

**INTERVIEW: GEOFFREY LEECH**

## **‘LEAVE NO STONE UNTURNED IN THE SEARCH FOR LINGUISTIC REALITY...’**

by Jelisaveta Milojević



Geoffrey Leech (born 16 January 1936) was Professor of Linguistics and Modern English Language at Lancaster University from 1974 to 1996. He then became Research Professor in English Linguistics. He has been Emeritus Professor in the Department of Linguistics and English Language, Lancaster University, since 2002. Professor Leech's main academic interests

have been: English grammar, semantics, stylistics, pragmatics, corpus linguistics and corpus-based natural language processing. He is a Fellow of the British Academy. To mark his retirement the Department at Lancaster set up the Geoffrey Leech Scholarship fund for MA students.

Professor Leech kindly took time to give this extensive interview for the second issue of *Belgrade BELLS*.

**BELLS:** You have written, co-authored over 25 books (and more are in preparation), and over 100 papers. The numbers are quite impressive. It is common knowledge that one cannot make a fortune on books and we can therefore assume that there must be (have been) a motive of some sort other than the money. I wonder what linguistic creed or message that you have wanted to get across has kept you so avidly devoted to writing on and about language?

GEOFFREY LEECH: Well, I guess very few academics are motivated purely by gain – and yet very few are totally uninterested in making a living! I cannot help thinking how lucky I am, to have pursued a career where I can get paid for doing something I enjoy – something like the best hobby you can have, full of interest and challenge.

I cannot claim any great moral mission or creed has spurred me on ‘to scorn delights and live laborious days’ (as Milton puts it in *Lycidas*). No, I’ve simply been an opportunist – and I was lucky that in my early days the academic world was much less crowded with people and ideas than it is today. After doing a bit of school-teaching in the early 1960s, I was lucky to get back into academic life – I have to thank Randolph Quirk at UCL for that. In those days, nobody asked a young aspiring scholar ‘Have you got a PhD? How many publications have you produced? How many are in peer-reviewed journals? How much lecturing have you done? What are your plans for winning research funding? No - Quirk interviewed me in an amiable fashion, and the next thing I knew was that I was offered a job as Assistant Lecturer at UCL (University College London) - at the best centre, as it happened, for studying and teaching English linguistics in the UK.

My areas of academic interest mostly grew out of the opportunities I got. The first new task they gave me at UCL was to teach a lecture course on ‘Rhetoric’ – the lecture topic I had found to be the most boring when

I was an undergraduate. But I took the opportunity to teach rhetoric in a ‘modern linguistic’ fashion, and out of that came my deep interest in the marriage of linguistics and literature – stylistics, as it is usually called today – and I wrote a book called *A Linguistic Guide to English Poetry*.

I could tell similar tales about the beginnings of my research interests in English grammar, in the English verb and modality, in semantics and pragmatics. There was no great mission – only opportunity.

**BELLS: When you were young you worked with Randolph Quirk. In what way was his influence significant?**

GEOFFREY LEECH: Well, as I’ve already explained, Randolph Quirk offered me my first chance of a university career. Without him, who knows where I would be? But, also, of course, RQ more or less invented the idea of a modern English ‘corpus’. The department I was teaching in happened to be the department where Quirk’s Survey of English Usage had started up. I couldn’t help being gripped by the fascination of recording and collecting real language data, in a project led by Quirk’s charismatic zeal. I mixed with RQ’s research acolytes – they included Jan Svartvik, David Crystal and Sidney Greenbaum – and when the opportunity came to move to a new university, I was soon starting a corpus of my own – this time with the help of the computer. But that’s another story.

