

David Crystal (Bangor)

Learning to be a linguist: the growth of linguistics in Britain in the 20th century¹

Abstract: In this paper the author traces the development of linguistics in Britain, gradually emerging from the philological tradition to become a multi-faceted discipline in its own right, giving language a reality hitherto underdeveloped in traditional philology. Drawing many examples from his own experience, various features of British linguistics are identified, e.g. the empiricist tradition working with “real” linguistic data, resulting in a critical attitude to theories lacking a sound empirical foundation, the recognition of the usefulness and applicability of linguistics in society, and finally the collaboration between linguistics and other disciplines with an applied emphasis of the subject, a strong feature carried over into the 21st century. Looking to the future, the fascination of linguistics is seen in the fact that society and language are permanently changing, resulting in a constant flow of new, challenging and often unpredictable fields of study.

It is a delight to be here in Kiel to talk about “Classics Revisited”. However, what do you have to be in order to be classic? Dead – I thought. That’s the usual criterion. Thomas A. SEBEOK, when he was editing the *Classics in Linguistics*² series many years ago, faced a real difficulty over this same question. One of the volumes was to be on biographies of famous linguists, and they all had to be dead of course – but then the editorial board had a problem concerning Chomsky. Can one have a *Classics in Linguistics* series without Chomsky in it? He told me that the editors had a discussion about whether to classify Chomsky as being ‘honorarily dead’, just for the purpose of the book!

In order to prove that I am still alive, I will begin with a look at the history of linguistics, as I have experienced and perceived it, but then I want to extrapolate from that, to relate to things in language study that are going on now in which I have been mainly involved. The historical perspective is an essential starting point, and fascinating in its own right, but I want to use it as a source of guidelines about how to behave now, as linguists, and how to plan for the future. I will illustrate from two new topics in linguistics, topics which I would never have imagined to

1 This chapter is an edited transcription of a lecture David Crystal gave at the University of Kiel on 20th June 2011.

2 Editor’s note: The reference here is to SEBEOK’s series *Current Trends in Linguistics* (1963–1976).

exist, even ten years ago. This always seems to me to be the best aspect of linguistics: you never quite know what is going to happen next. Whatever language is like today, it is going to be different tomorrow.

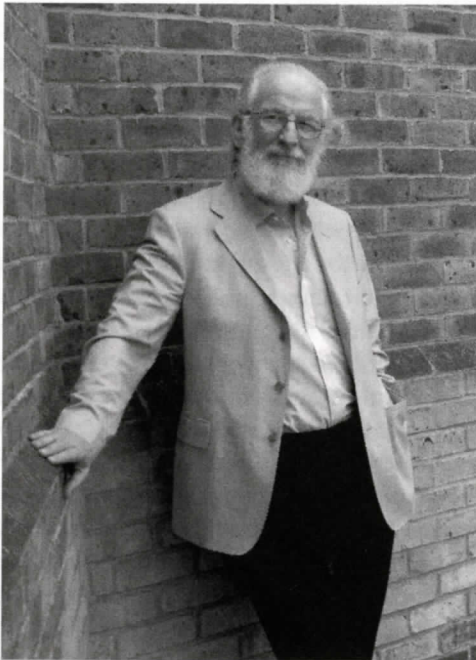
To take a modern example: imagine that someone had told you in 2005 that the next big thing on the internet was going to be online text messaging, that all those text messages we send in our mobile phones, just between ourselves, were to be put up on the internet so that *everybody* could see them. The number of characters would need to be reduced slightly because on the internet personal identity would need to be shown. So instead of 160 characters it would probably have to be 140. And the whole thing would be called by some enticing name – like Twitter, for instance. Would you not have said that the idea was crazy? Who would ever want to read someone's texts on the internet? But what has happened? Twitter – a short-messaging internet service – has become the most rapidly growing area of internet activity in recent years. And it presents linguists with a new domain of study. As I say, one can never predict what is going to happen next, and that is the excitement of linguistics.

Way back in the middle of the 20th century, there was an even greater degree of unpredictability. That period was marked more than anything else by remarkable diversity among those who were involved in language study. There was no linguistics, as we know it today, in Britain. There were no taught degree courses in linguistics in any British university, no introductory textbooks, and very few people who would have wanted to call themselves 'linguists' in the modern sense. This became especially obvious when, in 2002, the *Philological Society of Great Britain* produced a volume of personal histories (BROWN/LAW 2002), and all the important linguists of that period – those that were still alive, that is – contributed essays. A comparison of the views shows that, almost without exception, their story was one of not having a clear idea of what linguistics was. Virtually everybody started with the same general background: an interest in languages. That was it. The only kind of intellectual perspective they had available to study language in those days, as in most other parts of the world, was that derived from philology or modern language teaching. So they were aware of the historical tradition in language work, but little more than that. In the volume, people repeatedly say, for example, "I started out in Classics". Virtually two-thirds of the people in the book began as classicists. Some began in modern languages, and I am no exception.

