Effects of auditory, visual, and audio-visual training on nonnative perception of English fricatives

Yue Wang\textsuperscript{a}, Dawn Behne\textsuperscript{b}, Angela Cooper\textsuperscript{a}, Haisheng Jiang\textsuperscript{a}, Nina Leung\textsuperscript{a} and Jung-Yueh Tu\textsuperscript{a}
\textsuperscript{a}Simon Fraser Univ., RCB 9224, 8888 Univ. Dr., Burnaby, BC, Canada V5A 1S6
\textsuperscript{b}Norwegian University of Science and Technology, Psychology Dept, NO 7491 Trondheim, Norway

This study examines the effects of auditory (A), visual (V), and audio-visual (AV) training on nonnative speech perception. Mandarin Chinese natives were trained to perceive English voiceless fricatives (in monosyllabic words and nonwords) of three visually distinct places of articulation: interdentals nonexistent in Mandarin, labiodentals and alveolars common in both languages. Participants were randomly assigned to a control group or one of three 2-week (six sessions, 40 minutes/session) training groups with a different input modality: A, V, or AV. In pre- and post-tests, the fricatives are presented in four ways for an identification task: A-only, V-only, AV congruent (AV\textsubscript{c}), and AV incongruent (AV\textsubscript{i}). Additionally, three generalization posttests are administered testing voiced fricatives, new real words, and a new speaker. Results show that post-training, the trainees reveal: (1) improvements corresponding to training type (e.g., the V-training group improves most for the V-only stimuli), (2) greater improvements for the familiar (but less visually distinct) alveolars than for the new interdentals, (3) decreased AV-fusion for the AV\textsubscript{i} stimuli, and (4) consistent patterns in the generalization tests. Results are discussed in terms of the effects of speech input modality, experience, and L1 on L2 AV speech learning. [Research supported by SSHRC]