

When teaching phonology isn't enough: insights from mondegreens

Helen Fraser

All her life, Sylvia Wright had known a sad Scottish ballad with the lines 'They have slain the Earl of Moray, and Lady Mondegreen'. Her discovery that the real words were 'They have slain the Earl of Moray, and laid him on the green' led her to coin the term 'mondegreen', now used for any humorously mis-heard song lyrics. You probably have your own favourites. If not, google 'mondegreen' and you soon will.

Mondegreens are great fun, but, like many kinds of word play, they also give us useful information about speech and how it works. Pronunciation teachers sometimes use them in class to demonstrate linking, schwa, etc. Here I would like to use mondegreens to demonstrate some aspects of speech that one would not teach directly in pronunciation lessons, but that might be useful as background knowledge for teachers, especially when phonology taught in class does not transfer well to learners' spontaneous speech.

Mondegreens and mishearings

What makes a mondegreen funny? Of course the quirky meaning is important but surely a big part of it is how a phrase that seems perfectly obvious to us is heard by someone else in a completely different, yet strangely justified, way. Consider a famous Australian mondegreen: hearing the national anthem as 'Australians all like ostriches' instead of 'Australians all let us rejoice'. The meaning is crazy – yet when you play the phrase over in your mind, you can kind of see how someone might hear it that way.

Strictly, mondegreens involve songs, but something very similar happens in speech. Mishearings of speech are rarely entertaining enough to go viral – though there was a bit of publicity for the teacher who thought he had been asked to send an 'all star female' rather than an 'all staff email'. But mundane mishearings happen all the time. In fact if we pay close attention to our perception we realise we often change our initial interpretation of what was said. For example, listening to the radio recently, I heard 'I'm not getting Zimbabwe'. I soon realised that didn't make sense, and replayed the phrase in 'echoic memory' (our ability to 'hear it again' for a short time) till I found an interpretation that fitted the context better: 'I'm not yet in Zimbabwe'.

Studying mishearings can teach us a lot about speech (Tang & Nevins, 2013). Here I want to concentrate on how they show us a side of speech that is usually hidden to us, and reveal the unnoticed work our minds must do to interpret it. Understanding how our own minds interpret speech gives useful pointers to what and how to teach second language learners.

The double life of speech

When we speak, we feel we are producing a sequence of distinct words, each separate from the others, and each made up of a sequence of distinct phonemes and syllables. However, what comes out of our mouths is quite different to that. You may have observed yourself, for example, that 'did you' often comes out as 'didja'. But the differences go far deeper than that. Speech really has a hidden life of its own, quite different to how it appears on the surface.

It is hard to recognise the true nature of speech just by listening to it as it passes. A better impression can be gained by recording and transcribing (Shockey 2003). Richard Cauldwell's resources give an excellent opportunity to do this yourself (Cauldwell, 2013). If you don't have the patience to transcribe, a quicker way to gain insight into the nature of speech is to make a recording of conversational speech (radio talk shows make a good source) and use a sound editor (such as Audacity) to mark the boundaries between words and phonemes (without transcribing). It seems that should be an easy task, but I can personally attest that it is not.

When I was a graduate student, I was involved in a large project aiming to train computers to recognise speech (or, as we used to say, 'to wreck a nice beach'). At that time, the method involved identifying acoustic cues the computer could use to detect boundaries between words. I spent many hours glued to a screen in the basement, laboriously trying to 'segment' recorded sentences (i.e. locate the boundaries between words, phonemes and syllables). It is not easy to appreciate how hard this is without trying it yourself. When we listen, the boundaries seem quite distinct, but when you go in and look, phonemes, syllables, even whole words are blurred together with no boundaries to be found (one reason that this method of computer speech recognition was abandoned).

Speech is actually a continuous stream of sound. Indeed the existence of mondegreens depends upon that fact. To see that, consider what happens when we misread print, which actually is a sequence of distinct units.

Reading errors usually involve omitting or transposing units [sic]. Hearing errors, however, can give a completely different phrase, composed of totally different units. Consider the mishearing 'Got an opal candy?' for 'Got a notebook handy?'. We could never have a mishearing like that if speech was a sequence of units.

Listening without the mind

Transcribing gives useful insights, but it is hard to really appreciate the continuous nature of speech unless we can listen to it without knowing in advance what the words are supposed to be. 'Gating' is one way to do that.