My other debt to Randolph Quirk came from his leading role in developing the ‘Quirk grammars’. Strangely enough, this began not with RQ at all, but (as I remember it) with Sidney Greenbaum and myself lamenting the vast gulf between grammar theorizing within linguistics (think of the types of formal grammars that were available in the 1960s) and grammar teaching within EFL. There was virtually no connection between the two, and we thought that we could try to write an English grammar which mediated between the two – a kind of grammar which was informed by theoretical advances and which built on the work of the Survey – but at the same time was going to be useful to the EFL student. We enrolled our close colleague Jan Svartvik (who had by then returned to Sweden) as a potential co-author, and went to RQ seeking his involvement in the project, hoping that he would take a kind of editorial role or supervisory role in vetting our work. It was typical of him that he not only accepted the idea, but embraced it wholeheartedly, insisting on taking the part of a full author. Out of that came *A Grammar of Contemporary English* (1972) and

later *A Comprehensive Grammar of the English Language* (1985). Quirk's energy was the leading inspiration of the grammar project, and since then *GCE* and *CGEL* have been justly known as 'the Quirk grammars'. In those days before e-mail, to write and agree the final version of the grammar we had to get together in a single room – Quirk's office in UCL – to hammer out every paragraph. We virtually lived together for six weeks. The second grammar was hundreds of pages longer, so we needed even more time of co-habitation – eight weeks – and by that time RQ's position had become elevated to that of Vice-Chancellor of the University of London. In this capacity, he occupied a rather grand terrace house in Gordon Square, near where the Bloomsbury Group used to meet, about a stone's throw from UCL. In the front room we thrashed out the intricacies of the adjuncts, subjuncts, disjuncts, conjuncts and all the rest... The remarkable thing was that somehow RQ continued his deep involvement in the project for most of the summer vacation of 1983, while running the University of London in his 'spare' time.

**BELLS:** When one takes a look at the history of linguistics one sees that there have always been those brilliant and fluent in ideas (like Chomsky or yourself trend-setters if you like), and those who are just able to follow the trodden path. You left Quirk fairly early in your academic career and decided to take a less travelled way. You engaged yourself in what was totally new at the time: corpus linguistics. Please tell us more about what ideas about (and beyond) language made you make such choice.

GEOFFREY LEECH: Thank you for conjoining me with Chomsky – I'm deeply flattered – but I'm certainly far inferior to him as a 'trend-setter', and all my academic instincts have directed me in the opposite direction from Chomsky. By the way, I met him in the 1960s when I was a MIT as a Harkness Fellow, and found him very pleasant socially. But on returning to the UK in 1965 I drafted an article criticizing his three levels of adequacy (observational, descriptive, and explanatory) and sent it to him for comment. Instead of throwing my efforts in the waste-bin, he honoured me by writing about 14 pages of closely-typed refutation. I found him, in academic debate, an implacable opponent – quite different from his mild and amiable social self. But I still went ahead and published the article.

Yes, one of the ways I took the opposite road to Chomsky was in seeking the empirical evidence of language use, wherever possible, to back up linguistic claims – and to enlist the help of computer technology – although it was very primitive at the time.

It happened like this. In 1969 I was invited (by Norman Fairclough – is the name familiar?) to apply for a job at a very new University at Lancaster. It meant moving from the metropolis to the north of England – to the fringes of civilization, as it seemed then – to a university with virtually no research record. But it was a promotion – I would have had to stay at UCL for a very long time to gain such a job. And, once again, it was an opportunity: it is very much easier to start new things in teaching, new things in research at a new university. At Lancaster I joined a small group of young English language academics eager to experiment and prove themselves. We sat round a table one day and asked ourselves a question: What can we do in linguistic research, to put Lancaster on the map? With my experience at UCL, I suggested that we start a new computer corpus of British English. I had met Nelson Francis, creator of the million-word Brown Corpus – the first computer corpus of the English language – at UCL, in Quirk's room, and I thought it would be good to create a British 'clone' of the American corpus. But the computing facilities were primitive, and we had no expertise in using them. Also, we needed funding to support the input of data (using punch cards in those days) and other tasks needed to compile and use a corpus. It was eight taxing years before it was finished, and if I had known the time and effort it would take, I probably would never have started!

One of the difficulties of developing a corpus was that the idea of studying the real data of language use was totally out of fashion at that time. It was difficult to explain the value of the corpus, and the uses to which it would be put. But we soon started using the data of the LOB Corpus, as it was called, in our teaching, and in no time postgraduates were using the data to study areas of English grammar for their dissertations and theses. It was like a window into the English language that hadn't existed before.

