

To the memory of John Sinclair, Professor of Modern English Language

by Julia LAVID

With the death of John Sinclair English Language studies have lost a leading figure of world rank, a “giant” in the field. John Sinclair was an outstanding scholar, a brilliant mind, and clearly one of the most open-minded and original thinkers in areas such as discourse analysis, lexicography, and corpus linguistics, two of which he helped to create, contributing to many others. His writings and books are widely read and constitute compulsory reading for any serious language student.

After an initial appointment to a lectureship in the Department of English Language and General Linguistics in Edinburgh in 1958, John was elected in 1965 at the early age of 31 to the foundation chair of Modern English Language at the University of Birmingham, the post from which he retired in 2000. He fulfilled both posts with immense energy and drive, increasing the profiles of both departments immeasurably and building up their reputation by sheer application and scholarship. During his time at Birmingham the university's English department became internationally recognised as a world centre for English language studies.

At a time when English linguistics was excessively preoccupied with theory-building, John Sinclair began his work on ‘collocation’ and ran one of the very first research projects in computational linguistics. These two early steps testify to his nature, provocative and innovative.

His definition of collocation -the tendency for words to associate with each other in partly arbitrary ways- is now deemed central in descriptive linguistics but was certainly challenging for the then current theories of language.

As a founder of the ground-breaking COBUILD project in lexical computing, whose name stands for ‘Collins Birmingham University International Language Database’, John revolutionised lexicography in the 1980s by proposing a new kind of dictionary for advanced learners of English, which involved the creation of the largest corpus of English language texts in the world. As explained in the web page of the University of Birmingham English Language Research site, the new research project was at the time “the largest single research project the University of Birmingham had ever had; one of the first ever text scanners was bought at a cost of £70,000 and impoverished students worked day and night scanning in texts”. The result of this massive effort was a corpus of amazing size for that age, some eight million words of contemporary English text (the Bank of English), and the production of the *Collins COBUILD English Language Dictionary* (1987, 2nd edition 1995, 3rd edition 2001, 4th edition 2003, 5th edition 2006), based on the study and analysis of the COBUILD corpus. Since then, a “steady stream of corpus-based dictionaries, grammars and usage books followed, based on principles which have

radically changed the way all publishers produce foreign-learner reference books” (see web page at <http://www.english.bham.ac.uk/research/events/sinclair/series.htm>).

In her posting to the Sysfling-list and Sys-func distribution lists on the occasion of John’s untimely passing, Louise Ravelli, a COBUILD scholar, describes John’s truly innovative nature, and the magnitude of the task which he undertook at that time:

He was truly innovative in what he brought to Birmingham University, lexicography, and linguistics. For those of you who are young enough to think that computers quite naturally can sit on your lap and go anywhere with you, I would like to point out that in the 1980s, when COBUILD was underway 'the' mainframe computer sat in its own special room, about the size of a fat telephone booth, and whoever was in first in the morning had to 'boot it up'. When COBUILD wanted to do something 'big' in terms of computer processing, the job had to be run over the weekend, and the *entire* University network was switched off, to accommodate it. Just over 20 years ago, not exactly ancient history but nonetheless an immensely different era. It was John's vision and understanding which enabled all this to happen, and which provided so many opportunities and insights for us as a result. (Posting at Sys-fling, 14/03/2007)

Whoever is familiar with the *Collins COBUILD English Language Dictionary*, an extraordinary piece of academic research, containing a phenomenal 56 million words of contemporary and written text, knows that it is an invaluable research tool that should be at every grammarians’s right hand. Robin Fawcett says in this respect: “I remember that, when I first turned the pages of my copy, I realized that it would be a research tool that I'd ensure I had by my desk for the rest of my working life. Like Quirk et al 1985 (and now Biber et al 1999), it is invaluable as a quarry in which to go looking for relevant data. And with the addition of the diamonds indicating frequencies in the second edition, it has virtually double the value.” (Posting at Sys-fling, 14/03/2007)

The *COBUILD Dictionary* was innovative in many respects that set it apart from other learner’s dictionaries of English language. By its own account, "the techniques used to compile [the dictionary] are new and use advanced computer technology. For the user the kind of information is different, the quality of information is different, and the presentation of information is different" (from the introduction to the *COBUILD Dictionary*).

Though John Sinclair gained his widest renown as a lexicographer, his expertise stretched far wider, based on his comprehensive knowledge of linguistics. Thus, in 1975 he published, with Malcolm Coulthard, one of his greatest works, *Towards the Analysis of Discourse*, the first truly systematic account of a genre of spoken discourse. Using Halliday's description of grammar as a hierarchy of units, Sinclair and Coulthard created a model for spoken discourse analysis, developed as a tool for systematic study of classroom discourse, concentrating mainly on interactions between the teachers and individual students. Slightly revised by its authors in 1992, and used extensively by a number of scholars who have modified it to account for less structured discourse patterns, the Sinclair and Coulthard model, sometimes called Discourse Analysis (DA), has contributed immensely to our understanding of the structure of discourse beyond the sentence.

