

ENTREVISTA CON ... / AN INTERVIEW WITH ...

DR. GIBSON R. FERGUSON

University of Sheffield, UK

by **M^a Carmen Pérez-Llantada Auría**

Universidad de Zaragoza

Dr. Gibson R. Ferguson has a PhD from the University of Edinburgh in the area of Language Planning in Education. Following an extended spell at the University of Edinburgh (1982-2000), during which he lectured on the M.Sc. programme in Applied Linguistics, he took up a new post as lecturer in applied linguistics at Sheffield University in September 2000. He has previously taught in Malawi and Zambia, and has undertaken consultancies on behalf of the British Council and other agencies in a wide range of countries: Hungary, Mozambique, Cuba, for example. From 2001 he has been the course director of the MA in Applied Linguistics in the Department of English Language and Linguistics of Sheffield. He has carried out extensive research in a number of fields of applied linguistics: curriculum design, pedagogical grammar, ESP, language testing and language teacher education.

In January 2003, Dr. Ferguson received the recognition of 'researcher of excellence' from the Universidad de Zaragoza (Spain) and visited this university to assess research on corpus linguistics led by Dr. Pérez-Llantada (project funded by the Vicerrectorado de Investigación de la Universidad de Zaragoza no. ICE, 245-86). The interview was held at the Faculty of Science the day before Dr. Ferguson led the seminar "An overview of corpus studies. Implications for research and applications for LSP teaching" for a group of teachers and researchers in Applied Linguistics from the Department of English at the Universidad de Zaragoza.

Carmen Pérez-Llantada (CP): How do you define yourself (as a linguist, a sociolinguist, a corpus researcher...?)

Gibson Ferguson (GF): I think I would define myself above all as an applied linguist, which I see is different from being a linguist, and different from being

sociolinguist. Applied linguists are interested not only in language, but also in language learning and language teaching.

CP: **Corpus studies or corpus linguistics?**

GF: I think both; but they are interpreted in different ways. Corpus studies are specific studies of particular words or particular grammatical features. Corpus linguistics, on the other hand, is the whole field. Within corpus linguistics you have different specific studies.

CP: **What do you understand by corpus linguistics?**

GF: Corpus linguistics, above all, is a methodology, a methodology for studying language. I don't think it is a different kind of linguistics, just a different methodology.

CP: **Is corpus linguistics a theory, a discipline, a methodology, none of these or something else?**

GF: It's above all a methodology. A way of doing linguistics, which is different from that followed during the previous 30 or 40 years. It's a new approach, maybe only 10 to 15 years old. A different way of studying linguistics.

CP: **What can you tell me about your own beginnings in corpus analysis? How did you get involved in it? Why did you get involved in it?**

GF: I first started to read some articles and papers where people were using corpus techniques to investigate features of scientific English, and then I decided, well I'm interested in medical English, so I need to use some of these techniques for my own research. So then I started to read about it, and then I applied it in my research in the field of medical English.

CP: **Why do you think corpus linguistics is important for undergraduate students, postgraduate students, ESP teachers and researchers?**

GF: It's particularly important for ESP teachers and for researchers, because corpus studies are very useful for studying language variation, and in ESP we're interested in differences of language; what distinguishes general everyday language, from the language of computer science; engineering language from history and so on. Corpus studies allow you very easily to investigate this variation, these differences, and that makes it a very important

tool for ESP teachers and researchers who are interested in language variation, language change and ESP teaching.

CP: How can corpus linguistics improve research in disciplines like sociolinguistics, genre studies, pragmatics, critical discourse analysis, translation studies, cultural studies, contrastive rhetorical studies, cognition theory, ...? Is corpus linguistics a good tool for all of them, for some of them, is it more consolidated as a methodology in some of these disciplines than in the others?

GF: I think that corpus linguistics is more suitable for certain topics than for others. It is not very easy to use corpus techniques in sociolinguistics or pragmatics. It can be done, but there are some obstacles. On the other hand, it is very useful for studying genre, and register, and also lexicography, grammar, and semantics. All these are very good areas for applying corpus techniques. As I said, corpus techniques are not so useful, or perhaps more difficult to apply in pragmatics, and in sociolinguistics. It can be done, but these are more difficult areas.

CP: Are you going to use any corpus in your next book?

GF: The book that I'm writing at the moment is in the area of language ideology and language politics. Corpora are not such a useful tool in this area.

CP: Do you consider corpus analysis as an overlapping model of analysis with these disciplines or as a complementary or converging model of analysis?

GF: Corpus analysis techniques should always be considered complementary to other methods of linguistic study. In applied linguistics we use a lot of interview data, elicited data, questionnaire data, as well as corpus data. It is just another kind of data which is complementary to interview data, observation data. It's a good complement, but cannot substitute for the others. They go together in partnership.