My background, in a nutshell, was this: I was brought up in Wales in the 1940s; and in Holyhead, where I lived, which is in northwest Wales, the port that you'd use if you were taking the ferry to Dublin, Welsh was spoken quite substantially. It still is: 60 percent of the population in that corner of Wales speak the language, along with English. During my childhood, Irish too was spoken in Holyhead,

because Liverpool and the North Wales coast had long been places where the Irish had come over to get jobs, and Holyhead, at the end of the railway line, was as near as Irish immigrants could get to the 'Old Country' without, as it were, getting their feet wet. My mother was English-speaking, but my Uncle Joe was Welsh, and so he began to teach me a bit. Of course, when you went to school, you learned some Welsh there as well; and so I became confidently semilingual by the time I was six or seven, and would have become a fluent speaker of Welsh, if things had not taken a turn in a different direction. At the age of ten my family and I had to move to Liverpool where I encountered the local English dialect, *Scouse*. I arrived in Liverpool with a strong Welsh accent, as Welsh as you could possibly imagine, with a rising tone at the end of statements, distinctive vowels, and so on. Within ten days, that accent was beaten out of me; I, the newly arrived *Taffy*³, was forbidden (by the other children) to talk differently from them, and I quickly became a *Scouser*. My Welsh of course began to deteriorate at that point, which is what inevitably happens when you are away from a language for too long.

Abb. 1: David Crystal (Photo courtesy of Hilary Crystal)



3 *Taffy* is the colloquial name for a Welsh person.

The other language I was exposed to very early on was Latin. I was brought up Roman Catholic and served Mass as a young boy, which was a linguistically interesting experience in those days. The interaction between the priest and the altar server was entirely in Latin, and so I had to learn all the responses by heart. I became fluent in liturgical Latin without understanding a word of it. It was phonologically excellent, grammatically excellent, semantically very poor, and pragmatically absolutely useless. But it did intrigue me, this Latin. All languages do, at that age; they are calling to you. Chomsky used to say we are born with a language acquisition device (a LAD) in our heads. That was never an adequate specification. We are born not with a LAD, but with a MAD, a multilingual acquisition device in our heads. Given that three quarters of the world's population are naturally bilingual and that half the world's population is naturally trilingual, I see young children as having a brain that is looking for languages and wanting stimulus from languages. If they get it, then it is a matter of opportunity, motivation, and other well-known factors how far they are able to develop these experiences. In my case I had encountered four different languages by the time I was seven – English, Welsh, Irish, and Latin – and would later have the chance to explore all of them further. I found languages curious. I was intrigued by the strangeness of the situation in church, by the fact that I could understand the priest at one level while not understanding him at another, and wondering about the reason for that ambiguity. Of course, in secondary school I soon came to realise what was going on.

There I learnt French and Latin, properly this time, and Greek, and swapped expressions with people in the school who were learning other languages, including German. My German is what one might call 'bicycle German'. It is a German that is learnt when you are travelling home on bicycles with your best friend. You are in the class that learns Greek; he is in the class that learns German. And so on the way home, you swap phrases. He would say *ich bin* something or other, and I would say the equivalent in Greek: Thus he would learn a smattering of Classical Greek and I would learn a smattering of modern German. For us, that was good enough.

This kind of background which I have illustrated for myself is absolutely typical of all the early backgrounds of the people that you find in the *Philological Society* publication: varying levels of proficiency in a varying number of languages. After degree courses, we would all of us end up in linguistics, but not usually having any clear idea of what it was – knowing that it was different from philology, but not being entirely sure about its status or future direction. We knew there was quite a lot of relevant literature 'out there', written mainly by Americans (such as

Bloomfield), sometimes by Europeans (notably, Saussure), and we would read as much as we could. But there was a tension, as there has often been since, between the old discipline of philology and the emerging discipline of linguistics. There was a kind of mutual suspicion between those who had been reared in the philological tradition and the 'neo-linguists', who sometimes did not respect that tradition and who wanted to do things that were very different.

I encountered that clash very early on in my university career. It was when I went to University College in London (UCL) to start my English degree, a course which was ideal for me because it was fifty percent language and fifty percent literature, and thus balanced my two interests. We were taught Old English and Old Norse and several other languages of that period, but in a very philological way. Sound changes were letters on the page rather than vibrations in the air. I remember very early on in my Old English course going up to the lecturer and asking him about the word *cyning* (OE 'king'). I had encountered this word and read the descriptions of it, which included the fact that it had a high front rounded vowel. But I was not quite sure what that was, never having done any phonetics. So I went to my lecturer and said to him, "How do you pronounce this word?" pointing it out. He replied: "Oh yes, yes. That's the word for 'king'." Whereupon I said, "Yes. But how do you pronounce it?" He answered: "Well it's a very interesting form. The high front rounded vowel is the result of..." And he went on about mutations and so forth. At the end I asked again, "Yes, but how do you pronounce it?" and he told me, "Well, we don't actually know for certain, because there are many opinions as to just how high and how front the sound change was, and so nobody could actually give you a definite decision about the actual pronunciation of this vowel, and ..." I tried once more: "But could you just give me an example, just once?" He would not.