**BELLS: Please tell us about Lancaster University language research as it was shaped in the eighties and as it is at this moment.**

GEOFFREY LEECH: That's a long story. But I'll try not to bore you with too much past history.

Let's begin where I left off in answering the last question. In my own research, the drudgery of corpus work began to pay off in 1977, when we (1) set up an international organization called ICAME (originally called the International Computer Archive of Modern English). (2) In the 1980s, with the help of computer scientist Roger Garside, we became seriously computational, developing automatic part-of-speech tagging programs and corpus parsing. The aim was to produce not just text corpora, but corpora annotated with various kinds of linguistic information which could be a springboard for more advanced research. In 1991-5 we joined with a consortium led by Oxford University Press to compile the British National Corpus – a collection of over 4000 written texts and spoken transcriptions – a hundred times larger than the LOB Corpus. Many different kinds of corpus-based project followed. This was the time when it was not so difficult to find funding from the government or from industry to develop corpus-based research, and we wanted to take advantage of it – that's opportunism again!

A great thing about corpus research, in my experience, is that it cannot be done individually, and so corpus linguists tend to work in teams, which is very stimulating and productive of new ideas. Five corpus linguists now teaching at Lancaster (Tony McEnery, Paul Rayson, Andrew Wilson, Paul Baker and Andrew Hardie) began their careers as researchers working on funded projects and have since branched out into new fields. Since we began that very unpromising trail of corpus compilation in the 1970s, a 'corpus revolution' has taken place not only in the UK, but in many other countries as well.

That's enough on the computational theme. Our department has managed to grow on a number of different fronts since the four or five of us met around that table in 1969. This is the story not only for Lancaster, but for other 'new universities', as they were called, founded in the 1960s – York, Surrey, Stirling, Essex, Sussex, Kent – they read like the names of lords in a Shakespearean history play.

In 1974, after some strife in the Department of English, our small section became a separate department, the Department of Linguistics and Modern English Language. (3) One of the unfortunate results of this was that English literature and language split up and have been taught in different departments, which have relatively little contact. The literary department, now called the Department of English and Creative Writing, has flourished as ours has, but along an entirely different track.

In those days our fledgling linguistics department had nine members of teaching staff. Now there are more than thirty. This expansion has come about as a result of a continuing increase in student numbers, but also because of increasing research activities. In the eighties, three centres of linguistic research were established, and remain with us today: UCREL (Unit for Computer Research on the English Language) (4) for corpus work; CLSL (Centre for Language and Social Life) and CRILE (Centre for Research in Language Education). These represent three broad areas of research which reached a 'critical weight' early on, but of course they are far from the only areas in which research is done. After some of us had set up UCREL, CLSL was set up mainly by the initiative of Chris Candlin, before he left for Macquarie University in Australia in the mid-1980s. His dynamic role in shaping the department in its earlier days should not be forgotten. At present CRILE counts such names as Martin Bygate, Charles Alderson and Keith Johnson among its luminaries. In CLSL at present among the well-known names are Ruth Wodak and Paul Chilton (who are carrying on the research tradition in CDA – critical discourse analysis – pioneered by Norman Fairclough), Paul Kerswill and Mark Sebba (sociolinguistics) and Greg Myers (language in the media). More recently, a newer centre, led by David Barton, the LLRC (Lancaster Literacy Research Centre) has branched off from CLSL and become a national hub for literacy research.

It should be emphasized that these centres are not hermetically sealed compartments – far from it. We encourage multiple-membership of the centres – they're like loose confederations of people with related research interests. The department also has many smaller research groups which have meetings every week or fortnight or so during the term, often with invited speakers. An example is the PASTY (Pragmatics and Stylistics) research group, led by Mick Short and Elena Semino, reinforcing our connection with literature which has existed ever since the early days. Another research group is RITL (Research in Theoretical Linguistics) – Anna Siewierska is its best-known member – and its the focus is mainly on typology, cognitive linguistics and related theoretical approaches. Again, people can attend any group meetings that interest them. The three groups that I sometimes attend are CRG (the Corpus Research Group), PASTY and RITL.