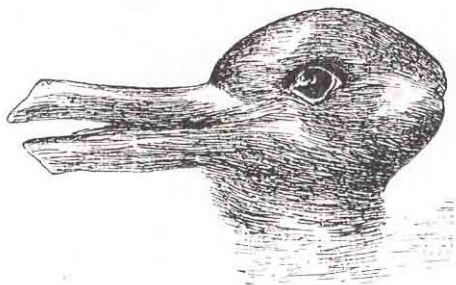
Gating involves playing a short recording of speech, not as we usually play it, continuously from the beginning, but in sections from the middle out. This reduces our ability to guess what the words are going to be, and the experience is quite remarkable. (You can find examples of gating and other intriguing demonstrations at rethinkspeech.com.au/demonstrations). Typically, we hear an uninterpretable noise – with no phonemes, syllables, words or any other recognisable units – until suddenly we hear an entire sequence of meaningful words.

Most people find this surprising. Even if we know in principle that speech is continuous, gating really brings home to us that the units that seem so clear in perception simply do not exist in speech itself. This can give useful empathy for the difficulty second language learners face in recognising and reproducing the sounds of English. But it also raises some interesting questions about our own perception. How do native speakers hear the units of speech if they are not objectively there to be heard? The answer is that the units are supplied by our minds, and again, mondegreens can give us useful insights into how that works.

Conceptualising speech

We see the role of the mind especially clearly in induced mondegreens, where we affect people's hearing by suggesting an interpretation. A classic example is the hilarious video that makes us hear 'All by myself' as 'Obama's elf' (www.youtube.com/watch?v=A_B5UrI7nAI). More disturbing examples come from legal cases, where a transcript can induce a jury to hear a 'confession' the speaker never actually made: (<http://forensictyping.com.au/category/experiments/>).

These things seem surprising, but the mind plays a crucial role in all perception. Consider the famous picture below (http://en.wikipedia.org/wiki/Optical_illusion).



When you first look at it, you see a duck or a rabbit. Either way, you feel you are simply seeing what is there to be

seen. It is only when someone else sees it differently that you realise all that is really there are some lines on a page. Making the lines into a picture takes a contribution from our minds. Fascinatingly, however, though our minds clearly play a hugely active role in perception, we typically disregard our own contribution, believing we are simply picking up information from outside (Frith 2007).

We can call this contribution of our minds *conceptualisation*. Conceptualisation is the process of deciding what something is, and it affects not just perception, but also production. If I flashed the duck/rabbit picture in front of you for a few seconds and asked to you reproduce it, your drawing would be greatly influenced by whether you had conceptualised it as a duck or a rabbit.

All of this has exact parallels in speech. Speech itself is a continuous stream of sound, which we conceptualise as a sequence of meaningful words and phrases, and reproduce on the basis of our conceptualisation – yet we disregard our own contribution, believing we have simply copied words and phrases that are 'out there'.

What about the phonemes, syllables, intonation contours and all the other units, you might be asking? Don't we need to conceptualise those first, before we can conceptualise the words? That is a common misunderstanding, but the truth is the exact opposite, as we can see by exploring gating and mondegreens a little further.

Words first

We saw above that in gating, we hear no units at all until we hear a sequence of words. What happens next is even more interesting. As further gates reveal more speech, the sequence of words we just heard appears to alter radically. The phonemes and syllables apparently shuffle around to make a new phrase – like a series of mondegreens unfolding in time. Of course we know the speech itself didn't change. What changed was our conceptualisation of the speech. And our conceptualisation of the phonemes and syllables follows, rather than preceding, our conceptualisation of the words and phrases.

To confirm this, it is worthwhile to round off a gating experiment by going back, after you know what the whole sentence is, to listen again to the earlier gates. Usually, the phonemes seem far clearer this time, confirming that it is knowing the words that helps us hear phonemes and other units, not (as is usually believed) the other way round.

This is often found to be surprising, but, again, it is quite in line with other aspects of human cognition, as we see from the duck/rabbit picture. It is only after you have conceptualised the whole as a rabbit that the part below can be conceptualised as ears.



Well, this is all very interesting and you may be already thinking about how to use it in class. Before doing that, it is best to pursue another question about ourselves. How is it that native speakers hear phonemes and other units so clearly when they are not really there at all?

