

Sinclair Lecture and Seminar

Wednesday 2nd September 2015

To celebrate fifty years of research in English Language at the University of Birmingham, we are combining the annual Sinclair lecture with a one-day seminar devoted to corpus linguistics and its current and future applications.

The theme of the seminar is: Corpus Linguistics: Pattern, Variation, and Change. The seminar will focus on how corpus linguistics can be used to identify patterns in language, how these patterns vary across different genres and registers and how they change over time, in both language acquisition and language development.

Time	Speaker	Topic
9.30–10.00	Registration and Coffee	
10.00–10.10	Prof. Jeannette Littlemore	Welcome from the Head of Department
10.10–11.00	Prof. Susan Hunston	Keynote Talk Pattern Grammar Research: Past, Present and Future
11.00–11.50	Prof. Nick Ellis	Keynote Talk Usage-Based Approaches to Grammar Acquisition and Processing
11.50–12.30	Prof. Susan Hunston, Prof. Nick Ellis and Prof. Ute Römer <i>in conversation</i>	Pattern Grammar reassessed
12.30–1.45	Lunch	
1.45–2.15	Prof. Michaela Mahlberg	Discovering patterns with CLiC: Dickens and 19th century fiction
2.15–2.45	Dr Rosamund Moon	'New facts about words': Cobuild and the corpus revolution
2.45–3.15	Dr Nicholas Groom	The phraseology of knowledge
3.15–3.45	Coffee Break	
3.45–4.15	Dr Paul Thompson	Disciplinary and interdisciplinary variation in research discourse
4.15–4.45	Dr Mel Evans and Dr Caroline Tagg	Corpus investigations of language change: from Tudor letters to today's text messages
4.45–5.00	Break	
5.00–6.00	Prof. Ute Römer	The 10 th Sinclair Lecture Corpora, Constructions, Collaboration: Providing new insights into the use and acquisition of verb patterns in English

The Sinclair Lecture is sponsored by the CfBT Education Trust

Sinclair Lecture and Seminar 2015

Department of English Language and Applied Linguistics
University of Birmingham

Seminar abstracts

Pattern Grammar Research: Past, Present and Future

Susan Hunston

'Pattern Grammar' was developed during the 1990s as a way of capturing the interaction of lexis and grammar, as revealed by corpus investigations. It contributed to the development of COBUILD dictionaries and grammar reference books. This paper briefly reviews both the original research and related concepts, in particular Local Grammar, Construction Grammar and Corpus Pattern Analysis. Current activities such as the development of the Patterns On-Line website and the semantic annotation of patterns are described, and their implications discussed. The viability of future research in this area is assessed.



Usage-Based Approaches to Grammar Acquisition and Processing

Nick Ellis

Usage-based approaches believe that we learn language through the episodes of our using language. Our linguistic ability emerges as a result of our cognitive learning mechanisms analysing this experience.

Corpus Linguistics provides relevant evidence. Pattern Grammar is based on the identification of grammatical patterns in text. Cognitive Grammar is concerned with how people acquire, represent, and process this knowledge.

Our research analyses (1) usage patterns of English verb-argument constructions (VACs) in terms of their grammatical form, semantics, lexical constituency, and distribution patterns in large corpora, (2) patterns of VAC usage in child-directed speech and in child language acquisition, (3) psycholinguistic investigations of VAC processing. It demonstrates:

1. Language usage is highly patterned in ways that support learning.
 2. Language acquisition is guided by this patterning.
 3. Language users have rich implicit statistical knowledge of these patterns.
-



Discovering patterns with CLiC: Dickens and 19th century fiction

Michaela Mahlberg

This paper will explore the fuzzy boundaries between literary and non-literary texts. Corpus methods make it possible to see similarities between fictional speech and real spoken language. At the same time, corpus methods help us discover patterns and linguistic units that are specific to the way in which narrative fiction builds textual worlds, e.g. suspensions or lexical patterns of body language presentation (Mahlberg 2013). Such literary patterns also relate to features of the real world. Importantly, the study of the nature of literary texts highlights that we need to complement corpus linguistic methods with a range of other methods and interpretative frameworks, such as psycholinguistic research (Mahlberg et al. 2014), cognitive poetics (Stockwell 2009), literary criticism and approaches in social history. Drawing on examples from Dickens's novels and other nineteenth century fiction, this paper will argue for a mixed methods approach to the study of literary texts. The paper will specifically illustrate some of the functionalities of the CLiC tool that is being developed to support the corpus linguistic analysis of fiction as part of such an approach.