It may be symptomatic of the development of linguistics at Lancaster that I have mentioned theory last of all. There is little doubt that compared with other linguistics departments in the UK, Lancaster is very much an

*applied* linguistics department – using *applied* in the broadest sense – and the teaching of syntax, phonetics, phonology, and semantics tends to be on the sidelines of the department, rather than in its centre. I have always regretted this limited presence of ‘core linguistics’, and it’s interesting to consider how this applied tendency has grown up. I attribute it to the fact that our department was never established by a decision of the ‘University Authorities’ – as happened with important ‘prestige’ departments like Psychology and Law. These grew up while Lancaster was gradually expanding from small beginnings, and was gradually making itself into a ‘proper university’. Instead, like Topsy in *Uncle Tom’s Cabin*, linguistics ‘just grow’d’. We’ve been allowed to grow above all through increasing student numbers, and students have a tendency to choose courses with a strong human interest, practical applications, and (dare I say it?) no particularly strenuous involvement with maths, logic and theory. The areas popular with students are of course the areas in which new staff – to cope with student demand – have tended to be appointed. So the staff’s research interests have a tendency to reinforce the students’ preferences and vice versa.

But there are also some strengths in this applied tendency. In recent years, research funding has tended to go more towards practical outcomes than towards purely theoretical investigations. So that has helped us. Also, applied linguistics lends itself to another trend favoured by funding bodies – what people are starting to call ‘interdisciplinarity’. The kind of linguistics our department specializes in naturally leads to collaboration across disciplinary boundaries – with psychologists, with sociologists, with computer scientists, with media specialists, and so on. I would argue, in fact, that part of the appeal and influence of linguistics as a discipline is that it interacts with so many other disciplines. All disciplines use language, after all.

**BELLS: What was your academic connection with Birmingham University, ELR, and particularly with professor John Sinclair who also showed strong interest in corpus linguistics and was a key figure in that field?**

I got on well with John Sinclair on a personal level, but on the level of ideas and methods, we used to clash. We were both pioneers in corpus linguistics, but the ways we got into it were different. John’s initial interest in corpora was research-oriented, while mine was resource-oriented.



Let me explain. John, when working with Halliday in the 1960s, saw the need to develop a theory of lexis – with word collocations and phraseology at its foundation. For this, he needed a large amount of text that could be processed statistically by computer. So for him a computer corpus was a necessary tool for a ground-breaking research programme, which he developed in the succeeding forty years with his colleagues and students. In the 1970s and 1980s he built up the Birmingham Collection of English Texts – a much bigger corpus than LOB – and the Cobuild Dictionary. Later the enormous Bank of English came along, and it enabled his theory of lexis to be tested and developed far beyond what could have been imagined in the 1960s. More importantly, the theoretical outcome was a wide acceptance among corpus linguists – and some others, such as construction grammarians – that grammar and lexis are not separate divisions of language (as the grammar and the dictionary traditionally make them), but that the co-occurrence of words at the lexical level is fundamental, and grammar is some kind of superstructure built ultimately out of word-cooccurrence. His watchword was ‘Trust the corpus’ – that’s to say, if your corpus tells you something, you believe it, but if it doesn’t, you don’t.

To explain my response to this, I will first suggest that in many aspects of my thinking about language, I am a middle-of-the-road person. My position is middle-of-the-road, for instance, with respect to theory and data: I think that abstract theory and observable data are both important, and the greatest challenge is to connect the two. I am against what I consider to be the extreme position of Chomsky – that theory is supremely important, and that the empirical evidence of language use is of little or no value. Now, I considered John Sinclair’s viewpoint to be extreme in the opposite direction. The data-driven approach that he advocated opposed any application of pre-existing theoretical concepts to corpus data, the argument being that no such concepts could be validated unless they emerged from corpus evidence. For him, the corpus data was all-important, whereas for Chomsky it is useless. My position (somewhere between what I would call extreme empiricism and extreme rationalism) is that we need both theory and the data of real language in use.