CP: Bhatia (*Ibérica*, 2002) describes three levels of analysis: textual, genre, and sociocultural, How can corpus linguistics operate in these three levels?

GF: I think corpora are very suitable for studying genre and register, because here we're talking about lexico-grammatical differences, different text types. This is a very suitable area for quantitative work, for counting the frequency of

different words, the frequency of different patterns, grammatical structures, discourse markers and so on. So the textual and the genre levels are both very appropriate for using corpus techniques, which place emphasis on looking at the distribution of features and at their frequency.

However, I think the social-cultural level is a bit more difficult because the ideology or the cultural implications of a text is often not on the surface, and if it's not on the surface, it cannot be seen by the computer. It can only be seen by the human being who looks deep into the underlying ideology or the culture of the text. This needs much more qualitative human interpretation. However, corpus techniques can help by identifying cultural bias: for example, by looking at the associations of words - which adjectives go with certain nouns- you could detect some cultural or ideological bias. So I would say corpora are one of the tools for the cultural study of language but not the most important tool.

CP: **Is corpus analysis a consolidating trend in the field of linguistics or just a trend borrowed from academic research in hard sciences (I mean, providing evidence by means of specimens, patients, devices, experiments)?**

GF: I think linguistics has been dominated in the last forty years by a particular tradition of rationalism led by Chomsky, one that's very much based on individual intuition and introspection. Now corpus linguistics has introduced something of a revolution into linguistics, so that it is now becoming a much more empirical discipline. By '*empirical*' I mean that linguists are currently more interested in observing language in natural situations and in different varieties of text. This has had a considerable impact in linguistics, which has thus become more similar in its epistemology to other observation-based disciplines.

CP: **Where does the relevance of corpus analysis lie in pedagogical contexts? Do you have any experiences you can tell us about?**

GF: I think it is an extremely useful tool for teachers and learners, and also for translators. It can be used in two different ways in language learning and teaching. If you have a very rich university with very good computer laboratories, the student can go to the computer and access the corpus data directly. But if- for reasons of lack of resources - that's not possible, the teacher can print out concordances from the computer, and get students to analyse the concordances in class. For example, to explore different meanings of words, the

grammatical patterns which different words have, to investigate the collocations of different words, to distinguish between similar words like *'begin'* and *'start'*. Corpora and concordances are also useful for translators. For instance, translators often want to translate a word exactly, let's say from Spanish into English, but the correspondence is not perfect because it has different connotations in one language from another. To give you an example, in English the word *'to incite'* (*'incite something'*) has a negative connotation. *'Incite something'* always means something bad. In Italian there is an apparently equivalent word *'incitare'*. Crucially, however, in Italian it does not have the same negative connotations, and this can be revealed by studying many corpus examples. This may mean, in turn, that translators should not necessarily translate *'incite'* as *'incitare'*. So corpora are very useful for learning, teaching, translating.

CP: **Is the use of corpora useful for rich instruction approaches in ESP teaching, for bringing authentic texts to the classroom, for teaching disciplinary discourses or specialised registers in general?**

GF: Yes, I think it is very helpful, perhaps more important than in General English. As I have said, corpora have great potential because in LSP it is very important to know how engineering language is different from the language of chemistry or the language of sociology: and with very specialised corpora we can investigate different meanings of words in these different academic subjects, and how they are used. Corpora also make possible quantitative studies of the distribution and frequency of language features. So I think LSP researchers and teachers will gain many advantages from using corpora - as a tool for their research but also for their teaching - for example, to give a lot of authentic examples to their students.

CP: **What are the advantages or disadvantages of using of corpora for writing, speaking, listening or reading? Is it better for one particular language skill? How about integrating skills in the classroom using corpora?**

GF: I think corpora can be useful in all areas. It fits well with current trends in teaching methodology because it has the potential to give more independence to students to investigate things for themselves. Perhaps the most appropriate use of corpora in classroom teaching is for teaching writing, grammar, and vocabulary – less so for listening. In these areas I have mentioned, corpora can support a kind of discovery-based methodology, which is very popular

nowadays. In other words, rather than being told everything by the teacher, corpora give students potential autonomy for discovering patterns for themselves and I think that's very helpful.

CP: How do you manage to get your students work or get involved with corpora?

GF: I give them small tasks. I call them mini-research tasks. I tell them, ok, I want you to investigate the adjective '*discreet*' and compare it with the adverb '*discreetly*'. Please go to the computer and check the differences between these two. So with mini-projects, the teacher can say to the students: "Ok, I want you to investigate this feature or this topic using your corpora and your software. Next week give a report about what you have found." Then they give a presentation about what they have found and you can assess how well they are performing the research.

As for statistics, I think it would be very useful for students to learn how to use some descriptive statistics, and also to pay attention to sampling considerations. But one very important thing is that you have to train them not only how to look in the corpus, but how to interpret. That's very, very important - the most important skill, how to interpret what the data reveals.

