ACOUSTICS2008/1607
Speech- synchronized visual cues release speech from informational masking

Mengyuan Wang[a], Jingyu Li[b], Ying Huang[a], Yanhong Wu[a], Xihong Wu[b] and Liang Li[a]

[a]Dept. of Psychology, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China

[b]Dept. of Machine Intelligence, Speech and Hearing Research Center, 2 Science Building, Peking Univ., 5 Yeheyuan Road, Haidian District, 100871 Beijing, China

Visual speech information, such as lipreading cues, can assist listeners to segregate a target voice from competing voices (Helfer and Freyman, 2005). However, because signals contained in lipreading are multidimensional, it is not clear whether a simple visual cue, such as the light flash that is synchronous to the onset of each syllable in target speech, is sufficient to release target speech from noise or speech masking. In this study, when target speech was of a constant rate, the speech-synchronized light flash had no effects on speech recognition under either speech or noise masking condition. However, when the rate of target speech was artificially manipulated unstable or an intense noise burst occurred in the middle of the target sentence, the speech-synchronized light flash improved speech recognition when the two-talker speech masker but not the speech-spectrum noise masker was co-presented. These data suggest that only when the rate of target speech cannot be predicted and the masker is speech, speech-synchronized visual cues play a role in helping listeners attend to the target voice and follow the stream of target speech, leading to a release of target speech from informational masking. Supported by the National Natural Science Foundation of China.