Geoffrey Neil Leech – An Academic Autobiography   12 June 2009

1. EARLY YEARS: 1936-1960

I was born in Gloucester, in the west of England, on 16 January, 1936. My father was a bank clerk, the son of a dispensing chemist (or pharmacist). My mother was the daughter of a butler. My parents had two children: an older brother Martin, and myself.

My father became manager of a bank in the small country town of Tewkesbury, near Gloucester. There I began my secondary education, at Tewkesbury Grammar School, a very small school of some 120 pupils, which was however over 400 years old.

After two years of National Service (in the R.A.F., where I reached the rank of senior aircraftman, and spent most of my time shorthand-typing in West Germany) I began my undergraduate studies at University College London in 1956. It was by accident that I went there to study English. Being interested in languages, I really wanted to study French at the university. My father happened to drink in the same pub as Professor A.H. Smith, who was Quain Professor of English at University College London (UCL), and who happened to own a weekend cottage at a village near Tewkesbury. As a favour to my father, Professor Smith gave me an interview at his country cottage, but I must have offended him when I said I really wanted to study French! However, he offered me a place in his department, and I duly began my undergraduate career at UCL.

My undergraduate career was undistinguished, and I graduated with an Upper Second Class B.A. honours degree in English Language and Literature in 1959. During my undergraduate years, I had become particularly interested in the linguistic part of the syllabus, and had opted for what was then called 'Syllabus B' – a set of courses which contained a large component of language work, more historical than contemporary. For example, in Syllabus B, we had to study the *whole* of *Beowulf* in the original, not just a part of it. Among the courses I took were Old English, Middle English, Old Norse, English Philology and Phonetics. This last course was taught by A.C. Gimson and J. D. O'Connor, distinguished phoneticians who were among the senior teachers at UCL at that time.

While thinking of famous teachers, I should mention that as an undergraduate I was fortunate enough to attend a series of lectures by J. R. Firth, the first British professor of linguistics, and in many ways the father of linguistics in the UK. He gave a series of lectures at the University of London during my first year, and made an indelible impression on me as a personality. At that time, I could scarcely understand his message, although I remember that the term 'context of situation' figured prominently in it. Another great man whose lecture

I was privileged to attend was Daniel Jones, the first professor of phonetics in the UK, and the father of the British school of phonetics. He was about 80 when I attended a lecture of his on – predictably enough – 'The Phoneme'.

I regard it as a very happy accident that I went to UCL to study English, not knowing at that time that this was a college well-known for English language studies, which was to provide the entrée to a circle of distinguished language scholars.

One of my contemporaries in the English Department at UCL was Roger Fowler, later Professor of Linguistics in the School of English and American Studies at the University of East Anglia. His career and mine followed similar courses: having known one another at Tewkesbury Grammar School, he and I both followed the 'Syllabus B', which gave us roughly equal doses of language and of literature. Probably this is why we both ended up taking a deep interest in the relation between linguistic and literary studies, and in the interdisciplinary field of stylistics. (Later in our careers, our paths diverged – Roger's moving into critical discourse analysis, and mine into computers and corpus linguistics. Regrettably, Roger died soon after his sixtieth birthday, in 1999.)

My interest in scholarship had grown in my third and last year as an undergraduate. Roger Fowler and I competed for the Quain Essay prize – for the then magnificent sum of £50 – and both wrote at length on the set topic of 'The persistence of the medieval conception of tragedy in post-medieval literature'. Roger won the prize, and I had to be content to receive (as a consolation) a lesser prize, which entitled me to £25 of books.

2. AN M.A. STUDENT: 1959-62

After graduation, I wanted to continue my studies as a research student at UCL. By this time I was becoming interested in modern linguistic research, but knew very little about it. Linguistics had so far made little impact in the UK, and there was no teacher in our department who could adequately supervise me in that area. However, at that time (1959) there was an initiative at UCL to promote the study of communication. An interdisciplinary conference on communication was held, and a new Communication Research Centre (CRC) was inaugurated. But there were two severe handicaps in the work of this Centre: first, the Centre had no funds or research staff; and second, scholars could not agree on what 'communication' was, and how it might be studied. Everyone generally agreed that 'communication' was important – but different disciplines had differing approaches to it.

As a modest beginning to the work of the CRC, two or three postgraduates in the English Department at UCL began to study the use of language in public communication.

One student took as his province the study of public information documents, another – Eugene Winter, well known later for his work on textual structure -- began to study the language of press advertisements, and a third – myself – began to study the language of television commercials, then a relatively new medium of advertising in the UK.

I had been granted a State Studentship enabling me to study for an M.A. (then a research degree at London University). However, we three students made little progress, since none of us knew what techniques would be appropriate. Little supervision was offered: we were left to find our own way. At this time, I grew disheartened with the work, left the university, and began teaching at a secondary school. I continued school teaching for about 18 months, making a very indifferent shot at being an English teacher in an overspill estate near London, and keeping up my M.A. studies as well as I could in my spare time.

On 29th July, 1961, while teaching at that secondary modern school near London, I got married to Frances Anne Berman, a psychology graduate I met at about the time of my own graduation at UCL. Soon after that, on 1st Janaury 1962, I was fortunate enough to be granted a research studentship in my UCL department: this was a meagre sum of £750 per year (slightly less even than I had been earning as a teacher), but I was overjoyed to have the opportunity to abandon school teaching and take up full-time research. I owed this 'break' to a commercial television company, ATV. How fortunate I was that some television magnate happened to donate to UCL a moderate sum for research into the language of advertising, at the instigation of Professor Smith, at that time!