CP: Do you think it would be useful for engineering or science students as well?

GF: Yes, I think so because using corpora is useful not only for language skills, but for general cognitive skills of inference and interpretation, and these are transferable. But I also think students need to be given some guidelines on how to interpret. They have to be a bit familiar with organizing data. The only problem may be that using corpora directly with language students is more suitable for intermediate to advanced levels, it is not such a suitable tool for elementary or beginning level students.

CP: Which is the greatest advantage of using corpora?

GF: The greatest advantage in my opinion is that you can discover things about language which will surprise you, even if you have been speaking that language the whole of your life; using corpora you can discover things that you didn't know.

CP: Any favourite corpora?

GF: I have no particular favourite. There are different corpora suitable for different purposes; there are specialised corpora for studying medical discourse, and general corpora for studying general aspects of language, and then there are corpora for studying spoken language. Each corpus has its own purpose.

CP: What disadvantages, or rather difficulties one can come across when using corpora for research or for pedagogical purposes? Any advice from your own experience?

GF: I think whenever there is something new in applied linguistics and language teaching you can get very enthusiastic. Of course, that's good, but when you are very enthusiastic there is always a danger of exaggerating something as more important than it is. Some people in language teaching have fallen into this trap in the past with new methodology or new technology. So you have to be careful. Corpora have some limitations or restrictions. You cannot investigate everything using corpora. For instance, you cannot very easily investigate sociolinguistics or ideology with corpora. This is because corpora take the language out of the context. You see the words on the screen but you don't know who was the speaker, how old are they, who they are speaking to, what was the place where they were speaking. What you see is language which is out of context, and this is a limitation with some corpora. Another important point is that you need to take corpora and complement it with your own intelligence and your own interpretation. So it is an instrument, it is a tool. We can compare a corpus to a microscope in a scientific laboratory. It's a good analogy, because before the discovery of the microscope we couldn't look at very small things. Now that we have the microscope, which is just a tool, it allows us to look at very small things. A corpus and the software is the equivalent to the microscope. They allow us to look at small things in language, but, of course, the things you see, you have to interpret. The tools cannot interpret, so you need to have a good linguistic framework, and good interpretive skills to interpret what you see. The microscope was discovered in about 1750 by Leuwenhoek in Holland, and now we have computers and corpora. It's the twentieth century version of the microscope.

CP: Key readings for getting involved in corpus studies?

GF: Yes, there are some very good books: there is a very recent one from 2002 – *Introduction to Corpus Linguistics*, by Meyer, Charles Meyer, and there is a very

useful book by Biber and Conrad (1988), *Corpus Linguistics: investigating language structure and language use*. These are useful books to start with.

CP: What are Key tools for working with corpora?

GF: First, you need a corpus. You can either make your own corpus. In this case, it is better to construct a small, highly specialized corpus. Or, if you can't afford to construct one, you can obtain one from the Internet or by buying a CD-ROM, or by subscribing. It depends on the expense. Once you have the corpora, you need the software. There is a very popular software package called *Wordsmith Tools*, the successor to *MicroConcord*. However, some corpora have their own specialised software: for example, *MICASE* has its own software. The BNC, the British National Corpus, has software called *SARA*, which is very good. I think you'd be advised to start with *Wordsmith Tools* or *MicroConcord*.

CP: In your opinion, what are the future prospects of corpus studies?

GF: I think there is still a lot of work we are going to see in the field of grammatical studies with corpora. We are also going to find that corpora are going to be tagged, grammatically tagged, in a much more sophisticated way in the future, and they are going to have parsing programs. So, corpora will be more sophisticated in the future and probably the software will develop. So we'll have more powerful tools. I think this will happen quite soon in the next five to ten years.

CP: Can you describe your latest book *Language in Language Teacher Education*?

GF: This book, published by John Benjamins, is mainly for and about people who train teachers to teach languages. The intended audience is applied linguists and language teacher trainers. Its main focus is on concepts of language, social concepts, ideological concepts, semantic concepts of language, and also on methods or techniques of raising teachers' awareness of language as a phenomenon. It contains a number of papers on these topics.

CP: Can you give us details about your forthcoming book on sociolinguistics?

GF: This book will, I hope, be completed in about one year's time. It will be called *Language Planning in Education*. Language planning is mainly about language policy and language planning, and language education in different countries. Language planning lies at interface between education, policy and politics, and language

ideology. The book discusses such issues as code-switching, world Englishes, language imperialism, language death, and language revival. The latter is an area where there is a lot of discussion in Spain: for example, in relation to the different languages of the country. In general, then, the book is about multilingualism and the implications of multilingualism in today's world.

CP: Well, Dr. Ferguson, thank you very much for everything.