ACOUSTICS2008/1147
Effect of speech rate on speech-on-speech masking

Jing Chen[a], Xihong Wu[a], Xuefei Zou[a], Zhiping Zhang[a], Lijuan Xu[b], Mengyuan Wang[b], Liang Li[b] and Huisheng Chi[a]

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

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

Target speech can be better recognized under speech-on-speech masking conditions if certain differences between target and masker (e.g. in loudness, pitch, location) can be used as cues for streaming. This study examined whether the speech rate can be used by listeners as a cue for unmasking target speech. The rate difference between target and masking speech was manipulated by changing the rate of masking speech using the Synchronized Overlap-Add Fixed Synthesis (SOLAFS) algorithm, and consequently, the ratio of target speech to masking speech (the speech rate ratio, SRR) was quantified. Both target and masker speech were Chinese nonsense sentences and they were co-presented with the signal-to-masker ratio of -7 dB. The results show that speech recognition was significantly increased with the SRR increase from 1 to 1.5 or the SRR decreased from 1 to 0.5. Moreover, the unmasking effect of precedence-induced perceived spatial separation on target-speech recognition was increased monotonically with the increase of the SRR from 0.5 to 1.5. These results suggest that the speech rate is one of the factors influencing both energetic masking and informational masking of Chinese speech. [Supported by the NSFC 60435010; 60535030; 60605016; 30670704]