O brave new audio world

David Crystal

This year I found myself doing something new in the field of pronunciation, and it has left me reflecting on why something so obvious should still feel so innovative. The third edition of my *Cambridge Encyclopedia of the English Language* (*CEEL*) came out at the end of 2018, and for the first time the publisher proposed and implemented an accompanying audio recording of all the pronunciation features in the book - a reading in original pronunciation of the various stages in the development of the language, from Anglo-Saxon times on, and a rendition of all the vowels, consonants, syllable structures, prosodic features, and paralinguistic features in the phonology section. This was 'speaking out' indeed.

I still can't quite believe that it has taken over 20 years to get to this stage. The first edition of *CEEL* came out in 1997, and its companion volume, *The Cambridge Encyclopedia of Language*, also in its third edition (2010), still has no audio incarnation. I look along my shelves of books on phonetics and phonology, and can't see one that is available as an audio book. In the world of language teaching, an audio dimension to a textbook has been routine for years. Why hasn't the same thing happened in relation to topics of a general descriptive character?

Is the answer to be found in the state of technology? Partly. When *CEEL*'s first edition came out, Google was still two years away, and most people had not yet attempted an email. The term *podcast* didn't exist. However, a CD in the back of a book was an option, or an accompanying cassette tape, and indeed I do have a couple of such books on historical phonology and regional dialectology that did precisely that. But there were problems with that approach which made publishers reluctant to use it: CDs and tapes regularly got damaged or lost, or - worse - were copied without permission.

It wasn't just a technological issue, though. Audio books have been around for ages, especially for the blind - I recorded one myself in the 1990s - and the Audio Publishers Association was formed in 1996. In the EFL world, audio back-up for textbooks has been with us for quite some time. But in phonetics and phonology, the medium has been conspicuous by its absence. Think of all the textbooks in those fields you have read, and ask: is there an audio book? When I was preparing *CEEL3* I looked around for models to help me decide how to proceed, and couldn't find any. Two books with a CD in the back out of a thousand isn't anything to be proud of. Yet surely, if any subject is the number one candidate for an audio presentation, it has to be pronunciation.

Having had my *CEEL3* experience, I am beginning to see why. The technology side is no longer an issue, now that there are digital downloads. And indeed, these solved my problem when it came to my section on regional accents, national and international. Thanks to the book's online presence, I was able to add links to the excellent IDEA (the International Dialects of English Archive) that Paul Meier has been compiling for several years: a simple and elegant solution. And for the other dimensions to the work, CUP is one of several publishers that have accumulated considerable experience in the audio world, so there was plenty of professional advice on offer.

Every double-page spread of *CEEL*3 has a link to some supporting content, and much of this is audio in character.

No, technology is not the barrier; the problem comes with the authors. Can we handle the recording effectively? When reading extracts from historical texts such as *Beowulf, Canterbury Tales*, and Shakespeare, we need to be phonologically accurate, obviously, but we also need to make the reading interesting, entertaining, and plausible, and this means the reader has to be a bit of an actor. The texts have to be respected, and presented with a degree of feeling. So either we need an actor who can do historical phonology (I know of none) or a historical phonologist who can act. The latter is a rare breed, which perhaps is why there are so few examples to date. I'm fortunate, in this respect, having had a close encounter with both worlds in my earliest academic days, when a phonetics training with Gimson and O'Connor coincided with drama opportunities in the UCL English department, and in more recent years working with my son Ben's theatre company at Shakespeare's Globe. But I looked in vain for kindred spirits when I was developing original pronunciation practices at the Globe a few years ago.

A further issue is that, if a book contains texts authored or spoken by both women and men, we need different voices to do them justice. And *CEEL* does have a mix of texts - such as in the Middle English period, where I use a letter from Margaret Paston alongside a prologue by William Caxton. Then there are dialogue extracts where two voices are needed, such as a piece from a Shakespeare play. We also have to consider the desirability of having several speakers for a series of readings, to avoid possible monotony if the same voice is used repeatedly. But costs, as I mention below, can make such diversity prohibitive, let alone the question of whether it is possible to find actor/academics willing to take time out of their busy lives to read texts for a project that is not their own.

The acting side is irrelevant when it comes to the basic topics in phonetics. We do not expect the cardinal vowels to be read with feeling! Or a list of the RP vowel phonemes. Indeed, to put some emotion into them would be a distraction. But the issue is not so clear-cut when it comes to intonation, stress, and other tones of voice. In *CEEL*, as in any standard account of the English phonological system, there are sections that illustrate the main tones and tunes, stress contrasts, assimilations, elisions, rhythmicality, and so forth. Presenting intonational contrasts inevitably involves taking semantic considerations into account and pronouncing them effectively, so that they sound real.

Then there is the question of confidence. We would expect anyone trained in phonetics to be able to manage the basic sound contrasts well enough, but faced with the tension of a recording situation, how many of us would be certain that we had got the articulation of a sound exactly right, as we worked our way through a long sequence? The ideal situation would be one where there was unlimited time to pronounce each sound, so that it could be articulated several times and the speaker could then listen to them all and choose the best exemplar. Or, even more ideal, there would be a back-up phonetician in the studio to provide a second opinion. But the practical realities of audio recording rarely allow this. Studio time is expensive; sound editors are expensive; additional expert listeners are expensive. I am not talking about a few minutes. Several hours of material were involved in the case of *CEEL*3. And

when I recorded my *Oxford Dictionary of Original Shakespearean Pronunciation*, it took a week.

Another factor is voice consistency. A recording situation isn't like a classroom, where the articulatory dynamic is variable and flexible. It doesn't matter if you demonstrate cardinal 3 with a bit of a sore throat, or if *pit* is spoken with a rising inflection and *bit* with a falling one. But there has to be reasonable auditory consistency in a recording. And it isn't easy to maintain a constant voice quality, loudness level, and intonational range for each word in a long series of examples. The voice inevitably tires as the day proceeds. It can be noticeably different even after sipping a glass of water. A good sound technician in the recording studio will notice when the voice is shifting quality - sounding 'dry', for example - and suggest a pause or a drink. Speakers need this help. They are concentrating so much on getting the words and sounds right that they simply don't realise something else is going awry.

There are of course limitations to what can be done, even by the most expert phonetician in the most ideal of circumstances. It isn't possible for someone with normal vocal apparatus to illustrate many kinds of speech disorders, for instance, especially those where the conditions are anatomical (as in cleft palate) or neurophysiological (as in dysarthria). And similarly unperformable are the earliest stages of child language acquisition, such as infant crying, cooing, and babbling. Even imitating regional or foreign accents can be problematic: we have to be careful not to present a pastiche, especially in cases where there is a sensitivity over racial stereotyping. But even an accurate rendition of a regional or social accent is not going to stop someone, in these hyper-sensitive days, claiming that they have been offended by it. In all such cases, links to websites where natural speech can be heard become invaluable.

The process of recording phonetic examples is full of challenges, therefore, but they need to be met. The more I think about it, the more absurd I find the concept of a textbook on general phonetics or on the history of English pronunciation in which the readers are left to imagine the sounds from the descriptions that the writers, doing the best they can, have provided. If audio books continue to be rare, the answer has to be online - and we do increasingly find publications (as in Cambridge Core) or dictionary entries (as in the *OED*) where writers or compilers have provided spoken illustrations. I sense that we are entering a brave new audio world.