References

- Mahlberg, M. 2013. *Corpus Stylistics and Dickens's Fiction*. New York & London: Taylor & Francis
- Mahlberg, M., Conklin, K. and Bisson, M.-J., 2014. "Reading Dickens's characters: employing psycholinguistic methods to investigate the cognitive reality of patterns in texts", *Language and Literature*. 23(4), 369-388
- Stockwell, P. 2009. *Texture. A Cognitive Aesthetics of Reading*. Edinburgh: Edinburgh University Press.

Cobuild and the corpus revolution

Rosamund Moon

The Cobuild project shaped and dominated the linguistic landscape in the Department of English at Birmingham from 1980 until the mid-1990s, though its influence extended far beyond Birmingham and its legacy is incalculable.

Under the leadership of John Sinclair and funded by Collins Publishers, its immediate goal was to produce a new EFL dictionary; what made it different was its methods, since it began by examining *de novo* the current English lexicon, drawing on a new and very large



(for its time) corpus of English to analyse words and their usages. This led to a radical re-evaluation of the lexical behaviour of English – and not just a series of award-winning dictionaries, grammars, and other materials, but also, from linguistic theoretical perspectives, new understandings of phraseology, grammar, and meaning that challenged the orthodoxies.

This talk explains why the Cobuild project is counted amongst the ten most influential contributions to be made by researchers across the University of Birmingham in the last 100 years.

The phraseology of knowledge

Nicholas Groom



In his recent book *Discourse and Knowledge*, Teun van Dijk (2014: 311) observes that while “much, if not most, of human knowledge is acquired by discourse and communication ... the [study] of the role of discourse in the acquisition, reproduction and diffusion of knowledge is still in its infancy, it if has even been born yet.” In this paper I shall focus on an aspect of the relationship between knowledge and discourse that is curiously absent from van Dijk’s own theoretical account: the concept of phraseology. I will argue that phraseology, understood in Sinclairian terms as “the tendency of words to occur in preferred sequences” (Hunston 2002: 138), is crucial to the study of knowledge and discourse as it encapsulates the key idea that preferred ways of saying are also preferred ways of knowing. Drawing on examples from a number of recent studies, I will argue that the corpus-based analysis of phraseology should be seen as having a central role to play in both the empirical analysis of the relationship between knowledge and discourse, and in the development of theoretical models of this relationship.

References:

- Hunston, S. (2002) *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
van Dijk, T.A. (2014) *Discourse and knowledge: A sociocognitive approach*. Cambridge: Cambridge University Press.

Disciplinary and interdisciplinary variation in research discourse

Paul Thompson

Disciplinary variation has been a focus of research into academic discourses in the last decade, but little has been written about the features of variation in interdisciplinary (rather than 'disciplinary' research discourse). In this talk, I will discuss findings from a large scale ESRC project into interdisciplinary research discourse, based on a series of analyses of a corpus of journals that the scientific publisher Elsevier has provided access to. I will look in particular at linguistic and rhetorical patterns in the data, and how at they vary across time and journal.



Corpus investigations of language change: from Tudor letters to today's text messages

Mel Evans and Caroline Tagg

In this talk, we explore how corpus methodologies can aid our understanding of how people in different time periods have used written resources such as spelling for identity work and how these individualised practices changed within their own lifetimes. We investigate this by comparing and contrasting two datasets, an early twenty-first-century corpus of text messages (2004-2015) and an early sixteenth-century corpus of personal letters (1544-1556). Our findings reveal interesting similarities in practice despite divergences in form and in social context such as technology and surrounding ideologies of standardisation that have implications for our understanding of written language repertoires.



The Sinclair Lecture and Seminar event is sponsored by the CfBT Education Trust.