His interest in language teaching was not negligible either, as Robin Fawcett rightly reminds us:

I remember him addressing a group of teachers of linguistics in tertiary education in Britain (which flourished for a while, in the 1980s). John is the only person I've met who has made the point, quite explicitly, that if you are going to teach an undergraduate course about language - or the English language - the students need, AS A MINIMUM, about 200 hours. (That translates, in today's terms, into at least three 10-week modules of two lectures and one seminar a week for each student.) This was a rough guide that I have followed all my teaching life, and that we implemented most completely - with a range of further options that took us far beyond that minimum - when Paul Tench, Gordon Tucker and I were all teaching together at Cardiff University. Many of those reading this will be working in institutions which ask you to do your job in less than 200 hours. Well - I have tried that, and I have even moved institutions to ensure that I could do the job properly. I reckon that John got it about right. (Posting at Sys-fling, 14/03/2007)

His second masterwork, *Corpus, Concordance and Collocation*, was published in 1991, and soon became “the bible for a generation of corpus linguists”, in the words of one of his colleagues in Birmingham, Michael Hoey. The book reveals new understandings of how English works leading to new descriptions of the language. The two key notions are ‘collocation’ and the ‘idiom principle’. Collocations are the words that are placed or found together in a predictable pattern. Examples range from two word combinations such as *naked eye* to extended combinations such as *He's recovering from a major operation*. These language patterns comprise much of speech and writing, having far-ranging implications for a proper description of English and for second-language learning. Here too he proposed his revolutionary ‘idiom principle’, the tendency towards idiomaticity in language, where words tend to go together and make meanings by their combinations. For him, units of meaning are largely phrasal and phraseology is due to become central in the description of English.

In 1995 John took partial early retirement from the University of Birmingham and moved to Italy, where he founded and run, with the help of Elena Tognini-Bonelli, his second wife, the Tuscan Word Centre, an internationally recognised association devoted to promoting the scientific study of language. The centre made its name by offering short, high-level courses for language researchers and workers in the language industries, with the participation of distinguished topic leaders in the field of corpus linguistics. One of those topic leaders, Üte Röhmer, has recently written the following statement: “on the short intensive courses that the Tuscan Word Centre offered, John very generously shared his original ideas about language and linguistics with generations of younger scholars, introduced numerous students to the fascinating world of corpora and inspired many new ideas for future research in linguistics”. (Linguist List 18.835, 19/03/2007)

A flow of influential publications appeared after his early retirement from the University of Birmingham. *Reading Concordances* (2003) is an invaluable practical handbook with tasks that introduce the user to simple and reliable ways of extracting information about language from a corpus. But it was with *Trust the Text* (2004) that Sinclair extended and refined his original theories and analytical practice of the

1990s into the twenty-first century. In this book, considered as his third masterwork, Sinclair argued that linguistics has relied very heavily on speculation and inadequate evidence, probably due to shortage of data, but that now that there are large quantities of data available – particularly written language in electronic form but also substantial quantities of spoken language– our theory and description should be re-examined “to rebuild a picture of language and meaning which is not only consistent with the evidence but exploits it to the full” (p.10). On the basis of this available evidence, Sinclair proposed a model of language based on extended units of meaning that involve larger patterns than the individual words. To account for these patterns he proposes structural categories such as ‘collocation’, ‘colligation’, ‘semantic preference’ and ‘semantic prosody’, whose study and interrelationships, he claimed, will assume a central rather than a peripheral role in language description.

His last book, *Linear Unit Grammar* (2006), co-authored with Anna Mauranen, challenges current language descriptions which, in his view, overemphasize the hierarchical patterns of language, by proposing a linear model, in that it avoids hierarchies, and concentrates on the combinatorial patterns of text.

John was an honorary life member of the Linguistics Association of Great Britain, a member of the Academia Europaea, the holder of an honorary doctorate from Gothenburg University, and honorary professorships at Glasgow, Shanghai and Jiao Tong Shanghai universities. But beyond all that, John was a very special and original mind. Robin Fawcett puts it this way:

I think that what I shall remember about John, as a presence, is the combination of wise scholarship based on a close familiarity with language (‘Trust the Text!’), a slightly mischievous sense of fun, and a healthy Scottish scepticism that often found expression in a slightly raised eyebrow. For John, it was always important to challenge ‘establishment’ ideas - even when the ‘establishment’ was itself a challenge to the larger establishment of formal linguistics. And, perhaps above all, he always brought out the fun in studying language. (Posting at Sys-fling, 14/03/2007)

Those of us who had the privilege to meet him and exchange ideas with him know how inspiring and thought-provoking his conversation was. He was a rare mixture of intellectual creativity and personal kindness. The memories and views of the colleagues cited in this obituary, and many others not quoted here, reflect the widespread esteem in which John Sinclair was held.

He is now gone and will be sorely missed. But, perhaps, he is not gone forever. His intellectual legacy will remain with us forever. May many younger generations have the pleasure of discovering the fascinating world of language by reading John Sinclair’s writings!

Dr. John McHardy Sinclair, academic and lexicographer, was born on 14 June 1933 and died on 13 March 2007. He leaves his second wife, Elena Tognini-Bonelli, and their two children, as well as three children from his first marriage to Margaret Lloyd in 1956.