INTRODUCTION

Our language experience often involves face-to-face interaction through simultaneous perception of the speaker's voice as well as facial and body movements. While it has been well established that watching a speaker's lips in

Stimuli and Recording

The training word list contained eighty Man	ndarin monosyllabic real	l words (20 tone	quadruplets),	which	was a
combination of the one used in Liu & Samuel ((2004), as well as Wang	et al. (1999)			

It is important to note that the positive role of hand gestures in the current report should be interpreted with caution. First and foremost, the three training groups did not start at a similar level; as the AG group scored much lower in the pretest than the two other groups. Testing the true level of improvement across training modalities requires starting levels of all groups to be similar, and the current results may be skewed due to the outliers in the