But we still had the problem of a lack of research tools. At that time, Randolph Quirk, who had been a student and teacher in the English department at UCL, had accepted a chair there. He was about to return to his old department once again, after spending a number of years at the University of Durham. He suggested to our supervisor that we should read the new linguistics at that time coming out of the USA, in order to arrive at the best analytic categories for describing the language of television. 'New linguistics', for us, included books now largely forgotten: books on English syntax by Paul Roberts, W. Nelson Francis, A.A. Hill and James Sledd. These works showed the influence of American structuralism: we had yet to catch up with the new generative grammar associated with Noam Chomsky.

3. TEACHING AT UNIVERSITY COLLEGE LONDON: 1962-1969

In the summer of 1962, I had another piece of immense good fortune, when a temporary assistant lecturer's post became available in the English Department at UCL. My head of department, Professor Smith, was apparently ready to appoint me. (In those days, the say-so of a head of department was enough to grant someone – or to lose someone – a job.) But

before the decision was made, he offered his new professorial colleague, Randolph Quirk, the opportunity to vet me, and decide my fate. This interview was my first meeting with Professor (now Lord) Quirk, who was later to play a most important role in my developing career. At the interview, I was ready to be overawed, but his manner was so cordial that he soon put me at my ease. It seems that he was satisfied with my performance as an interviewee, for I was offered the post, much to my surprise and delight.

My most important task as a novice lecturer was to plan and deliver a series of lectures on 'Rhetoric' for first-year students. Previously, this lecture series had been on the history of rhetoric from classical times, and had been reputedly the dullest course offered by the Department. I was given *carte blanche* to teach the course as I wished, and chose to treat literary language (especially the language of poetry) from the modern linguistic point of view, rather than from that of the rhetorical tradition.

In 1963 I finished my MA thesis on *The Language of Commercial Television Advertising*. Having listened to so many commercials, and studied them *ad nauseam*, I was tired of the whole subject. I should have been more grateful to my ATV sponsors, without whom I could scarcely have put a foot on the academic ladder. At least I was grateful enough to send them a copy of my thesis, but there was no evidence that they read it or found it useful.

At this time I found Professor Quirk extremely helpful and encouraging. He invited me to embark on a book, intended for publication in a new series he was editing: the Longmans English Language Series (ELS). The book was to be based in part on my M.A. thesis, but was to be extended to a more general treatment of the language of advertising. It was eventually published in 1966 under the title of *English in Advertising: A Linguistic Study of Advertising in Great Britain*. Three years before that, in 1963, I had published my first article – also on an aspect of the language of advertising – in an obscure though reputable continental journal.

After working on my new 'Rhetoric' course, however, my favourite subject then was the language of literature, and this led to the publication of two papers in 1965 and 1966. This was a period when, for the first time, modern linguistics was being applied to the study of literary language in the UK. Often, I felt, this approach led to misunderstanding and even animosity between literary and linguistic scholars. However, I had been much influenced, as an undergraduate, by the lectures of the textually-oriented literary critic Winifred Nowottny (author of *The Language Poets Use*), who was now a senior colleague in my own department. I felt a rapprochement could be achieved between these two approaches – the linguistic and the literary. This thinking eventually became the leitmotiv of my book *A Linguistic Guide to*

English Poetry (1969), also written with the editorial encouragement of Randolph Quirk, for the Longman ELS series.

My period as an assistant lecturer – then lecturer – in the Department of English Language and Literature at University College London lasted from 1962 to 1969. I have mentioned two strands of my academic development in that period – the study of register (particularly advertising) and the study of literary style. I will now backtrack a little to introduce a third strand – semantics.

In 1963 (I believe) Dr M.A.K. Halliday was appointed the first full-time Director of the Communication Research Centre at UCL, and under his influence the whole direction and thrust of the CRC underwent a transformation. Soon after, indeed, Michael Halliday (as I soon came to know him) became the first Professor of Linguistics at UCL. As he was a charismatic teacher and delightful, approachable person, I benefited greatly from close contact with him in 1963-4, when he was Director of the CRC, and I was Assistant Director. I should explain that UCL was reluctant to establish a Department of Linguistics, although linguistics was then becoming a popular and 'fashionable' new subject in the UK. Hence the CRC, of which I was a sort of caretaker at that time, was conveniently regarded as a stalking horse, an incipient Linguistics Department which could safely be launched after Michael Halliday was installed. Halliday had made his reputation in Edinburgh, and it was considered a great coup that UCL had managed to entice him down to London. After he had been at UCL for a few months, the CRC faded into the background, and the Linguistics Department came into its own. At that time I was greatly influenced, as were many in the country, by Halliday's linguistic theory, then called 'Scale and Category Grammar', later renamed 'Systemic Linguistics' or 'Systemic Functional Grammar'. I was interested in exploring Halliday's concepts of system and structure in new directions, and asked his advice about which branch of linguistics I should tackle – morphology or semantics – as neither of these had so far been sufficiently investigated. He advised me to take up semantics, and indeed I did, soon finding the opportunity to teach a new course in the subject to postgraduate students. (How easy it was to launch a new course in those days!) However, my ideas on semantics, which veered towards the integration of componential analysis and logical semantics, were rather different from those of Halliday, for whom the notions of 'context' and 'situation' (related to his teacher J.R. Firth's concept of 'context *of* situation') were the basis for the study of meaning.