Things have changed now. Teachers are happy to speak old languages aloud, notwithstanding the uncertainties. You can hear *Beowulf* read aloud, plausibly and confidently, in reconstructed pronunciation, as well as Chaucer, Shakespeare, and other writers. But at that time, it was different: the language was 'on the page', intellectually supported by a huge amount of linguistic metalanguage, tradition, and orientation, but not related to the way language was encountered in the real world. And more than anything else I wanted to encounter languages, ancient or modern, in the real world.

That sort of experience almost put me off linguistics forever. But I was saved by the place-name people. The *English Place-Name Society* was based in University College at that time, and the *Survey of English Place Names* was well underway, run by people like Hugh Smith and John Dodgson. What Dodgson did was very

clever. He would say, "We're going to study place names this week" – of, for example, some place in Oxfordshire.⁴ "I'll be in such-and-such a pub ...", and he would tell us a location and invite anyone interested to meet him there at a particular time. I remember going by bus one weekend to the selected place and pub, to find Dodgson sitting in the corner. On the table in front of him were clearly more pints of beer than there were people. When asked what the pints were for, he explained that they were "elicitation materials." I did not quite understand, but all soon became clear. Looking around the pub, he spotted a man who to us students looked ancient and who was identified as our target. Dodgson went over and invited him to come and have a drink, and the old man accepted and joined us. Then the interrogation started:

"You from around here?"

"Yes."

"Where do you live?"

The man would name some place.

"That sounds an interesting place. What does it mean?"

By degrees, our informant told us about the names of the fields, lanes, hedgerows, and all sorts of local places. Sometimes we were getting names for locations that we would never have thought would even have a name. Everything was recorded. And I realised that here was a reality to language which was entirely different from that encountered in philology classes.

Due to these classes, at the end of the first year of the university course I managed to achieve a kind of balance between language on paper and language in the field. As I then continued through that University College period, I realised something that is also mentioned several times in the memoirs of the other linguists from the *Philological Society* publication: that there was tension even within linguistics. It had nothing to do with rivalry with scholars from across the Atlantic (as one might have expected, given that it was 1959, and Noam Chomsky's *Syntactic Structures* had come out two years earlier).⁵ Instead there was a tension within London itself, between two big figures who had established themselves in linguistics in those days: on the one hand, Daniel Jones, who had started the phonetics department at University College and who had built it into a hugely

4 In the 1970s to 1990s John Mc N. DODGSON published seven volumes on dialects in Cheshire with the *English Place-Name Society* (volumes XLIV-XLVIII, XIV and LXXIV together with A.R. RUMBLE). The Oxfordshire dialects found their way into earlier volumes from the 1950s, written by Margaret GELLING (XXIII and XXIV).

5 This tension between competing research across countries would only develop later.

important centre; and on the other hand, J. R. Firth at the School of Oriental and African Studies, also part of the University of London, who was developing his ideas and building a department focusing on language generally. The resulting conflict was between phoneticians on the one side and general linguists on the other. Although they theoretically complemented each other, and might thus have profited from each other, they were severely divided by the question of definitions. Daniel Jones – D. J., as everybody called him – believed in the phoneme. He lived it and accepted no other way of looking at phonology. Firth, however, had developed a totally different view, which he called ‘prosodic phonology’. Its aim was not to identify the individual units, but to see the way in which sound features stretched over syllables and words and even longer stretches of speech. These two approaches were not just seen as alternative models. They were a starting point for serious turf battles. I remember being told not to go anywhere near the School of Oriental and African Studies (SOAS) and its ‘heretical’ notions, while people from SOAS were warned against attending seminars in University College for fear of being turned into phonemicists. Personally, I found it extremely enervating that people would be so passionate and that they would harangue each other to such an extent. In the local Marlborough Arms, midway between the two institutions, I would hear people whose books I read and valued, and who were leaders in the field, being condemned out of hand – in a way that I now know to be perfectly normal academic criticism!

I graduated in 1962, having begun studying at University College in 1959. This was a very important year for British linguistics, because the *Linguistics Association of Great Britain* was formed. It had emerged from a small group of linguists located at the University of Hull, but it quickly became a national institution. It was the first professional organisation of its kind in the country. The year 1959 was also important because Randolph Quirk arrived at University College to start his *Survey of English Usage*, which gave me my first job some years later.