From what we know to how we learned it

The ability to conceptualise phonemes, syllables and other units is called phonological awareness (Gillon, 2007). Studying the process whereby native English speakers gain phonological awareness throws up some interesting observations.

Little babies conceptualise speech as a continuous stream of sound. Reproducing this allows them to communicate with us very effectively, and endearingly, with intonation alone. But gaining real linguistic competence requires them to abandon the rambling 'conversations' of pre-language, and learn individual units of language (Berko-Gleason & Ratner, 2012).

This starts with the 'first word', when the child recognises a recurring meaningful unit in the stream of speech, and reproduces it (at least approximately). We call these early utterances 'words', but really they represent a whole phrase or sentence. For example, 'up' might mean 'pick me up'. Gradually children learn to produce sentences in the adult way, till by the age of four, they can talk with near adult competence. But they have little ability to conceptualise sentences as sequences of words. Ask 'how many words' in a favourite rhyme, and they will look at you blankly. It takes experience with the written language to give a clear concept of 'word' in the adult sense.

Once they can identify words, the child's next task is to learn to break words into parts. The first parts they recognise are morphemes – units of meaning. For example, children recognise the /s/ at the end of 'books' as a separate unit before they can identify the /s/ at the end of 'box' (Byrne, 1998). Only later do they learn to identify meaningless units, such as syllables, 'beginnings of words' or 'rhymes'.

Last of all comes 'phonemic awareness', the ability to identify English phonemes. This is extremely hard, only emerging as a result of extensive experience with reading and writing, and even then rarely before the age of ten.

The strange inversion

This brief overview highlights two key features of phonological awareness. First, it is very hard. Second, it progresses in a sequence from larger, meaningful units to smaller, meaningless ones, with phonemic awareness mastered last.

However, as children move from 'learning to read' to 'reading to learn', something strange happens: they forget

the long, arduous process of phonological awareness, and come to feel that recognising phonemes is 'as easy as ABC'. Further, phonemes, the last learned unit, come to seem the most basic, the building block of all other units.

Due to this 'strange inversion', people confidently believe that they recognise words by first recognising phonemes (Port, 2007, gives insightful discussion of just how strange this idea is), and that they recognise phonemes simply because they are 'there' in speech. This is definitely not right, as we have seen in the discussion above, and as has been well established by the disciplines of phonetics, phonology and psycholinguistics.

Unfortunately, however, very few people study these disciplines. For most, the idea that speech is a sequence of words made up of a sequence of simple sounds just seems so obviously true that they never question it. Instead, they build a whole range of other assumptions on top of it. One of these assumptions is that teaching second language pronunciation should be quite easy: all you have to do is demonstrate the sounds of English and give rules for putting them together into words. This belief (inaccurate, as any teacher knows), can make people impatient with foreign accents, for example.

Phonetics phonology and psycholinguistics

It is only when we actually try transcribing speech with the symbols of the International Phonetic Alphabet that we realise the 'phonemic awareness' of literate native speakers is very limited. For this and other reasons, mere 'native speaker competence' is not a sufficient qualification for teaching pronunciation.

These days, many teachers receive extensive education about the sound system of English (Celce-Murcia et al., 2010), learning not just segmental but also suprasegmental aspects, and discovering how combining phonemes and syllables into words and phrases involves complex modifications of the sounds.

However, it is rather rare for teachers to be introduced to the ideas described above. The fact that speech is continuous, and word-level interpretation typically precedes recognition of any other units is certainly taught (Ladefoged & Disner, 2012), but generally in the context of theories explaining how unconscious mental processes convert discrete phonemes and other units into a continuous stream of sound in speech production, and back to discrete units again during speech perception (Byrd & Mintz, 2010).

Such theories are, rightly, considered to be of limited relevance to pronunciation teachers, who must interact with the conscious, rather than unconscious, minds of their students, so teacher education usually concentrates on phonology suitable for passing on to learners. This is certainly useful, even essential, for teaching English.

However, it does have a limitation familiar to many teachers, especially those whose students come from language and literacy backgrounds very different from English: in some cases, mastering phonology in class does not transfer well to learners' spontaneous speech in real conversation.

To handle this well, teachers can benefit from non-technical understanding of ideas like those described above. Although it is initially confronting to discover that units of speech (including units of intonation and rhythm) are products of our minds, with no independent existence in speech itself, this knowledge can make sense of classroom observations, and of research findings (Fraser, 2009). Understanding how to apply the knowledge well requires specialised workshops (Fraser, 2001), but a few suggestions are given here.