While I was trying to develop my Hallidayan approach to semantics, I was given the opportunity to spend a year in the USA as a Harkness Fellow (1964-5). At the interview for this fellowship, I was confronted by a 10-man panel of 'the great and the good' of the academic world, of which one, Sir Isaiah Berlin, had assumed the task of interrogating me about my research programme, in his well-known gruff-barking manner of delivery. I can

only assume that his bark was worse than his bite, as I was granted the fellowship, which by my standards was amazingly generous. My wife Fanny and I travelled to the USA (with our baby son) on the liner *Queen Elizabeth*, and we travelled home a year later on the *Queen Mary*, by which time our Tom had become an obstreperous toddler of 18 months. We toured the USA for three months as tent campers. More importantly for my career, the Fellowship gave me the opportunity to study a subject of my choice at the American university of my choice: so (who would not at that time?) I decided to study linguistics at the Massachusetts Institute of Technology (MIT). Ironically, Chomsky was not there at the time, and I found to my chagrin that he was on leave in London! However, he returned to the USA later during my stay, and I had the opportunity to meet him and attend one or two of his lectures. I was struck by the contrast between Chomsky's public persona and his private personality. In lectures, he was as sharp and uncompromising in defending his own ideas and dismissing those of others as in his writings. As a private man, he was mild, diffident, and easy to talk to. Having lunch with him in a diner near Harvard University with Jerrold Katz, I was nonplussed yet fascinated to find all the talk to be of politics and how to keep Senator Barry Goldwater out of office, and not about linguistics and the latest models of transformational grammar.

Although I learned a great deal at MIT, particularly about the habit of rigorous thought and application of theory, I must confess that the MIT approach to linguistics was too constraining for my taste, and that I found the intensely intellectual atmosphere there somewhat uncomfortable. Perhaps I was not young enough to imbibe the powerful drink of transformational grammar uncritically. Or perhaps it was that as a visiting student (rather than a regular member of the PhD program) I was necessarily on the fringes of the world's powerhouse of linguistics. Nevertheless I got to know many of the leading linguists of what was then the young generation: J.R. (Haj) Ross, Perlmutter, McCawley, George Lakoff and other notables all passed through the MIT graduate school at that time. Barbara Hall (later Barbara Hall Partee) taught an excellent course on the mathematical and logical basis of linguistics – and I found that the most valuable of my MIT courses. I also attended courses by world-renowned figures in linguistics at that time: Morris Halle (phonology) and Paul Postal (syntactic theory).

Going to MIT taught me a great deal about many areas of linguistics. But one of the things one easily learned at MIT at that time was a sense of conviction – the assumption that MIT led the world (as indeed it unquestionably did), and that others' heterodox opinions need not be treated too seriously. Although as a visiting student I remained something of an outsider at MIT, I suppose I must have carried something of that arrogance back to the UK with me. However, any sense of superiority was soon punctured when I gave a paper on

semantics at the Philological Society, where my nascent semantic theory met with some scepticism and hostility. (By this time Halliday had left UCL.) Later, I sent my work to Professor John Lyons at Cambridge, to see if he would publish it in the newly-founded *Journal of Linguistics*, but he was similarly unimpressed. It was evident that my ideas on semantics needed more careful exploration, so I developed my theory further into a monograph. Even then, I could not get it published. At that point, it was accepted that I could use this material as a basis for my PhD thesis. During the period1965-1969, after my return to UCL from MIT, I rethought and developed my work on semantics until it became a PhD thesis with the title: *An Approach to the Semantics of Place, Time and Modality in Modern English*. Finally, I revised the material yet again, and it was published in 1969 in the Longman Linguistics Library as a book entitled *Towards a Semantic Description of English*. This book was out of print in a very few years, and is hardly read today. But its publication did more to win me a reputation in linguistics than any other volume. McCawley recommended it for publication in the USA, by Indiana University Press.

This was a period when work in semantics as a sub-field of linguistics was developing extremely quickly. In the USA, it was the era of the dispute between generative semantics (Lakoff, McCawley, Ross, etc) and interpretive semantics (Chomsky, Jackendoff, etc). My thinking was in some ways parallel to the generative semantics school, but at the same time I developed my own theory, based on autonomous semantic and syntactic representations, linked by mapping rules.

Another important benefit to my academic career in 1962-9 was more empirical and practical, and ultimately more lasting, than the influence of MIT. I was fortunate to have a close association with the Survey of English Usage: a research centre founded by Randolph Quirk in 1959, and attached to our department at UCL. The most important part of the Survey's work at this stage was the compilation and analysis of a large corpus of modern English texts, both spoken and written. Three of Professor Quirk's leading researchers in the early years of the Survey were David Crystal, Jan Svartvik and Sidney Greenbaum – all subsequently well-known in English linguistics. Svartvik and Greenbaum later collaborated with Quirk and me as co-authors of *A Grammar of Contemporary English* (*GCE*), a detailed descriptive grammar published by Longman in 1972.

The writing of *GCE* was a large enterprise, drawing on the work of the Survey of English Usage. We felt at that time that there was a large gap between the type of academic and theory-driven grammar that was studied in linguistics departments, and the type of grammar which was needed for the English language classroom. There was a consequent need for a reconciliation between theory and practical pedagogy in the study of English grammar. It was this reconciliation that *GCE* tried to achieve. Largely because of Quirk's

leadership, and in spite of countless arguments between members of the team, the collaboration was more successful than we had dared to suppose. The book, in spite of its weaknesses, became well-known throughout the world as a source of descriptive information on English grammar.

