Changes in audio-visual speech perception during adulthood

Dawn Behne[1], Yue Wang[2], Magnus Alm[1], Ingrid Arntsen[1], Ragnhild Eg[1], Ane Valsø[1]

[1] Psychology Department, Norwegian University of Science and Technology, Trondheim Norway
[2] Department of Linguistics, [2]Simon Fraser University, Vancouver BC, Canada
dawn.behne@svt.ntnu.no

Abstract

Audiovisual speech perception research has shown an increasing use of visual information from infancy to young adulthood. The current study extends these findings by examining audiovisual speech perc

2.2. Participants

Participants were 10 young adults between 19 and 30 years old (mean=23 yrs) and 10 middle-aged adults between 49 and 60 years old (mean=53 yrs). Each group had a balance between male and female participants, all of which had Norwegian as their native language. All participants reported having normal hearing and normal or corrected-to-normal

2.3. Stimuli

The stimuli were developed from consonant-vowel (CV) audiovisual syllables (/pi/, /bi/, /ti/, /di/, /ki/, /gi/, /pa/, /ba/, /ta/, /da/, /ka/, /ga/) recorded from an adult male native speaker of Norwegian using a Sony mini DV video camera and an external Røde NT3 microphone.

Based on these recordings, the incongruent audiovisual CVs presented in Table 1 were prepared with a labial consonant in one modality and a velar in other modality. The consonants were either voiceless (/p/ or /k/) or voiced (/b/ or /g/), although within any given audiovisual stimulus voicing of the two modalities was the same (e.g., audio /b/ with visual /g/). The vowel was either /i/ or /a/ to allow for differing effects of vowel context (e.g., [19]), but was the same across modalities for a given stimulus.

Table 1: *AV stimuli incongru*

Young adults Mid-aged adults

mid-aged adults. This pattern may reflect the additional ca 30 years of experience with integrating AV cues that the mid-aged adults have over the young adults.

3.4. Voiceless and voiced stimuli

Previous research has consistently shown a greater likelihood for AV-fused responses with voiced than voiceless stimuli (e.g., [3]). This is also observed in the current study as is illustrated in Figure 3 [$F(1,18)=39.46$, $p<.001$]. In addition, analyses of A and V responses show greater use of A [$F(1,18)=14,27$, $p<.001$] and V cues [$F(1,18)=8.68$, $p=.009$] for voiceless stimuli than voiced stimuli. That

Young adults **Mid-aged adults**

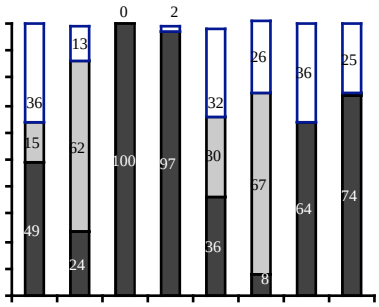

Figure 6: *Percent audio, AV-fusion and video responses in voiceless and voiced* $A_{labial}V_{velar}$ *and* $A_{velar}V_{labial}$ *stimuli by young adults a*

Australian Workshop on Auditory-Visual Speech Processing. 1998.