

RESEARCH DIRECTIONS

Speaking of babes and songbirds

Dr Paola Escudero, of The MARCS Institute, is leading a multinational research team to investigate how human infants, human adults and songbirds crack the variability in the speech signal. The project, which is supported by the Australian Research Council, aims to unlock the secrets of speech comprehension, an important component of human thought.

'It seems miraculous that we understand one another, given the differences in the speech of people we encounter on a daily basis,' says Dr Escudero. 'Women speak differently from men, children from adults, and most people in the street have different accents. Surprisingly, this variability in speech is not a barrier for communication.'

Influential models of early word recognition are predicated on the view that humans learn to adapt to the speech of different talkers through exposure to large amounts of acoustically different instances of the same words. Thus, an infant's failure to cope with speaker and accent variability is explained by their small brain-bank of words. However, more recent evidence challenges this research. The finding that children as young as 7½ months can normalise speech across speakers suggests humans may be born with the ability to encode only relevant aspects of speech, ignoring differences between voices and accents alike.

Dr Escudero's project, a collaboration exploring English and Dutch language and examining human infants and adults and zebra finches, will test two alternative hypotheses. The first is that the same single procedure handles speaker and accent variability, requires learning and occurs at the same time during speech processing. The second is that speaker and accent normalisation are two different procedures. The former occurs early in speech processing, may be an innate property of the human auditory system and shared with other nonhuman animals. Whereas the latter takes place



later and requires the learning of higher-order information such as word meaning. Innovative eye-tracking, electrophysiological and computational technology will be used.

While humans are able to handle the variation in the speech signal, the most sophisticated automatic speech recognition (ASR) systems are not. If different mechanisms are found to underlie the handling of speaker versus accent variation, different computational algorithms could create ASR systems that mimic how humans crack the speech signal. The same could be said for the algorithms used in hearing devices and cochlear implants, which would be more effective if they more closely reproduced human listening.

Project Title: Understanding different speakers vs. different accents: apples and apples or apples and pears?

Funding has been set at: \$460,000

Research Team: Professors Carel ten Cate & Niels Schiller (Leiden Uni.), Richard Aslin (Uni. of Rochester) and Paul Boersma (Uni. of Amsterdam)

Contact Details: paola.escudero@uws.edu.au
URL - <http://marcs.uws.edu.au/>

February 2014