4. ESTABLISHING A NEW DEPARTMENT AT LANCASTER UNIVERSITY: 1969-1974

In 1969, I applied for the post of Senior Lecturer at the new University of Lancaster, in the north of England (about 240 miles north west of London). I was appointed to this post, and on my arrival at Lancaster was further promoted to Reader in English Language (equivalent to Associate Professor). After this move, Lancaster became my permanent academic home. Even today (2009), forty years later, I remain on the books of Lancaster University as Emeritus Professor of Linguistics and English Language.

During my early years at Lancaster, much of my research time was spent in the collaboration with Quirk, Greenbaum and Svartvik on *GCE*. However, I also found time to continue my work on semantics, with the publication of *Meaning and the English Verb* (Longman 1971) and *Semantics* (Penguin 1974) – both books that have subsequently been published in revised editions. When working on *GCE* I remember being frantically busy, not only because of the big task of writing the book, but because of a heavy teaching load and, worst of all, a battle that was taking place in the English Department, between two factions. The 'left' was supported by the students, and the 'right' supported the Head of Department. This disrupted the whole campus and became a *cause célèbre* in the national press for a time. The cause of this strife, which need not be elaborated here, seems in retrospect rather trivial, but that was a time when student unrest took hold on university campuses throughout the world. At one point I just had to plead with my three co-authors to give me more time to complete my chapter drafts.

After *GCE* was published in 1972, the four authors decided, with the agreement of the publishers, to write two advanced students' grammars based on the approach of the larger book. One of them, written by Quirk and Greenbaum, was in effect a shorter version of *GCE*, entitled *A University Grammar of English* and published in 1973. The other book, written by Svartvik and me, was *A Communicative Grammar of English*, published in 1975. In this book we tried to develop a somewhat fresh approach to English grammar, based on the idea that grammar, to be useful to the learner, should be 'communicative' in the sense of relating the forms and structures of the language to their meaning and use. In terms of number of copies sold (over 250,000), this has been my most successful book, and, like the books mentioned in the previous paragraph, it has been published in a revised editions (1994 and 2002).

Lancaster University had been founded only five years previously in 1964, and during my early years there the university was growing at a rapid rate. New disciplines were being established, new departments founded, new buildings erected. As student numbers rose, we could appoint more staff, and this gave us the opportunity to develop the study of linguistics almost 'by stealth' within the English Department. I was head of the 'Linguistics Section' of the Department, and in 1974 (in the aftermath of the 'Craig affair' which had caused political strife in the Department and the University in 1971-1972) this became the Department of Linguistics and Modern English Language, one of the three constituent departments of a new entity called the School of English. When the new Department was established, I was promoted to the post of Professor of Linguistics and Modern English Language. At about that time, also, we began to offer Linguistics as a major undergraduate subject: Lancaster was among the first universities to offer a B.A. in Linguistics.

5. THE BEGINNINGS OF COMPUTER CORPUS LINGUISTICS: 1970-1978

As early as 1970, before linguistics became a separate department, the small group of young linguists ('the Linguistics Section') at Lancaster got together round a table, and considered how Lancaster could make its mark in the world as a new centre for research. At my suggestion, we decided to develop a computer corpus of British English, one which would match in every possible respect the Brown University Corpus of American English which had recently been completed and distributed, and which was the first computer corpus of modern English. Like the Brown Corpus, the 'Lancaster Corpus' would consist of more than a million words of various registers of written (printed) English. As its director, I found this project onerous and time-consuming. Our computing facilities were primitive. We received some financial help from the publishers Longman and later from the British Academy, but the money was soon used up. There were also great and apparently insurmountable problems concerning copyright. In 1976 I was about to abandon the whole project, but a former student of Jan Svartvik, a Swedish scholar named Stig Johansson, offered to take the project to Norway (where he had recently obtained a post), to complete the corpus there. He managed, writing as the secretary of an important-sounding international organization (see next paragraph) to obtain permissions from UK copyright holders who had withheld such permissions when requested by an inmate (myself) of a provincial northern university of no great repute. At last, in 1978, the corpus of written British English was completed, through the combined efforts of three universities: Lancaster, Oslo and Bergen. It was therefore called the Lancaster-Oslo/Bergen Corpus (or the LOB Corpus), and has since been widely used throughout the world, along with the Brown Corpus.

In the previous year, 1977, a group of English language specialists including Randolph Quirk, Jan Svartvik, W. Nelson Francis, Stig Johansson and myself had met in Oslo and founded the International Computer Archive of Modern English (ICAME), an organization to develop and promote the use of computer corpora in English language research. Oddly enough, the original impetus for this initiative was the need to achieve copyright clearance for the LOB Corpus. In 1976 the London publishers and literary agents appeared to be 'ganging up' against the LOB Corpus, consistently refusing to grant permission we needed without a fee for each 2000-word sample (which we could not afford). If one wanted to persuade the London publishers and other copyright holders to grant free copyright, it helped to have an address in Norway rather than Lancaster, and it helped to write as 'secretary of the International Computer Archive of English Texts'. This stratagem succeeded, and more importantly, ICAME (later re-entitled the International Computer Archive of Modern *and Medieval* English) has continued to flourish, as an organ for the promotion of corpus-based research. It has an annual conference, a journal (*ICAME Journal*) and a distribution and information service, based in Bergen, supplying computer corpora and related documentation and software. During its lifetime, the use of computer corpora, from being a fringe maverick activity, dismissed as valueless by MIT-inspired linguistics throughout the world, has become a mainstream methodology, both in computational linguistics and in English language research. Only in some varieties of theoretical linguistics has the use of a computer corpus remained suspect.

