Oral and extraoral facial movement in audio/visual speech perception

Thomas & Jordan 2004
The Question

• What are the contributions of various types of facial movement to AV speech perception?

• Previously investigated using window displays
  – Disrupt spatial relationships between mouth and nose, cheeks, etc.
  – Attention/gaze patterns changed
  – Not very “ecologically valid”
Conclusions

• Oral movements contain most useful information for AV speech perception

• Facial frame contains some non-vital movement information

• Some dependence on spatial context with other features

• Inversion affects efficacy of frame movements
Methods (Exp. 1)

• Four visual conditions
  – No movement
  – Oral movement
  – Frame movement
  – Normal (whole-face) movement

• Three A/V conditions
  – Visual only
  – Audio/visual matched (congruent)
  – Audio/visual mismatched (incongruent)
Results (Exp. 1)

N=20
Methods (Exp. 2)

• Same as Exp. 1, but with non-moving facial features systematically removed

Frame movement, mouth static

Mouth movement, frame static
Results (Exp. 2)

N=14
Methods and Results (Exp. 3)

• Same as Exp. 2, but with each condition also presented in an inverted position
Critiques

• Stimuli not necessarily more “ecologically valid” than window displays

• “Frame” vs. “mouth” fails to examine what aspects of extraoral movement contribute to improved AV speech perception

• Baseline perception pretty low
Possible Follow-ups

• Examine role of individual extraoral features in AV perception
  – e.g. Are jaw movements more effective than eye movements?

• How do these manipulations affect performance with sentences?
  – Can syntactic/semantic context alter the importance of facial movement cues?
Discussion Questions

• The authors seem quite concerned with presenting natural stimuli. How natural are their stimuli and can you think of more natural ways to answer their questions?

• Was experiment 3 important to determining the effects of inter-feature spatial relationships on AV speech perception? Why (not)?