

Interview with
Peter Roach

RG

*Your name is connected particularly with the worldwide popular textbook *English Phonetics and Phonology* which has already had four editions and the first one was published 30 years ago. Did you expect such a success?*

PR

It was a complete surprise. I wrote the original material when I was teaching as a visiting professor in the University of Seville (Spain) in 1975-6, and my Head of Department there asked me if I would write up my lecture notes and tape exercises so that they could continue to run my course after I left to return to my post at the University of Reading. Before I went there, students of English in Seville were expected to read Gimson's *Introduction to the Pronunciation of English* as their introduction to phonetics. It's an excellent book, but it isn't an easy way into the subject, so it seemed a good idea to have something which was simpler and also structured as a course-book. In the end, this project turned into a book, and I needed to find a publisher for it. Cambridge University Press took it on, and they have been very good to work with. I think the book has sold about 300,000 copies so far, including a lot in low-cost editions aimed at third-world countries. Its latest version is the recently-published "enhanced ebook", which can be used on the Apple iPad, iPhone and iPod (it is being sold through the Apple iBook store). This ebook is the same text as the printed book, but all the examples are printed in a different colour, and when you tap them you hear my voice pronouncing them. I am very excited about this development, which I have been pressing for many years.

RG

In one interview you have mentioned that you fell in love with phonetics during your teacher's training at Manchester University. Why phonetics? Why sounds of speech? It is well-known that language teachers tend to prefer grammar or vocabulary building to pronunciation.

PR

Well, when I went to Manchester I had already worked briefly as a teacher of English pronunciation to foreigners and then been through a degree course at Oxford University which laid great emphasis on language, both in the study of Classical languages and, later, in the fields of psycholinguistics and philosophy. But phonetics was something else. I think the nearest comparison would be with music. Imagine you had a hidden talent for music but had never had any experience of performing or studying it, then one day someone discovers your talent and you realize that this is what you want to spend the rest of your life doing. Some people find they are just naturally gifted at music, and it's not a matter of intelligence or force of character - it's just something that a few people happen to have. I was like that with phonetics, though of course it's not nearly as wonderful as music. I was lucky to have great teachers - Alan Cruttenden at Manchester, A.C.Gimson, J.D. O'Connor, Gordon Arnold and John Wells at UCL- and they encouraged me to make a career in the subject. Initially I was

particularly interested in practical phonetics - ear-training, learning to make difficult sounds, working with foreign learners of English and so on, but when I started my first teaching post, at the University of Reading in 1968, I found that I had a well-equipped laboratory in my department and I could start working on experimental phonetics. The training in experimental methods that I had been given when I studied psychology at Oxford was very useful for getting started on this, and I had been taught some experimental and acoustic phonetics at University College London by distinguished professors such as Dennis Fry and Adrian Fourcin. I soon realized that computers were going to be important in this area of phonetics, and took my first course in programming in 1969, when computers were very different from today's little wonders. I loved experimental phonetics and the study of speech with laboratory instruments - my colleagues used to refer to the laboratory equipment as "Peter Roach's toys". The phonetician I admired most throughout my career was Peter Ladefoged, who was brilliant both as a teacher and as a researcher, and who combined traditional phonetic research with advanced experimental work.

By 1971 I had persuaded the University to buy a minicomputer for the Phonetics Laboratory and I spent a large amount of time writing programs for applications such as speech synthesis, reaction time measurement, electropalatography and pitch analysis. So from that time onwards I knew that I was working in a field that interested me more than anything else, both in the teaching I did and also in my research. It was incredibly lucky for me that I found my way into phonetics at a time when British universities had lots of money and well-stocked libraries, and one was able to do research free from the pressure we have today to produce lots of publications and "measurable outcomes".

RG

You have been the senior editor of four editions of the Cambridge English Pronouncing Dictionary (from the 15th to the 18th edition). Which aspect of an editor's job did you find the most challenging and why?

PR

I and my co-editors had to learn a lot when we started work on the CEPD to produce the 15th Edition. The dictionary used to be produced by a small publisher using type-setting, a slow and laborious process. When Cambridge took on the dictionary it was decided to convert the existing EPD into a computer database which could then be used to produce and check new editions much more quickly and accurately. The printed EPD (14th Edition) was optically scanned and fed automatically into the computer database, which was done with astonishing accuracy considering that the process had to preserve the distinctions between normal text and IPA symbols, in text which included bold and italic symbols. Equally challenging was the addition of American pronunciation, which was not included in previous editions. Later we added the CD-ROM, which I think works nicely though I do regret the fact that it isn't available for Apple computers like the Macintosh. The CD-ROM was later given the "sound search" facility, which I find a very useful tool for checking facts about English phonology, as well as for finding errors and inconsistencies in the transcriptions (the actual database is managed by a very impressive piece of software that is kept in secure conditions by Cambridge University Press and takes a very long time to learn - the general public is not allowed to get access to this). The next challenge will be to produce a version of the dictionary in the form of an app which will work on tablet computers and mobile phones.

What I find particularly difficult is deciding what attitude to take on the question of the standard accent. Previous editors (Jones and Gimson) also worried about whether it was a good idea to use what they called Received Pronunciation as the standard, though it was always hard to see what else they could have done. Now more than ever people feel that the RP accent belongs to older times and past generations, but as always the problem is to find another accent that could be used instead. I have tried to make the model a bit more broadly-based by changing the name to “BBC pronunciation” in my own work, but this is certainly not ideal. If I could change the model accent for my book and for the dictionary I would be more concerned to base it on the speech of young people – there is too much material on English pronunciation which uses the speech of old people like me, even though the likely users of the material will be young people.

