

# Face by Association: Spoken Language and Accent Affects Face Categorization and Discrimination

Jeesun Kim & Chris Davis

MARCS Auditory Laboratories, University of Western Sydney  
Australia

{J.Kim;chris.davis}@uws.edu.au

Multimodal perception processing has been demonstrated and quantified by clever experiments revealing how perception based upon information transduced from one sense is affected by that transduced from another. A classic demonstration of such influence is where the visible aspects of articulation affect auditory speech processing. This has most often been demonstrated by showing that seeing the talker can facilitate speech recognition in difficult circumstances (e.g., in noise, Sumbly & Pollack, 1954) or with complex content (Reisberg, McLean & Goldfield, 1987), or can even alter phoneme recognition (McGurk & McDonald (1976).

In the last decade, a range of other auditory-visual perceptual effects have been demonstrated. For instance, Smith, Grabowecky and Suzuki, (2007) have demonstrated that a pure tone in the typical male or female fundamental frequency range can affect judgments of the sex of an androgynous face. Similarly, de Gelder and Vroomen (2000) have shown that the identification of facial emotion can be biased by an accompanying happy or sad tone of voice. These effects indicate that the relationship between the auditory and visual signals can influence processing mechanisms subserving a variety of cognitive judgments.

The question remains, however, as to whether in order to be effective, cross-modal relationships need to be governed by some common physical process or whether mere association between the inputs is enough. That is, arguably, the above mentioned examples of auditory-visual interaction all involve signals that are more or less lawfully related by some underlying physical property (visual and auditory speech by the process of production, fundamental frequency and a person's sex loosely by vocal tract length, etc). In the current study we employed auditory and visual signals that were related purely by association (i.e., speech accent and a person's race/ethnicity).

To do this, we used two head models (a male and a female) and constructed an eleven-step morph of each that ranged from a character judged to be of Asian appearance to one judged to be Caucasian. Then each morph was animated with gender matched English, Japanese and French speech or accented speech. The participant's task was to make a forced-choice decision for every animation as to whether the character was Asian or Caucasian. These videos were either played without sound or with sound.

The results showed an auditory influence on the visual judgment in both identification and discrimination tasks (i.e., when Japanese or English with a Japanese accent speech was presented compared to the no sound condition). These results indicate that simple association between auditory and visual signals is sufficient to engender auditory-visual perceptual effects.

## REFERENCES

- de Gelder, B. & Vroomen, J., 2000. The perception of emotions by ear and eye. *Cognition and Emotion*, 14, 289-311.
- McGurk, H. & MacDonald, J., 1976. Hearing lips and seeing voices. *Nature*, 264, 746-748.
- Reisberg, D, McLean, J. & Goldfield, A., 1987. Easy to Hear to Understand: A Lip-Reading Advantage with Intact Auditory Stimuli. In Dodd, B. and Campbell, R. (eds.) *Hearing by Eye: The Psychology of Lip-Reading* (pp. 97-113). London: Lawrence Erlbaum.
- Smith, M. Grabowecky, S. & Suzuki E., 2007. Auditory-Visual Crossmodal Integration in Perception of Face Gender. *Current Biology*, 17, 1680-685.
- Sumbly, W. H. & Pollack, I., 1954. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212-215.