The degree I obtained at UCL was a perfect kind of degree for somebody like me, and many other people have had the same experience. There was a balance first of all between language and literature: fifty percent *lang*, fifty percent *lit*. That has shaped me ever since, to the point where I think of my brain as being split between the two of them. They are two sides of one coin. I have never been able to separate them, and have never wanted to. Unfortunately many institutions, especially in mainland Europe, don’t share this view of the two disciplines. I was once invited by the language department of one university to give a talk on language play (‘ludic linguistics’, as I would later call it). Upon my arrival at the linguistics department, I was told that the literature department had heard I was there, and

they would also like to have a talk from me. I considered this and suggested that, as my speech was actually on language play and its relationship to literature, it would be equally relevant for both departments, so why not have them join together for this occasion. This was refused outright. I ended up giving the same talk twice on the same day, once to the language people and once to the literary people.

At University College, on the contrary, *lang* and *lit* were very much interconnected. There were also the beginnings of a desire to integrate philology and linguistics, the common ground being a focus on the detail of textual analysis in both speech and writing. It was possible to choose options that gave you a stronger emphasis on the *lang* side or the *lit* side, and phonetics was one of those options – the indispensable foundation, as many people have called it. That was the subject for me. I became a phonetician.

Working for the *Survey of English Usage*, which I joined after graduating in 1962, was to me the closest thing possible to heaven because it relied fully on corpus work. It brought together all the elements of my training, linguistic and phonetic, and made me part of the empiricist tradition – the one term that I think best characterizes the British approach to linguistics. British linguistic research stood out for its concern for description and for identifying, at levels of detail that were not available as intuition (and thus very different from the approach associated with Chomsky), points of interest to do with language structure and use. When I started work at Reading a few years later, this was a major motivation for the direction of the department.

Doing corpus work means being able to live and work with a certain messiness of data. There are problems; there are exceptions; language does not always work according to the rules one initially formulates. As a consequence, anybody who adopts that kind of approach quickly develops an early scepticism about schools of thought. And that is a second characteristic of the British linguistic tradition: a deep-rooted critical attitude. Any new theory is going to be met openly, but is also going to be immediately questioned. What sort of hypotheses does it generate, and do the hypotheses actually account for the data available? Do they account for exceptions? Most British linguists of the time displayed this kind of attitude. Of course since then the subject has evolved, and there are linguistic theories and theoreticians in Britain as anywhere else in the world. But a certain scepticism about the generalizations being made about language as a whole, especially those emanating from Chomsky, was a very definite feature of the linguistic atmosphere I found myself growing up in.

The third big feature of British linguistics was a general concern that linguistics was not simply fascinating as an end in itself, but actually useful. Long before

applied linguistics started up in many parts of the world, it was already strong in Britain. The *British Association for Applied Linguistics* became an institution in the 1960s, and since then has been very prominent. It began with a particular interest in English language teaching (to foreign learners), and also in such well-established fields as lexicography. The *Oxford English Dictionary*, for example, had a strong tradition dating back almost a century⁶, and the new generations of lexicographers were very concerned to develop their ideas in relation to the latest thinking in linguistics. But there was also a very real interest in applying language studies to educational issues, to mother-tongue education, in particular, and to such topics as literacy in schools. At Reading, another branch of applied linguistics opened up: we introduced a degree course where linguistics became the foundation of training in speech and language therapy, the teaching of the deaf, and remedial language teaching.

One other characteristic of British linguistics at the time is that it was very collaborative. Single-subject degrees in Linguistics were rare. The Reading one was the first, and then a few others followed. But for every single-subject degree course in Linguistics, there were several joint degrees: Linguistics and French, Linguistics and German, Linguistics and Philosophy, Linguistics and Sociology, and so on. The idea behind these seemingly indefinite combinations was to see to what extent there were points of connection between these different intellectual traditions and what kinds of 'product' would emerge. At the seminar level there was no limit. Students as well as staff members were actively encouraged to go to departments that had never talked to linguists before and have a seminar with them, just to see what would happen. So there were seminars between linguistics and chemistry, for instance. Against all odds, some of the outcomes were remarkable. Three were, to me, especially outstanding.

Firstly, there were some fantastic seminars between linguistics and music. At the time musicians such as Luciano Berio⁷ and Karlheinz Stockhausen⁸ were

6 The *Philological Society* of London had first initiated the project in 1857. The volumes of the first editions were published between 1884 and 1928.

7 The Italian Luciano Berio was a composer and pioneer of electronic music. He also taught in the US and founded the Julliard Ensemble there. His work is characterised by adaptations and reinterpretations of existing themes, musical pieces and literature into, for example, collages.

