Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

The Ability to Transiently Store Acoustic Details Is Associated with Speech Recognition against Informational Masking in Reverberant Environments

Huahui Li (1,3), Ying Huang (1,3), Liang Zhang (1,3), Xiangbin Teng (1,3), Xihong Wu (2,3), Liang Li (1,3)

(1) Department of Psychology, Peking University, P. R. China
(2) Department of Machine Intelligence, Peking University, P. R. China
(3) Speech and Hearing Research Center, Key Laboratory on Machine Perception (Ministry of Education), Peking University, P. R. China

PACS: 43.66.Pn, 43.66.Lj, 43.66.Mk, 43.66.Dc, 43.66.Qp

ABSTRACT

Frequency-dependent transient storage of acoustic fine structures is critical for both temporal integration of correlated signals and releasing speech from informational masking under (simulated) reverberant environments. To investigate whether the humans' ability to transiently store acoustic details is generally associated with recognition of speech in noisy, reverberant environments, we invited 30 young-adult listeners to participate in two experiments, one examining their abilities to temporally store fine-structure information of either wideband or narrowband noises and the other one examining their recognition of target speech against speech masking at various simulated reverberant conditions. In Experiment 1, a break in correlation (BIC) between interaurally correlated wideband or narrowband noises, which were presented by headphones, was detectable even when an interaural interval (IAI) was introduced. The longest IAI, at which the BIC in either wideband or narrowband noises was detectable, varied markedly across participants and decreased monotonically as the center frequency increased for narrowband noises. In Experiment 2, target speech was presented by two spatially separated loudspeakers, one simulating the source and the other simulating a reflection. Uncorrelated two-talker speech maskers were also presented by these two loudspeakers. Recognition of target speech markedly improved when the interval between target-speech source and its single-reflection simulation (inter-target interval, ITI) was reduced from 64 to 0 ms. Target-speech intelligibility under simulated reverberant conditions (ITI $\geqq$ 16 ms) significantly correlated with both the longest IAI for detecting the BIC in wideband noise and the longest IAI for detecting BIC in narrowband noises, particularly low-center-frequency narrowband noises. These results indicate that the ability to temporally store acoustic fine structures, particularly low-frequency fine structures, is functionally associated with recognition of target speech under informational-masking conditions in reverberant environments.