Why deaf people can have accents, too

March 22 2017, by Helen L. Bear

Credit: AI-generated image (disclaimer)

Most people have probably encountered someone who appears to use lip-reading to overcome a hearing difficulty. But it is not as simple as that. Speech is "bimodal", in that we use both sounds and facial movements and gestures to communicate, so deaf or seriously hearing-impaired people often use lip-reading or "speech-reading" – watching facial movement, body language and mannerisms – to understand what people are saying to them.

But are these visual cues enough to help deaf or hearing-impaired people learn to speak with a regional accent? The answer is complex and goes right back to when they learn to talk.

People learn to speak at an early age – and people who became deaf after they learned to talk (postlingual) learn to talk differently to those who have been deaf from birth (prelingual). It is people who are born without any hearing who tend to benefit the most from lip-reading – and are often better at doing it although lip-reading takes longer to learn.

How people learn to talk depends on the development age of the individual. Assuming we are talking about children with hearing loss, under the age of two to three, articulation and language knowledge is negligible, so practising sounds to put together into words is the general approach. If hearing is lost after this age and the speaker has a more developed understanding of language and sentences, this knowledge enables more reinforcement learning of prior known sounds into new arrangements.

This is a premise used by Audio-Visual Speech Recognition (AVSR) systems where both the audio and visual information is captured to recognise the spoken words. Where these systems are used in noisy environments the systems depend more on the visual cues of speech. But the visual information only gets us so far. This may be because we are unable to distinguish some visual gestures between different sounds.

This means that good human lip-readers are rare. It is a particularly difficult skill – and variations between speakers, languages, pronunciations, and local grammars make it all the more troublesome. Good lip-readers are often actually speech-reading rather than understanding speech solely from the movement of the lips. Even good lip-readers can fail to understand silent speech recorded on a video rather than from a speaker in front of them.

Babel talk

As infants, we primarily learn to talk by listening, but we are also watching the way adults around us articulate. We do not know how much visual information infants take in, but we do know that children as young as [six months old](#) can tell when someone begins to talk in a different language. So while an infant cannot yet articulate, they do respond to new accents and pronunciations.

The sounds of speech are known as "phonemes" – and are the smallest units of sound a human can utter within the context of language. Those who can hear learn to talk by mimicking articulation – so if parents use phonemes in a certain way to make the particular words, then the effect of this is perceived as an accent which is then mimicked by their children.

So, given that the way we use phonemes when speaking affects the way we pronounce words, can we assume that with different sounds, we also make different visual cues with our lips? Indeed, if you can't hear the different phonemes that cause different accents, how can they be perceived – particularly given that some [visual cues](#) appear to be the same for different phonemes?

Read my lips

There is exciting recent work emerging from experiments using a computer to lip-read. Researchers from the University of Oxford and Google DeepMind recently presented an [end-to-end lip-reading system](#) using examples of thousands of speakers with more than a million instances of different words.

They showed that, with enough training, a computer can achieve over

90% accuracy in lip-reading. So, if a machine can do it, there is hope that humans can also be trained to do the same because the experiment demonstrated that there is something in the visual speech information that makes it possible to correctly interpret words.

But there are important qualifications – the system was trained on whole sentences. Consequently, we do not yet know if this ability to distinguish sounds in visual information comes from language structure (the sounds that make up words which make up sentences from grammatical rules) – in other words, it's not certain whether the computer is able to deduce what is being said because it makes sense or from the visual gestures themselves.

And we have to stress just how much data this machine requires in order to achieve this level of accuracy. Most deaf speakers simply do not meet this many people, for that many hours, in order to learn to articulate with a specific accent. So those who have picked up an accent just from looking have learned to do so with less information and less training than the computer. And, in my opinion, this makes them particularly remarkable individuals.

So, while lip-reading probably does influence the accents of deaf speakers, the extent of this remains unknown. But if I were a gambling woman, I would bet that the majority of those who are deaf and talk with an accent are either not completely deaf, or they heard the sounds prior to losing their hearing.

The few exceptions to this – well, they are remarkably intelligent people. We should listen to them a lot more.

This article was originally published on The Conversation. Read the original article.