8 The German composer Karlheinz Stockhausen figures among the most important composers of the 20th century. He is one of the founders of Punctualism (or pointillist music), and his work is seminal in the areas of electronic music, aleatoricism in serial composition, and musical spatialization.

trying to explore the properties of the voice in their musical composition, but lacking a phonetic metalanguage in order to do it. There was interest in seeing how phoneticians actually transcribed things like the prosodic features of the voice: intonation, loudness, rhythm, speed, paralinguistic features, and so on. This mutual interest resulted in a series of fascinating seminars with the Canadian composer Istvan Anhalt, and explorations of fresh kinds of phonetically inspired musical notation.

Secondly, there was collaboration with typography. Typographers also lacked a metalanguage for describing the linguistic effects conveyed by the choices made in their field. They know the effect that a particular choice of a typeface has compared with the choice of some other typeface, but have difficulty describing exactly what effect that is. So there were some really interesting discussions about the kind of metalanguage that might evolve in order to talk about typographical differences.

The third example was in art, where a number of seminars explored ways of visualising language. This is also a problem all linguists have to face: how to present language in a visual way, in a textbook, on a screen, on a poster... How does one show the way all the bits of language come together – phonology, syntax, semantics, pragmatics, and so on? Linguists were, and still are, searching for good models, and their inspiration for model-building traditionally comes from the frame of reference developed in the philosophy of science. Artists, on the other hand, have a very different conception of model-building. As a result, they came up with all sorts of interesting visual representations of language which, as a linguist, I had never conceived of before.

These examples give an idea of the atmosphere of the Reading department in those days. It was not alone, because similar work was going on in Edinburgh, Sussex, York, and various other parts of the country. In my view, it is the interaction between disciplines, and the applied emphasis of the subject, that has carried through more than anything else in Britain into the 21st century. I will illustrate this by means of three areas which are novel, though displaying continuities with the tradition. All three are examples which, at the turn of the century, I had never dreamed were possible. And, as I said before, this is the excitement of it: to encounter domains of language use which have not been explored before.

Students sometimes tell me that they feel that in linguistics everything has already been said. On the contrary, I reply: new subjects are constantly coming along, and I use the internet as the prime example. With the development and vast success of the internet, virtually everything in language has to be rethought. The internet is now the largest linguistic corpus there has ever been: there is more written language on the internet now than in all the libraries of the world

combined, and little of it has been explored from a linguistic point of view. What is more, the major trends in current technology are going to offer linguistics even further horizons over the next twenty-five years – in particular, we will see its increasingly oral character (through audio and video). Keeping pace with it all will be the main challenge. The quantity of information on the internet is already increasing at an extraordinary rate: there was a time when, according to a study by IBM, the amount of information on the internet was doubling about every twelve hours.⁹ One reason is the way individuals are uploading and posting huge amounts of material. Trivial bits of information such as *Here is me, having breakfast and eating cornflakes* are posted on social platforms or filmed and uploaded to YouTube. With millions of people doing this every second of the day, the content of the internet is expanding in unprecedented ways. And there is a language element in everything.

Let me illustrate what I mean by saying “virtually everything has to be re-thought”. We rely chiefly on the *Oxford English Dictionary* (OED) to tell us the first recorded instance of a word in English. The original OED lexicographers, however, were able to do only so much with the very limited tools at their disposal. They chose certain authors that they felt were ideal sources for the dictionary – and Shakespeare was of course pre-eminent. So it is often the case that he is listed as the first recorded user of a particular word simply because his texts were studied more than anyone else’s. As a result we see people saying that Shakespeare introduced over 2,000 words into the English language. It only takes a few moments of analysis to realise that this could not possibly mean he invented them. For instance, he is the first recorded user of the swear-word ‘*Sblood*, meaning ‘God’s blood’. But that is not a Shakespearean invention: he was simply the first person that the lexicographers had found to have written it down. Today, as more and more literature from that period has come online, we can carry out the same task that the OED lexicographers did, only with a much larger database at our disposal and with far more rapid search techniques. A Google search for any particular item will quickly reveal whether it turns up in any of the plays, poems, or prose texts that precede Shakespeare. And what we find is that there are indeed often precedents. For example, the word *lonely* was for a long time a first recorded instance in Shakespeare, but this has now been discovered in several texts from the 1560s and 1570s. So the total number of word inventions attributed

9 In July 2006, IBM published a white paper in which, relying on a projection for 2010, they claim that by then “the world’s information base will be doubling in size every 11 hours” (online: www-935.ibm.com/services/no/cio/leverage/levinfo_wp_gts_thetoxic.pdf).

to Shakespeare is getting lower. It's still a large number, and it doesn't diminish Shakespeare's creative position one jot – in fact it enables us to be more precise about the nature of his linguistic creativity. But the general point emerges: thanks to the internet, even a traditional subject like historical lexicography can present fresh opportunities for investigation.

