An assessment of objective indicators of speech intelligibility in noise at low signal-to-noise ratios

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 paper concerns how well objective indicators of speech intelligibility correlate with the low percentages of words that are correctly identified when speech is embedded in high levels of noise with signal-to-noise ratios (SNRs) down to -50 dB. The indicators under consideration are Short-Time Objective Intelligibility (STOI), the Normalised Covariance Metric (NCM) and the Coherence Speech Intelligibility Index (CSII). STOI is suitable for noisy or degraded speech, including non-linear processed or time-frequency weighted speech. Unlike the NCM, which is based on the normalised covariance between the entire original and degraded envelopes, STOI involves the correlation of the envelopes of clean and degraded (or processed) speech signals that have been divided into overlapping short-time (384 ms) segments. In this paper, speech is degraded by four types of additive noise: white noise, a 400 Hz sine wave, white noise with a 400 Hz sine wave and white noise with a 400 Hz sine wave and harmonics up to 3200 Hz. Listening tests involving normal-hearing human listeners have been carried out for male and female talkers using four SNRs per noise type, ranging from -10 dB to as low as -50 dB. The results characterise the relationship between the objective indicators and the performance-based measure from the subjective tests for speech communication in noisy conditions.

Keywords: speech intelligibility, noise, STOI, NCM