6. OTHER ACTIVITIES AND PUBLICATIONS 1976-87

In 1977 I finished my three-year term as the first head of the Department of Linguistics and Modern English Language (a post I have more or less managed to avoid since that time). I then made an arrangement with the University of Lancaster whereby I worked half-time at the University, and worked half-time on my own research and publications, as a 'free-lance'. This arrangement continued for eight years, and enabled me not only to work on several books, but also to continue my computer research. If I had continued full-time at the University, I would not have been able to bring all these research interests to fruition.

One new research interest had already been developing in the mid-70s from my earlier work in semantics. This was the study of pragmatics, a fast-growing and popular new area of research in linguistics, arising from the controversies about meaning which had dominated linguistics around 1970, as well as from the work of philosophers such as Grice and Austin.

In 1976 and 1977 I published a number of papers on the border between pragmatics and semantics – indeed, a recurrent theme of these papers was that *was* such a borderline.

Pragmatics could not be absorbed into the all-embracing field of semantico-syntax, nor could it just be ignored as a waste-bin for things linguistics could not handle. These papers I eventually revised to make a book *Explorations in Semantics and Pragmatics*, published by John Benjamins in 1980. The book, dealing with themes such as metalanguage, performatives and politeness, was inconclusive, showing a train of thought beginning in semantics and ending in pragmatics. After this, I found the issues of the scope and methodology of pragmatics a large scale problem which could only be tackled by a separate book-length study. This was finally published in 1983 as *Principles of Pragmatics,* published in the Longman Linguistics Library. Perhaps the best known part of this book is where I develop a theory of politeness, a theme which began much earlier in a short monograph 'Language and Tact' (1977). This was also the time at which Penny Brown and Stephen Levinson were developing, in a more thorough-going way, their theory of politeness which has since been the most influential model in the field (published in 1978 and revised for re-publication in 1987). I have recently begun to resume my work in this field, which has developed hugely since my early contributions to it in the 1980s.

Alongside work in pragmatics, I worked with a Lancaster colleague, Mick Short, on a book dealing with literary stylistics: *Style in Fiction: A Linguistic Introduction to English Fictional Prose* (Longman English Language Series, 1981). This was a return to an area related to that of my earlier book *A Linguistic Guide to English Poetry*, and in fact we began by thinking of the two books as companion volumes. Like the earlier book, *Style in Fiction* was intended to be a course book for students, but it was also an attempt to develop a theory of prose style. It grew out of the teaching of stylistics to undergraduates, which Mick Short and I had shared for several years, and continued to share up to about 1990. Compared with other books, this book was particularly difficult *to write*, but also most satisfying *to have written*. Whereas my books on semantics and pragmatics had tended to receive very mixed reviews, *Style in Fiction* was relatively well regarded by reviewers. As with other books on which I have collaborated, I was especially fortunate in having, in Mick, a co-author with whom I could work closely and well, though inevitably not without disagreements.

In 1980 Mick took the initiative in founding a new academic association, PALA (the Poetics and Linguistics Association), for furthering the study of stylistics. I have been a largely inactive member of PALA since it began, but Mick and I were immensely gratified when, in 2005, our book *Style in Fiction* was awarded the prize for the most influential book on stylistics during the twenty-five year lifetime of the organization.

Again in collaboration with colleagues in the Department, I published in 1982 a book entitled *English Grammar for To-day: A New Introduction* by Leech, Robert Hoogenraad and Margaret Deuchar (Macmillan). This, like *Style in Fiction*, was based on the experience of co-

teaching courses at Lancaster. The book was actually commissioned by the English Association, a national body which at that time was concerned about a decline in the study of grammar in British schools. However, the book was more successful in other parts of the world than in Great Britain. It seems that the British educational world, or at least British students, were not yet entirely ready for the notion that grammar was worth learning.

I was getting used to co-authorship. My experience has always been that a co-authored book is more difficult to write (because of differences of approach, and because of issues to be negotiated with co-authors), but is more satisfactory in the end. *English Grammar for Today* was the third book I had co-written on English grammar. Then, in 1985, came yet another co-authored grammar – much larger even than *GCE*. From about 1978, the 'gang of four' (as we were sometimes jokingly called), Quirk, Greenbaum, Leech and Svartvik, began to work on a second edition of *GCE*. Since *GCE* had been published in 1972, ideas on grammar, and knowledge of English grammar, had moved forward considerably. Moreover, *GCE* had received many reviews, which detailed both its strengths and weaknesses. We authors felt already, then, that there was a need for an updated edition of the grammar. When we started work on it, however, we found ourselves rewriting the whole book, changing its organization, and introducing much additional material based on the Survey of English Usage experience. The second edition of *GCE* evolved into a new grammar, which we named *A Comprehensive Grammar of the English Language* (*CGEL*), and which Longman published in 1985.

In the initial stages of working on this new grammar, three of the 'gang' (Greenbaum, Leech and Svartvik) were secretly producing a Festschrift honouring the sixtieth birthday of the leading author, Randolph Quirk. The book was published as *Studies in English Linguistics: For Randolph Quirk* (Longman, 1980), and contained contributions from eminent linguists and English language scholars in various parts of the world. Unfortunately we failed to keep the secret until the day of publication: the sharp observation of Randolph sensed that 'something was up' a few months before the book was due to be published.