The internet has itself generated great fascination as far as language is concerned. Its corpus needs to be described, but very little descriptive research has been carried out so far, especially of a cross-linguistic character. There are new kinds of problems to be dealt with, such as the speed of change. To give one example: I know somebody who started work on Twitter in 2007. This was going to be cutting-edge linguistic descriptive work, using the latest technology. He collected a large corpus of tweets, and started to describe what was going on. He was making good progress, but in November 2009 Twitter changed its policy. The prompt for Twitter originally was *What are you doing?* This elicited tweets in the form of *I'm eating cornflakes* or *I am stuck in a lift*.¹⁰ Then in November 2009 Twitter changed its prompt. Instead of *What are you doing?* it became *What's happening?* This resulted in a fundamental change in focus, which had immediate linguistic consequences: instead of looking into yourself (lots of first-person pronouns and present tenses, for example), you are being asked to look outside yourself (lots of third-person pronouns and other verb forms). From that moment Twitter became more of a news service: there are now far more news reports and advertisements on Twitter than there were before. So, suddenly, the linguistic character of Twitter changed. As a consequence, the graduate student who thought he was doing synchronic linguistics turned out actually to be doing historical linguistics, because his corpus covered a period pre-2009.

This illustrates a core problem: there are many areas of internet activity which have grown up in the last twenty years but which are now strictly historical in character because the technology has moved on, and various other things have taken their place. Most of the early internet period is actually inaccessible now, which is a lost opportunity for linguistics. It is like losing important manuscripts from the medieval period in a fire. All these data have electronically disappeared, so some of the evolutionary trends that happened on the internet will remain linguistically undocumented. That is why it is so important that as much descriptive

10 This particular tweet became famous when posted by British comedian Stephen Fry in 2009 (online: www.telegraph.co.uk/news/celebritynews/4518473/Stephen-Fry-posts-Twitter-updates-while-trapped-in-lift.html).

work as possible should be done now, because change will recur constantly over the next few years.

That was one area which we could not possibly have anticipated a decade ago, or rather 1990, to be precise, because everything has happened basically since then: the World Wide Web in 1991; the first emails (for most of us) in the mid-1990s; chat rooms in the late-1990s; Google in 1999; text messaging around 2000; blogging in the early 2000s¹¹; Facebook 2004; YouTube 2005; Twitter 2006. This is a time span of 20 years, and it produced so much novelty awaiting description. The result was a new branch of the subject: internet linguistics. And then we see the arrival of an applied internet linguistics, in line with the British linguistic tradition described above, where the study of language because it is interesting is supplemented by the study of language because it is useful. Several new and interesting ways in which internet linguistics may be useful have emerged over the last few years.

Applied linguistics evolved in order to solve problems. Applied linguists use the theories, methods, and findings of linguistics to provide a solution to a problem that is not of a linguist's devising, but which has been identified by (for example) a language teacher, a speech therapist, or a lexicographer. The internet has already generated several problems for which solutions need to be found, problems that we encounter in everyday life. We might, for example, do an online search and, upon failing to find the hoped-for hits, think "couldn't Google do better?" The answer is that often, using the algorithmic techniques employed, a better result can't be obtained. Might an applied linguistic approach obtain better results? Here is an illustration of when it did.

The problems I meet often start with a phone call. I was rung up in the late 1990s by an internet advertising agency who told me about the following problem:

"We are putting ads on web pages, and the problem is that the ads are being inappropriately placed, which is embarrassing for everybody."

"An example, please?"

"There was a page on CNN the other day about a street stabbing in Chicago. The ads down the side said *Buy your knives here. Get your knives on*"

The implication was clear: *if you buy our knives, we can stab you better than anybody else's knives can*. The cutlery company had complained: they did not want their ads to be placed on that page, or on any page that had negative associations.

11 The word *blog* was invented in 1997, but blogging was not widespread until the creation of Blogger in 2000.

They blamed the ad agency, which then lost the contract. The agency was desperate to find a solution, to stop this kind of thing happening again.

That was the first of thousands of similar examples. They can still be found nowadays. I remember when the Iraq war started, the ads down the side of web pages were still saying *Come to sunny Iraq!* A page reporting a serious car accident on a motorway caused by a mechanical failure was accompanied by an ad for the same sort of car that was in the accident. These situations cause embarrassment for all the parties concerned.

