Speech in Noise (SpiN) 2019 Ghent abstract

Title: Effects of limited attenuation and signal replay on ideal binary masked speech with very low mixture SNRs

Simone Graetzer and Carl Hopkins

University of Liverpool, Acoustics Research Unit, School of Architecture, Liverpool, L69 7ZN, UK email: s.graetzer@liverpool.ac.uk

This study concerns the effects of limited attenuation and signal replay on intelligibility when speech is mixed with white Gaussian noise at low signal-to-noise ratios (SNRs) and subsequently enhanced with an Ideal Binary Mask (IBM). Such masks require a priori knowledge of both the target signal and the masker. The standard IBM takes values of zero and one, and is derived by comparing the instantaneous or 'local' SNR in each time-frequency bin against a pre-set threshold ('Local Criterion' or LC), e.g., 0 dB. Speech produced by four speakers of British English was mixed with white Gaussian noise at SNRs as low as -25 dB. These signals were subsequently enhanced using IBMs with LC = 0 dB or LC = SNR. The standard IBM was compared with an alternative IBM that took values of 0.2 and one, *i.e.*, using limited signal attenuation. To investigate the effect of replay on the intelligibility of enhanced speech involving very low mixture SNRs, each signal was presented three times consecutively to normal-hearing listeners. The results indicate the importance of mask density for speech signals mixed with white Gaussian noise at low SNRs, where density is measured as the number of ones in the mask. In this study, some masks were sparse due to high Relative Criterion (RC) values, where RC is defined as LC - (global) SNR. There were benefits of limited attenuation at low SNRs for LC = 0 when the masks were sufficiently dense (> 5%), but this did not occur for LC =SNR. For LC = SNR, a second and third presentation resulted in increases in intelligibility scores, whereas for LC = 0, a third presentation was only beneficial when masks were sufficiently dense.