992 in Stubbs: 221). The challenge (and I consider it a major one) is now to find out how to interpret all this new information. From the point of view of lexical semantics I consider it a very important result of corpus studies that many more words have evaluative connotations than previously recognised. And as to the practice and theory of stylistics and of translating this also has important implications (as pointed out, but not further discussed by Stubbs on page 243).

*Words and Phrases* provides a wealth of relevant suggestions for further reading, it is highly up-to-date and very well researched and discussions are carried out at a high level. However, its logical structure, internal references and user-friendly language make it comprehensible to everyone with an interest in the subject and basic semantic/linguistic knowledge. To get the full advantage of the book it is no doubt necessary to have more than a basic knowledge of semantics and the history of semantics, but the book never mentions a theory, a concept or a scholar without proper definitions, introductions and references – the reader is certainly not left to his own devices. There are many repetitions throughout the book of its most important arguments, but instead of being annoying it merely serves constantly to remind the reader of the framework of the book and to put the analyses into perspective. It happens once or twice that the author generalises perhaps a little too eagerly on the basis of a few examples, but at the same time he often points out that much research is needed to fully confirm many of the observations of the book. On the whole I can only agree with the promising recommendation on the back of the book: “The author successfully demystifies his own discovery processes, providing the reader with tools for further investigation. The book’s clarity and depth make it indispensable for students and researchers alike” (Susan Hunston, University of Birmingham). It is not merely a marketing ploy – it is actually true and I fully endorse it!
References


Karen Korning Zethsen