How do you solve it? This calls for a linguistic analysis which reveals that words (and thus, ads) turn up in the wrong places because the software used is not semantically sophisticated. The programme, in the first case above, clearly was looking only for the most frequently occurring words in the CNN article, and it found several instances of *knife*. The software then scanned an inventory of ads, found some that contained the word *knife*, and put the two together. Nobody had thought to take into account that *knife* is polysemic, has more than one meaning: *knife* as 'a weapon' and *knife* as 'an eating tool'. That kind of lexical awareness, although elementary, was missing from the simple algorithms that the search engines were using. The solution lies in developing a different kind of search engine, one that is linguistically aware. It's something that has now been done, thanks to the project which followed that initial phone call. It involved a comprehensive lexicographic investigation of English (and later, other languages) in which the entire polysemy of English was comprehensively analysed. The result I called a 'sense engine', whose role was to analyse all the words on a web page in order to find out what it is *really* about. The CNN webpage was really about homicide, policing, dangerous locations, and suchlike; so notions like personal safety are the underlying themes. If there are to be ads down the side of that page, they should be something to do with those. But only a full semantic analysis of the lexical content of a page is going to produce an understanding of what the page is about, so that relevant ads can be selected from the inventory. And as web pages can be on any topic, this means anticipating the polysemy in all available words. It took several years, and a large team of lexicographers, but it was eventually completed, in the form of commercially available products such as iSense and Sitescreen.

The name Sitescreen indicates the existence of even more worrying cases, where ads appear on pages that contain sensitive or controversial content. For example, most advertisers do not want their ads to appear on a pornographic page. An extreme case would be if a children's clothing manufacturer's ad appeared on pages talking about (or advocating) paedophilia. To stop the ads appearing in such cases, the exact same approach has to be followed: an in-depth analysis of

the semantic structure of the page. The word *children* turning up on a page is not enough to conclude that an ad for children's clothing would be appropriate. This is where applied linguistics actually becomes really interesting, if not dangerous, because the linguist has got to analyse the structure of the page, whatever its content. You have to do a descriptively detailed study, and not only of pornography, of course, but of any topic where the content is controversial, such as drugs, gambling, alcohol, racism, hate sites, they all have to be studied if one wants to stop ads appearing on those particular pages. Never having seen a linguistic description of a racist, gambling, or pornography site, it was one of the things I had to explore. I spent quite a long time compiling and analysing corpora of such areas as pornography, gambling, and drugs, to the extent that I worried sometimes about whether my activity might actually suggest, to an unaware observer, that I was engaged in nefarious practices!¹² Anyone engaged in such work has to make it absolutely clear in advance (not least, sometimes, to the security services) what one is doing, and why one is doing it. Or to one's spouse! My wife would come into the room as I was working on pornography data and ask what I was doing. I would reply: "Darling, I am not looking at her body. I'm looking at the words being used to describe her body." It's all applied linguistics.

This is the sort of 'real linguistics' I referred to earlier. Linguistic research can actually end up doing something extremely useful in trying to help a domain which was, previously, linguistically naïve. There has of course been a great deal of sophisticated grammatical work carried out in computational linguistics and natural language processing, but most internet technology approaches in the 1990s were semantically and pragmatically primitive, and still are. There is plenty of work to be done here in further developing the field that has come to be called 'semantic targeting'.

Let me now give one other illustration of the way applied linguistics moves in unexpected directions, from another project that began unexpectedly. I never anticipated that historical linguistics, or philology, would have a real-world application. Philology is one of those subjects that is studied because of its intrinsic interest and its contribution to knowledge. I had never thought it might be useful to anybody, as opposed to being just fascinating. But in 2004 I found myself

12 This work culminated in the foundation of *Crystal Semantics Limited*, a company founded together with Ian Saunders in 2001 and taken over by 24/7 Media in 2013. It uses a patented Search Engine™, which relies on the practice of Semantic Targeting to furnish appropriate advertising placement solutions. It has also developed SiteScreen^{ASP}, which aims at protecting brands. It has been translated into 11 languages (so far) and won the LT-Innovate.EU in 2013.

suddenly sucked into a world which I have since come to call ‘applied historical linguistics’. The initiative came from Shakespeare’s Globe in London. Again, it started with a phone call, this time from a director at the Globe who said that they were interested in putting on a Shakespeare play in original pronunciation. The Globe is an attempt to replicate the Globe Theatre of 1599. It opened in 1997 on London’s South Bank, and there are now other replica Globes in various parts of the world. The motivation was to go for ‘original practices’ as much as possible, original movement around the stage, original dress, original music with original instruments, and so on; but not until 2004 was there a proposal to put plays on in the original pronunciation. When I asked why it had not hitherto been attempted, I was told:

It’s a risk for us. An accent of four hundred years ago is going to be so different from modern English that people will not be able to understand it. The Globe, as an open-air theatre, is only open six months of the year, and we need to have full houses. If we put on a production which is unintelligible, people will not come, and it would be a disaster.