If I had to choose a point in time as the summit of my career, I would probably choose the publication of *CGEL*. The book made a big impact, and began to be treated as 'the authority' on English grammar. Soon after that book was published, I received honours which I might not have received otherwise: I was elected a Fellow of the British Academy (1987), I was awarded an honorary doctorate (of the University of Lund, 1987), and I was made a member of the Academia Europaea (1989). In 1988, I was also made a fellow of my old college, University College London.

7. LATER CORPUS RESEARCH: 1978-2000

Let me now return to the computational work in which I had been engaged on and off since 1970. In the late 1970s, the computational analysis and annotation of computer corpora of English became my main research preoccupation. This continued up to my retirement in 2001: indeed, we still have at Lancaster a small team of researchers, working in, or linked to, a research centre entitled UCREL. (Originally UCREL stood for 'Unit for Computer Research on the English Language'; but in the mid-1990s, to reflect a number of changes including an expansion of research interests, this was changed to the rather ungainly title: 'University Centre for Computer Corpus Research on Language'.) My chief collaborator in this research was Roger Garside, of the Computing Department at Lancaster, who became the Director of UCREL.

After the completion in 1978 of the computer corpus of British English, the LOB Corpus, we were lucky enough to win a research grant from the Social Science Research Council (now the ESRC) to undertake an *automatic grammatical tagging* of the corpus. That is, every word in the million-word corpus was to be assigned a tag indicating its grammatical category, and complex computer programs had to be written for this purpose. We completed the task in 1983, with the collaboration of Stig Johansson and his team based in Oslo, Norway. The software then developed – the tagger CLAWS1 – was the first tagger to employ a statistical algorithm, similar to that of a Hidden Markov Model. This has now become a commonplace method of grammatical tagging; but at the time it was discovered (by one of our researchers, Ian Marshall) we felt that a sudden breakthrough had been achieved. Success in automatic tagging leaped from c.77 per cent to c. 96 per cent. We were able to achieve this by using the Brown Corpus (which had been previously tagged by Greene and Rubin) as a *training corpus* – that is, CLAWS learned its frequency data from the Brown Corpus, and then applied it to the LOB Corpus, which in current terminology would be called the *test corpus*. This was another important advance of a kind: as far as I know, we were the first team to employ (without knowing it) the distinction between a 'training corpus' and a 'test corpus', which is now another commonplace of corpus-based natural language processing methodology.

During the 70s and 80s, the use of computer corpora for linguistic research was becoming accepted by a small group of researchers in the UK, and the research councils were beginning to respond to the need for funding to support programmers and other research staff.

Our next piece of good fortune (in 1983-6) was a grant from the Science and Engineering Research Council (now the EPSRC), to tackle the more exacting task of *automatically parsing* a corpus. In this case we had no training corpus – the Brown Corpus

had not been parsed – so we had to create our own training corpus by hand. A senior colleague at Lancaster at that time, Geoffrey Sampson, had become fascinated by corpus research, and was the first person to build a *treebank* – that is, a corpus, or part of a corpus, annotated for sentence structure. (Soon afterwards, Geoffrey left Lancaster for Leeds and then Sussex, where he continued his own treebanking research, continuing to act as a powerful advocate of the corpus-based linguistic methodology.) The term 'treebank' – invented at Lancaster – has since become established in computational linguistics. However, the automatic parsing task proved to be more difficult than we had imagined, and by 1990, only 13 per cent of the LOB corpus had been accurately parsed, using statistical methods. To do this more successfully, we would have needed a much bigger treebank, a better model of syntax, and more powerful computing facilities.

The task of automatic parsing of unrestricted text data – which is, generally, what the syntactic annotation of a corpus amounts to – was a tough nut to crack, and indeed even now, in the year 2009, in spite of much progress, the problems of automatic corpus parsing (or 'robust parsing', as it is often called) have not been solved. There was obviously a need for larger teams to tackle this important area of computational research, and in 1987 we were approached by a group led by Fred Jelinek, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, to engage in collaborative research funded by IBM. The IBM team was at that time leading the way in developing new technology for speech recognizers, and was also blazing a trail in new and highly sophisticated statistical methods, making use of enormous electronic text collections of more than 300 million words. We were the first to develop, during this collaboration, large-scale treebanks of three or four million words, from which more adequate statistics, and hence more accurate parsing results could be obtained. However, because of our funding by IBM, which also provided the data, we were not at liberty to make the treebanks freely available and to publish results in the way we would have liked. Our work on treebanks was superseded, in 1990s, by more advanced work by Mitchell Marcus's group at the University of Pennsylvania, and the term 'treebank' nowadays is more associated with the 'Penn Treebank' created by Marcus and his team in the 1990s, than with the IBM/Lancaster treebank created in the late 1980s. One book which did come out of this collaboration with IBM, however, was *Statistically-Driven Computer Grammars of English: The IBM/Lancaster Approach*, edited by Black, Garside and Leech, and published by Rodopi in 1993.