So my priority, following in the footsteps of Quirk and Francis, was first to create a body of data – a corpus – and then make it available for researchers to use for whatever purpose they want. In practice corpora like Brown and LOB have been used for amazingly varied investigations

– research into lexis, grammar, semantics, psycholinguistics, stylistics, and so forth. This is what I meant by a ‘resource orientation’. (It’s a truism that a corpus is likely to prove useful for many many purposes undreamed of by its compilers.) To make the corpus more useful, both for ourselves and other users, we were keen to annotate it with grammatical information such as part-of-speech tagging, so that syntactic information (for instance) could be extracted from the corpus. Our position was that unless you build certain kinds of linguistic information into the corpus, your ability to use the corpus for abstract investigations (for example, of syntax) was severely limited. So annotated is a kind of value added. John disagreed with this, as for him even categories like ‘noun’ and ‘adjective’ could be seen as impositions on the corpus of theoretical constructs. For him, the corpus in its pure form was the only thing that could be trusted.

There were other differences between us too, but that is enough to give you the flavour of the debate. There is no doubt that John’s contribution was enormous and original in establishing the importance of lexis and phraseological structure. But I formed the impression that he saw little value in my own approach.

**BELLS: Your academic connection with the Scandinavians seems to be as strong as ever. Please tell us something about that.**

GEOFFREY LEECH: Yes, I owe a lot to Scandinavians. The key names that provide the connection are Quirk, Svartvik and Johansson. I have already mentioned my friendship with Jan Svartvik when he was RQ’s senior researcher at the Survey of English Usage. Later we collaborated on the ‘Quirk grammars’, as well as a grammar of our own called *A Communicative Grammar of English*. When Jan returned to Sweden on completing his doctorate, he was soon appointed to the chair of English language at Lund University. In the mid-1970s, his brightest student, Stig Johansson, was awarded a visiting fellowship which, at Jan’s instigation, he spent with us at Lancaster. That was the time when I was in the depths of despondency about the completion of the Lancaster corpus. At the end of his stay in Lancaster, Stig (to my astonishment and delight) offered to take over the project. The main logjam we were suffering from at that time was the problem of copyright. British publishers were reluctant to give *free* permission to an obscure provincial university wanting to computerize and distribute their texts, and we could not afford to pay their permission fees.

But when Stig returned to Scandinavia (where he had secured a post at Oslo) he was able to write to the London publishers from a foreign vantage point, and to speak on behalf of an international organization. It seemed as if they were being invited to belong to a prestigious-sounding collection of British English text. His entreaties succeeded where mine had failed – so the corpus was finished, and still bears the name of the three cities where it was created: LOB stands for ‘Lancaster-Oslo-Bergen’.

Did I say an ‘international organization’? Yes, to get permission from the publishers, we (Jan, Stig, Nelson Francis and I) adopted the strategy of setting up a rather grand-sounding ‘International Computer Archive of Modern English’, which we thought was a very good idea anyway – I mentioned it earlier, and can now say that ICAME has kept going for over 30 years, and is the oldest association for corpus linguistics in the world. Stig stayed on for many years as the coordinating secretary of ICAME – he effectively ran the whole organization from Oslo –, and the *ICAME Journal* and the ICAME website have been run (by Knut Hofland) from Bergen. So that Scandinavian connection has continued without a break.

Meanwhile in the later 1970s Jan Svartvik masterminded an important corpus project of his own: he computerized most of the spoken data of the Survey of English Usage corpus, with its complicated prosodic transcription, and so created the London-Lund Corpus. Brown and LOB had contained only written texts, so this was another important milestone – the first computerized spoken corpus of modern English.

I just have to mention another key Scandinavian figure in ICAME – Matti Rissanen, who with his colleagues has made Helsinki the top centre historical English corpus research. As chair of the ICAME Board, Matti took over Stig’s coordinating role in the 1990s, and inserted the word ‘Medieval’ into the name of ICAME.