RG

Cognitive scientists, following the Chomskian generative approach, claim that all languages are English-like with some differences in sound systems and lexicons. And thus many generative phonological theories are considered to be language-universal just because they are appropriate for English. You have several years experience of teaching English as a foreign language and, moreover, you have been Professor of Psychology and Professor of Phonetics. What, in your view, is the source of this simplification?

PR

I suppose many of us who work in the field of phonetics feel suspicious of some of the theoretical work that goes on in areas of linguistics, particularly phonology. Phonetics is often seen as scientific and fact-based, while generative theorists have always seemed reluctant to expose their claims to empirical testing. In addition, the speed with which one linguistic theory becomes obsolete and is replaced by a newer one seems very strange to phoneticians who are used to a much slower pace of development in their science. It has often been suggested that in the early days of generative theory the framework was so attractive to non-linguists such as mathematicians and psychologists that people with little or no background in the study of languages became involved in its development. This may have been a factor in the early days, but I don't think there is much evidence today of theories of human language that are based only on English. Present-day generative linguists (as well as cognitive psychologists who work in the field of language) would surely reject the claim that they base their theories only on English. They would claim that they are searching for universal truths about human language, and English is just a convenient language to use to exemplify their discoveries. We should remember that Chomsky and Halle wrote the *Sound Pattern of English* as a sequel to Halle's *Sound Pattern of Russian*.

RG

You have done research into synthetic speech. What was the primary motivation for this? Do you think that one day the interaction man – machine will be more frequent than the interaction man – man?

PR

My research was aimed more at developing systems for computer recognition of speech. Although I often used synthetic speech in my research, I found it less interesting. There is a practical reason for this: if you use a computer to speak or read to you, you will find it takes more time than it would take for you to read the same information, so synthetic speech is really only useful if you can't use your eyes (for example, if you are blind, or driving a car, or working in the dark). But if you speak to a computer that can recognize speech, you will always find it much quicker than typing on a keyboard, so speech recognition can really save a lot of time.

I got involved in computer speech technology in the late 1970's and at the time many people thought it was going to be a big success. As the years went by, the widespread use of speech input and output was always being predicted by researchers in the field for five to ten years in the future. Now we are in the 2010's and still the technology has not been widely adopted, though I have seen the price of a speech recognition system drop dramatically - a working small-vocabulary system in the 1980's cost about the same as a new Mercedes, while today speech recognition and synthesis applications are given free with cheap computers. What we didn't realize back in the 1970's was that though most people like using computers and they like speaking, they don't like speaking or listening to computers. Even sophisticated applications like Apple's Siri are only used by a minority of the people who own an iPhone or iPad. Of course, there are applications for disabled users such as Professor Stephen Hawking who can only speak by operating a synthesis system, and some telephone banking systems that use speech recognition for basic interaction with customers, but these do not yet amount to a widespread adoption. Human reluctance to engage in spoken communication with a computer is a real psychological problem that we have still not understood.

RG

You have also tried to find the ways in which the speech recognition software may help improve language teaching. However, these new computer technologies are more used for forensic or various commercial purposes than for teaching. What, in your opinion, may be the reason for this?

PR

Many people (including me) have tried to develop computer-based systems that evaluate a learner's performance and correct their errors. Some have actually been marketed, though I don't think anybody has become rich in this way. It is relatively easy to get a computer to analyze someone's speech and compare it with some "ideal" template. The most difficult thing is to produce useful feedback based on this comparison. Part of the problem comes from inter-speaker variability: we all speak differently, and have our own distinctive voices, and the brain is very efficient at compensating for these differences so that, for example, the vowel /e/ spoken by a small child is heard as the same sound as /e/ pronounced by a large adult male. However, this speaker normalization process is enormously difficult for a computer to carry out, and a lot of speech-improvement systems have suffered from the "His Master's Voice" effect: the system works fine in the laboratory when its designer speaks to it, but when other speakers try to use it, it "falls over". Many design problems arise: for example, how much inaccuracy in pronunciation is the system to tolerate? A human teacher might decide to be kind to an insecure learner and not demand a very high standard, so as not

to discourage them, but to make stricter demands of a confident learner. Such a decision would be almost impossible for a computer to make.

More effort has gone into developing feedback for handicapped people. Mostly this has been based on the acoustic signal, and one such system for helping children with speech problems was the IBM SpeechViewer, which has been used in speech therapy clinics for some time. I was involved in the development of a similar system, called SPECO, as part of a group of researchers, funded by the European Union, based in the Technical University of Budapest (Hungary), KTH Stockholm (Sweden), the University of Maribor (Slovenia) and the University of Reading (UK). I very much enjoyed the work of collecting hours of speech recorded by children and helping to build a system with amusing animated graphics. The SPECO system was made into a commercial product at the end of the project (the product name in English is "Box of Tricks"), though I don't know if it is still available.

A system that used articulatory feedback instead of acoustic analysis was based on the Reading Electropalatograph, a system that detected contact between the tongue and the palate electronically and displayed the activity on a computer screen. Many children with articulation disorders were helped in this way in the Reading University speech therapy clinic.

RG

You like travelling and old cars. How do your hobbies help you in your professional life?

PR

My hobbies helped my professional life by saving me from going crazy! As I got older I had to become more involved in academic management and politics, particularly when I was Head of Department, and this left less time for phonetics. My hobbies all have one thing in common – they get me out in the open air and the countryside, away from phones, computers, committee meetings and publishing deadlines. So as well as driving around in my old car with the roof down, and going on camping trips in our camper-van, I love long-distance running and walking, cycling and sailing.

Thank you for the interview.

Renáta Gregová