During the 1990s UCREL continued to win grants for research projects. This was a time when, it seemed to me, the opportunities of corpus-oriented research were opening up left, right and centre. After word tagging and semantics, other areas such as semantic tagging, anaphoric annotation, and parallel corpus alignment were ripe for exploitation. Although our

work with Yorktown Heights ceased in 1991, when IBM underwent a financial crisis entailing drastic cuts, other projects continued: with a market research firm, with a telecommunications institute (ATR Kyoto), with the publishers Longman and OUP. It was refreshing in some ways to work with private enterprise, and thus to outflank the intense competition for funding by public research councils. But industrial and commercial collaboration brought severe pressures and constraints of their own, particularly restrictions on publishing and the use of research results, and the need to 'deliver on time' according to contract, whatever problems might have been turned up by our research.

The largest project in which I was involved during this period was the British National Corpus (1991-1995), a collaborative project between three publishers (OUP, Longman and Chambers), two universities (Oxford and Lancaster) and one national institution, the British Library. Here, again, the pressures and difficulties were enormous, and the result was a 100-million-word corpus of spoken and written British English, completed 12 months later than scheduled, and in far from perfect condition. Nevertheless, the BNC was an important national achievement, in which the UK led the way, and its lead has been followed by other countries. Only recently has an equivalent American National Corpus been undertaken on similar lines to the BNC.

Also during the 1990s my colleagues Roger Garside and Tony McEnery were engaged, with me, in research sponsored by the European Commission. This was a time when the 'corpus revolution' had really taken off internationally, and there was a big push to capture and annotate the data of other languages, apart from English, in computer corpora. We were involved, with continental collaborators, with the development of parallel (i.e. mutual translation) corpora of English, French and Spanish. Also, as a participant in the EAGLES initiative, I took a role, as chair of various committees and lead author of various documents, in proposing guidelines for corpus annotation for European languages. Between 1993 and 1999, I was engaged in three different projects of this kind, concerned respectively with word-class tagging, syntactic annotation, and dialogue representation and annotation.

8. THE CONTINUING IMPACT OF CORPUS LINGUISTICS ON MY OWN CAREER: 1986-1999

But here I should return for a little to the mid-1980s, to trace my personal career rather than that of the UCREL team. It will be clear that by this stage I had become deeply engaged in computer corpus work, which has a tendency to monopolise the time of anyone who becomes seriously involved in it. This work is very satisfying from some points of view: for example, working in a team of like-minded researchers is something which is normal for scientists, but both novel and stimulating for scholars trained in arts or social sciences. The

excitement of blazing a trail in a new and fast-developing area of research has been a second benefit. Another was that corpus research enabled UCREL to build something of a reputation in *natural language processing* by computer (also termed *language technology*), and to maintain a network of international links with scholars and computer scientists with similar interests. The corpus linguistics world now forms a well-established research community, with links all over the world. Even more amazing to the pioneer of the 1960s and 1970s is the way corpus-based methodology – using a corpus as an empirical basis for investigating this or that linguistic phenomenon – has spread like a contagion, to embrace almost all areas of linguistic research.

On the other hand, corpus research has disadvantages: a considerably amount of the work is painstaking 'donkey work', rather than challenging investigation. The labour-intensive nature of this work meant that for many years I had to put on the shelf some of the research interests I had previously been involved in – notably stylistics and pragmatics. Much time can also be spent seeking funds from research sponsors, to enable research staff to remain in post.

Further, there were until very recently only one or two journals in which one could publish work on corpora. (Now there are many.) One consequence of this was that publications often took the form of edited collections, and because of the team-built nature of corpus research, the editorship of these volumes tended to be a collaboration between two or more editors. Similarly, the individual contributions to the collections tended to be co-authored. There was less room for individual research initiatives; research tended to be steadily incremental, in contrast to the breakthroughs of the early days of corpus research. Yet another regrettable thing, for me, was that the field seemed to advance so quickly that UCREL was finding it difficult to maintain the forefront position it once held. The pioneering days were over, and as resources and software proliferate and become more generally available, we were in danger of being overtaken by the 'big battalions' – departments and teams, whether in the USA, the UK or elsewhere, better endowed in terms of funding, equipment, and staff than we were. In the computational world, we remained a relatively small fish in a big pond. My love affair with computational linguistics (or language technology) faded away, and I reverted to my interest in corpora simply as means of pursuing research into language for its own sake. But since I retired in 2002, UCREL has continued its achievements, under the directorship of Paul Rayson, who took over from Roger Garside on his retirement.

In terms of publications, I played a role in producing four edited collections on computational linguistic topics, all published by Longman, in the period 1986 to 1997. These volumes were *Computers in English Language Teaching and Research* (1986), edited by

Leech and Christopher Candlin; *The Computational Analysis of English* (1987), edited by Roger Garside, Leech and Geoffrey Sampson; *Spoken English on Computer* (1995), edited by Leech, Greg Myers and Jenny Thomas; and *Corpus Annotation* (1997), edited by Garside, Leech and Tony McEnery. Also, I should not forget to mention here another co-edited book in which I took no role apart from that of passive recipient: *Using Corpora for Language Research* (1996), edited by Mick Short and Jenny Thomas. This was a Festschrift to celebrate my sixtieth birthday: I was suitably speechless with surprise and delight when it was presented to me, in front of a hundred colleagues, friends and relatives, on an unforgettable occasion on 16th January, 1996. I had been told that I was to meet a group of visiting Japanese academics, and was completely taken in by this subterfuge.

9. AFTER RETIREMENT

My official retirement came on 31st December 2001, but I retained my position at Lancaster University as an Emeritus Professor. I still have a small office in the Department of Linguistics and English Language, and engage in research, with some PhD supervision and the occasional teaching engagement. My visits abroad for conferences or lecturing engagements also continue.

