ACOUSTICS2008/3111
Speech perception in noise with binary gains

Deliang Wang[a], Ulrik Kjems[b], Michael Pedersen[b], Jesper Boldt[b] and Thomas Lunner[c]
[a]Ohio State University, Department of Computer Science & Engineering, 2015 Neil Ave., Columbus, OH 43210, USA
[b]Oticon A/S, Kongebakken 9, DK-2765 Smørum, Denmark
[c]Oticon Research Centre Eriksholm, Kongevejen 243, DK-3070 Snekkersten, Denmark

For a given mixture of speech and noise, an ideal binary time-frequency mask is constructed by whether SNR within individual time-frequency units exceeds a local SNR criterion (LC). With linear filters, co-reducing mixture SNR and LC does not alter the ideal binary mask. Taking this manipulation to the limit by setting both mixture SNR and LC to minus infinity produces an output that contains only noise with no target speech at all. This particular output corresponds to turning on or off the filtered noise according to a pattern prescribed by the ideal binary mask. Our study was designed to test on speech intelligibility of noise gated by the ideal binary mask obtained this way. It is observed that listeners achieve nearly perfect speech recognition from gated noise. Only sixteen filter channels and a frame rate of one hundred Hertz are sufficient for high intelligibility. The results show that, despite a dramatic reduction of speech information, a pattern of binary gains provides an adequate basis for speech perception in noise.