What can we do when teaching phonology isn't enough?

Unlike children, adults can learn pronunciation through explicit teaching of phonological units and rules for their combination. Indeed such teaching is usually essential. However, it is often found to be not enough on its own. Learners may be able to state the relevant rules, but be unaware when their own pronunciation breaks the rules. For example, they may be able to mark 'take' as the stressed syllable of 'mistake', but still pronounce it as 'mis-take'.

It is tempting to assume the solution is exercises to help learners better understand the phonology of stress. In some cases this helps, but when it doesn't, it is worth considering the reason for learners' difficulty. Could it be, not that they don't understand the rules of stress placement, but that they simply can't reliably conceptualise the difference between 'mis-take' and 'mis-take'? It can be hard for native speakers to imagine anyone could not notice such an obvious difference – but of course that is because we learned it at a long-forgotten stage of phonological awareness.

Unfortunately, we cannot teach conceptualisation by giving learners information. We can only facilitate learning by providing appropriate experiences and discussion. In doing this, a useful tip is, when learners have a problem with an aspect of phonology, move to larger, not smaller, units of speech. Learners often benefit from exercises involving words and phrases rather than phonemes and syllables.

In discussing these units, remember it is not always obvious to learners which word you are referring to, so write them on the board (in ordinary spelling, to highlight meaning differences). Demonstrate the distinction several times, pointing to each one as you say it, and ask them what they notice as the major difference in the pronunciation. You may be surprised at the answers, but as with mondegreens, surprise can give an interesting window into how other people conceptualise speech in completely different, yet strangely justified, ways to your own. It also gives an excellent starting point to steer learners' conceptualisation in a direction more appropriate to English.

Couper (e.g. 2013) gives experimental evidence of the effectiveness of this kind of approach, and Couper (this volume) goes into more detail regarding classroom practice.

Helen Fraser studied phonetics and linguistics at Macquarie University (Sydney) and the University of Edinburgh, then lectured for many years at the University of New England (Australia). She now works as an independent researcher on cognitive phonetics, and as a consultant on intercultural speaking and listening, and forensic transcription.

Email: helen@helenfraser.com.au

References

- Berko-Gleason, J. & Ratner, N.B.** (2012). *The Development of Language*. Allyn & Bacon.
- Byrd, D. & Mintz, T.H.** (2010). *Discovering Speech, Words and Mind*. Hoboken, NJ: Wiley-Blackwell.
- Cauldwell, R.** 2013. *Phonology for Listening*. Birmingham: Speech in Action.
- Celce-Murcia, M. et al.** (2010). *Teaching Pronunciation: a Course Book and Reference Guide*. Cambridge: Cambridge University Press.
- Couper, G.** (2013). Talking about pronunciation: socially constructed metalanguage. *English Australia Journal*, 29(1), 3–18.
- Fraser, H.** (2009). Pronunciation as categorization: the role of contrast in teaching English /r/ and /l/. In A. Mahboob & C. Lipovsky, eds. *Studies in Applied Linguistics and Language Learning*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 289–306.
- Fraser, H.** (2001). *Teaching Pronunciation: a Handbook for Teachers and Trainers*. Sydney: TAFE NSW Access Division.
- Frith, C.** (2007). *Making up the Mind: How the Brain Creates our Mental World*, Hoboken, NJ: Wiley-Blackwell.
- Gillon, G.** (2007). *Phonological Awareness: from Research to Practice*. New York: Guilford Press.
- Ladefoged, P. & Disner, S.F.** (2012). *Vowels and Consonants*. Hoboken, NJ: Wiley-Blackwell.
- Port, R.F.** (2007). The graphical basis of phones and phonemes. In O.-S. Bohn & M. J. Munro, eds. *Language Experience in Second Language Speech Learning: in Honor of James Emil Flege*. Amsterdam: Benjamins, pp. 349–365.
- Shockey, L.**, (2003). *Sound Patterns of Spoken English*. Oxford: Blackwell.
- Tang, K. & Nevins, A.** (2013). Measuring segmental and lexical trends in a corpus of naturalistic speech misperception. *Proceedings, 43rd meeting of the North East Linguistic Society*.

