

## **Audio-Visual speech processing**

**Chris Davis**

The University of Melbourne

This tutorial will provide an overview and introduction to the area of Audio-Visual (AV) speech processing.

It will be suggested that understanding the range of AV speech phenomena calls for a multidisciplinary approach; one that provides the basis for more dynamic conceptions of speech properties.

The emphasis of the tutorial will be on showcasing the multimodal characteristics of speech from a variety of perspectives. These will range from consideration of the physical aspects of visible speech (outlining facial development and models of facial structure and motion) to reviews of recent studies of AV speech perception. Mention will also be made of the potential for AV applications in human-machine interaction and in ASR with several different approaches being considered.