As I move on into my seventies, I have extricated myself gently from the pressures of running large-scale research projects and large-scale research teams. Between 1998 and 2008, with the assistance of research associate Nick Smith, I was engaged in relatively small corpus projects: a comparison of two corpora (the LOB and FLOB corpora) representing written English in 1961 and 1991, enabling us to trace changes in English grammar over the intervening 30-year period. The project was supported by the Arts and Humanities Research Board (AHRB), and because it is on a smaller scale, it enabled me to do what I think I do best – which is descriptive study of the English language, not leading-edge language technology.

This collaborative work at Lancaster was part of a larger collaboration with two corpus linguistic scholars, Christian Mair and Marianne Hundt (now professors at Freiburg and Zurich respectively). They had initiated in the mid-1990s the creation of two corpora equivalent to the Brown and LOB corpora, except that the texts, of American and British English respectively, were sampled from publications dating from 1991 and 1992. Because they were created in Freiburg University in southern Germany, these corpora were named the 'Freiburg-Brown Corpus' and the 'Freiburg-LOB Corpus', or 'Frown' and 'FLOB' for short. The combination of the four matching corpora (Brown, LOB, Frown and FLOB) enabled us to undertake rather precise studies of grammatical change over the thirty-year period 1961-1991/2, across a generation-gap of American and British speakers.

The collaboration with Nick Smith (as senior researcher) and also with Paul Rayson of the Computing Department (Director of UCREL) continued through the next few years, with the help of small-scale research grants from the Leverhulme Trust. The theme of this work continued to be recent and contemporary changes in English grammar, particularly as revealed through comparable corpora of the design of the Brown and LOB Corpora: a group of corpora we have called 'the Brown family'. We have so far succeeded in creating a corpus of British written English 1928-1934 (centred on 1931) appropriately named B-LOB ('before LOB'). Work on a 1901 corpus is continuing, and a colleague in Linguistics, Paul Baker, has compiled from World Wide Web sources a matching corpus from 2005-6. Hence, by the use of these comparable corpora, we are able to trace how, in British and (to a lesser extent) in American English, the use of grammar has been changing from the beginning of the 20th century to the beginning of the 21st century: a century of linguistic history. But there is still much work to be done on these corpora.

10. CONCLUDING REMARKS

While these days I remain deeply engaged in computer corpus research, I hope to continue in the future my interests in English grammar, literary stylistics and pragmatics (particularly politeness theory). I still give lectures on politeness and pragmatics, and hope to have to update my research in this area (particularly adding the dimension of cross-cultural pragmatics). Two or three books – including new editions of old faithfuls – are in the pipeline.

During my years at Lancaster I have enjoyed the opportunity to visit many parts of the world as a lecturer and teacher. Sometimes I have spent longer periods overseas, of which the most memorable was a period of three months spent leading an academic delegation in China in 1977. At that time I taught for three or four weeks in each of four great cities: Beijing, Xian, Nanjing and Shanghai, and since then I have made three further visits to China. Earlier, in 1972, I had spent 6 months at Brown University, USA, as a visiting professor, at the invitation of Nelson Francis, the 'Grand Old Man' of corpus linguistics, whose brainchild was the Brown Corpus. Other countries where I have had visiting appointments are Japan (Kobe University, Kyoto University and Meikai University), France (Paris VI, the Sorbonne), Australia (University of Queensland, Brisbane), New Zealand (University Canterbury, Christchurch) and Singapore (National University of Singapore). Each of these extended visits has enriched my life and, I feel sure, deepened my understanding of languages and cultures. The same could be said of the visits to more than thirty countries I have made as a lecturer or conference participant.

It is inevitable that, as one gets older, one's interest in the past increases and one's interest in future trends decreases. The current year, 2009, has been something of a nostalgic year for my involvement in corpus research. Lancaster University, for the first time since 1984, hosted the annual ICAME conference this May, and there was a special celebration to mark the 30th conference and the 32nd year since the organization's foundation. Three of the five founding members of the organization (Stig Johansson, Jan Svartvik and myself) were present and took part in a historical presentation and exhibition entitled 'The Coming of ICAME'. Later this year, in July, will be another corpus-related celebration: a symposium hosted by UCL commemorating the fifty years since the founding of the Survey of English Usage by Randolph Quirk (who, it is hoped, at the age of 89 will put in an appearance at the symposium). This year also marks the fortieth anniversary of my arrival at Lancaster University. I have seen it grow from a small, undistinguished, young university (founded in 1964), to a large established university which (as its website tells us) is 'a leading higher education institution', and 'has won international recognition for the quality of its teaching and research'.

Meanwhile, I continue to publish and to work on publications. In 2008 Longman published a book of mine, *Language in Literature*, which was largely a collection of papers published in obscure places in earlier years, combined with three new chapters. Due out later this year, a book co-authored with Marianne Hundt, Christian Mair and Nick Smith (now at Salford University), and to be published by CUP, will be the culmination of our corpus-based researches on recent changes in English. It will be called *Change in Contemporary English: A Grammatical Study*. I am also working on a book for OUP on politeness in English, compared with other languages. While I am fortunate enough to remain healthy, I would like to continue my research and publication in the fields of corpus linguistics, stylistics, the pragmatics of politeness, and English grammar. As far as I am concerned, 'old professors never die, they merely fade away'.

<div align="right">Geoffrey Leech, 12 June 2009</div>