However, the director had persuaded the company to experiment with it by pointing out that if the Globe did not do it, some other theatre would. (Theatres are always in friendly competition with each other!) I was able to demonstrate to them that, despite the four hundred years that have passed, original pronunciation (OP) is not as different from modern English as they were expecting, and they found it perfectly intelligible. The decision was made to put on a production of *Romeo and Juliet* in original pronunciation. It was a great success, and since then, the OP movement has grown.¹³ Another OP play was produced in 2005, *Troilus and Cressida*, and since then there have been OP productions in other parts of the world. In September 2010 the University of Kansas put on an OP production of *A Midsummer Night’s Dream*¹⁴, and in November 2011 the University of Nevada presented an OP production of *Hamlet*.

Let me give you an example of OP compared to modern English pronunciation, so that you can appreciate the point about intelligibility. We’ll take some of the opening lines of *Henry V*. The chorus comes to the front of the stage and introduces the main topic of the play, namely the Battle of Agincourt. There are some important differences in pronunciation, but the overall effect is readily accessible to a modern audience:

13 More information can be found in CRYSTAL (2005) or on the website accompanying the book.

14 Information on the production is given by Paul MEIER (2011).

Text	Modern English (RP)	OP ¹⁵
<i>Oh, for a muse of fire that would ascend</i>	əʊ fɔ:ɹ,ə mju:z əf,fa:ə ðət wʊd, əsɛnd	o: fʊ,ɹ,ə mju:z əf fa:ɹ ðet wʊd əsɛnd
<i>The brightest heaven of invention!</i>	ðə braɪtəst hævn əf ɪnvenʃn!	ðə bɹɛtəst ɛvn əf ɪnvensjɔn!
<i>A kingdom for a stage, princes to act,</i>	ə kɪŋdəm fɔ:ɹ,ə steɪdʒ, pɹɪnsəs tə ækt,	ə kɪŋdəm fə,ɹ,ə ste:ɹdʒ, pɹɪnsəs tə ækʔ,
<i>And monarchs to behold the swelling scene!</i>	ænd mɒnɑ:kəs tə bihəʊld ðə swɛɪŋ si:n!	ænd monaɹks tə bihə:ld ðə swɛɪn se:n!

It is an interesting accent. It contains features that are recognisable to anyone familiar with modern English accents, yet it is unlike any of them because no modern English accent, for example, says [ɪnven'sjɔn] for *invention*.

I see OP as part of an evolving applied historical linguistics. This was not the first time that people have tried to reconstruct early accents, of course, as part of historical phonology. There have been audio recordings of Chaucer, *Beowulf*, and other texts. But this was different, as the reconstruction was not an end in itself but a means to a further end, the production of a play which would work successfully on a stage. I therefore had to do things that, as a 'pure' historical phonologist, I would never have had to do, such as training the actors to use the accent well. It is exactly parallel to what goes on in another area of applied linguistics, ELT. We are now talking of applying a new branch of linguistics to language performance in an artistic and dramaturgical sense.

I have given you two examples of novel domains of linguistic application – domains that were unpredictable a decade or so ago, but which are now being increasingly encountered in various parts of the world. So I wonder what is going to come next. What will the next ten years bring? This is the fascination, as it has always been, in linguistics. *We ain't seen nothin' yet*, and tomorrow is always another day.

References

- BROWN, Keith / LAW, Vivien (eds.) (2002): *Linguistics in Britain: personal histories*. Publications of the Philological Society, 36. Oxford and Boston: Blackwell.
- CRYSTAL, David (2005): *Pronouncing Shakespeare*. The Globe Experiment. Cambridge: Cambridge University Press.

¹⁵ A longer audio version of this prologue is available at www.youtube.com/watch?v=gPlpphT7n9s. Other examples of OP can be heard at www.pronouncingshakespeare.com/op-recordings/.

- CRYSTAL, David (2005): Pronouncing Shakespeare. (online: www.pronouncing-shakespeare.com).
- IBM GLOBAL TECHNOLOGY SERVICES (2006): The toxic terabyte. How data-dumping threatens business efficiency. (online: www-35.ibm.com/services/no/cio/leverage/levinfo_wp_gts_thetoxic.pdf).
- MEIER, Paul (2011): A Midsummer Night's Dream: an Original Pronunciation Production. In: Dudley KNIGHT (ed.): A World of Voice: Voice and Speech Across Culture. San Diego: University Readers. 209–224.
- MOORE, Matthew: Stephen Fry posts Twitter updates while trapped in a lift. In: The Telegraph, 5 Feb 2009. (online: www.telegraph.co.uk/news/celebritynews/4518473/Stephen-Fry-posts-Twitter-updates-while-trapped-in-lift.html).
- OPEN UNIVERSITY: Shakespeare. Original pronunciation. (online: www.youtube.com/watch?v=gPlpphT7n9s).
- SEBEOK, Thomas A. (ed.) (1963–76): Current trends in linguistics. The Hague/Paris: Mouton.