**BELLS: What is a current linguistic scene in Great Britain and what do you anticipate as future orientation?**

GEOFFREY LEECH: Well, I don’t feel well qualified to answer this, as I no longer see myself in the vanguard – if I ever did – and the academic linguistics community has become fragmented. I remember the founding of the LAGB (Linguistic Association of Great Britain) in the early 1960s, and my feeling then was that it represented all the linguistics that was really worth following in the country. Now the LAGB still exists, together

with a much older association – the Philological Society. They still publish their journals and have meetings. The LAGB represents core linguistics of a fairly theoretical orientation, but now there are many other organizations representing sub-fields of linguistics – including ICAME, BAAL (British Association of Applied Linguistics), the Sociolinguistics Symposium, the Association of Computational Linguistics, PALA (Poetics and Linguistics Association) and so on – and many conferences spring up via the internet without any supporting associations. So it is difficult to keep track of what is going on. Also all these conferences, although some of them were founded or organized in the UK, tend to have an international membership. One important international foundation of this kind is ISLE (International Society for the Linguistics of English), which was founded in Manchester a couple of years ago, and had its inaugural conference in Freiburg in Germany.

For what it's worth, I will hazard some impressions and speculations about what is happening in the UK, which is probably not too different from what is happening in other countries. Corpora are becoming mainstream and their use is no longer restricted to people who regard themselves as 'corpus linguists' – even theoreticians are consulting corpora from time to time. The Chomskyan paradigm, on the other hand, is not growing, and is probably in gradual decline. An increasingly important theoretical focus is cognitive linguistics and its associated models such as cognitive grammar, construction grammar, and usage-based linguistics. 'Hyphenated linguistics' is flourishing on many fronts (socio-, psycho-, clinical, forensic, historical...), so the interdisciplinary reach of linguistics is continuing. What all this amounts to I'm not sure, but perhaps there's a move towards a greater interest in observational methods and the real use of language in all its varied settings.

**BELLS: You have recently retired. Mission accomplished?**

GEOFFREY LEECH: No! I still have a little office in the Department and go there about twice a week. I supervise a couple of PhD students and do the occasional lecture. I do a number of academic visits to other countries – for example, last year I lectured in Thailand, Turkey, Japan and Montenegro. I keep busy writing articles and papers for various publications, and have four book projects in the pipeline. At present I am trying to write a book on linguistic politeness – a topic which engaged my attention quite a lot

in the 1970s and early '80s, and which I have been recently revisiting. The other three book projects are all new editions or anthologies of existing publications – perhaps it's a sign of old age that I am now spending more time looking back to what I have written earlier, trying to improve and update it, rather than looking forward to completely new ventures. There's an ancient refrain 'Old soldiers never die, they only fade away', and for me, the word 'soldiers' can be happily replaced by 'professors'.

**BELLS:** I am happy to see that there is a Scholarship Trust honouring your academic achievement. Not only have you been an academic but also a professor much loved and admired by your students and fellow researchers. I myself was a lucky beneficiary of your expertise and generosity in giving your time and patience when I came to see you in Lancaster. I was impressed by the fact that it took only three weeks to organize the meeting – you were already a celebrity and I was only a doctoral student. Apart from the message that you send by being what you were and what you are, is there anything that you would feel like saying when addressing the young academic audience world-wide and in Belgrade in particular?

GEOFFREY LEECH: Once again, you are flattering me! I can only reiterate the message, implied in what I have already said: 'Take every opportunity!' 'Leave no stone unturned in the search for linguistic reality'.

- 
- (1) We = Nelson Frances, Stig Johansson, Jan Svartvik, Arthur Sandved and myself.
  - (2) Later the title was expanded to: International Computer Archive of Modern and Medieval English.
  - (3) It is now called the Department of Linguistics and English Language, as since the 1990s we have taken on some historical language teaching and research.
  - (4) The title has changed (mainly because research has extended to many other languages) to: University Centre for Computer Research on Language. But we have kept the original acronym UCREL.

