Learning and adaptation in spoken communication

Matt Davis, MRC Cognition and Brain Sciences Unit, Cambridge, UK.

Spoken communication requires continual adjustments on the part of the speaker and listener to ensure optimal transmission of information. For instance, a speaker's production of speech must adapt to the needs of the communicative situation, to the properties of the communication channel, to the information requested and to the cognitive capacities of the receiver. We would speak very differently to a young child, or to a colleague or if we were in a noisy pub.

Similarly the perceptual system must adjust to the incoming speech to optimally receive the intended message. Adaptive processes in comprehension are most clearly apparent when we encounter speech produced in a foreign accent or an unfamiliar dialect, or when speech is badly distorted. In such situations, comprehension improves over the first few minutes of exposure to degraded speech, suggesting a fast-acting adaptation process. Even in extreme experimental situations, perceptual-learning appears to be rapid and automatic. However, for face-to-face dialogue between speakers from the same language community, adaptive processes may be so fast acting as to be undetectable by all but the most sophisticated behavioural measurements.

Current computational models of human speech recognition provide learning algorithms to enable the recognition system to learn from experience. These mechanisms are primarily assigned a developmental role – for instance, allowing the network to acquire vocabulary over a long time-course. Many training exposures are required to alter the computational properties of the recognition system. This slow learning is in stark contrast to the rapid, perceptual-learning processes that we have described. A greater understanding of these learning mechanisms will therefore help us to characterise a crucial component of human language comprehension, and may also assist in the construction of speaker-independent machine recognition systems.

One characteristic of perceptual learning processes that are observed in natural comprehension is that they involve 'top-down processes'. That is, computational mechanisms which, after words have been recognised, make alterations to lower-level perceptual processes. These top-down learning mechanisms are supported by neuroanatomy since most of the brain areas involved in comprehension are reciprocally connected. However, the poor time-resolution of current neuroimaging techniques (e.g. fMRI) cannot distinguish between neural activity that results from, bottom-up, stimulus driven processing, and from top-down processes that operate after recognition. Although, we can observe the consequences of perceptual learning processes in fMRI studies, it has thus far been difficult to isolate the neural processes that are directly responsible for perceptual learning. New imaging technologies such as MEG will be crucial if we are to fully characterise the neural mechanisms involved in these adaptive processes.

We therefore propose a project to investigate the neural mechanisms involved in learning and adaptation in speech comprehension. We will use a combination of behavioural and multi-modal neurophysiological measurements to characterise this process in humans, and then implement artificial recognition systems that are similarly able to adapt to incoming input. The ultimate goal of these investigations will be to create computational systems that exhibit the same powerful and flexible communicative behaviour as human speakers and listeners.