ACOUSTICS2008/216
A computational model of binaural speech intelligibility level difference

Kalle Palomäki[a] and Guy Brown[b]
[a]Adaptive Informatics Research Centre, Helsinki University of Technology, P.O. Box 5400, 02015 Espoo, Finland
[b]University of Sheffield, Dept. of Computer Science, Regent Court, 211 Portobello Street, S1 4DP Sheffield, UK

This study addresses two questions relating to the binaural intelligibility level difference (BILD). First, we ask whether the BILD is underlain by an equalization-cancellation (EC) mechanism, in which a disparity between the interaural time difference of the target and masker is exploited within each frequency channel, rather than across channels. Second, we consider the effects of three sources of internal noise on the EC mechanism: jitter in neural delays, noise in the equalization process and nonlinearities in the auditory pathway. These issues are investigated using a computational model consisting of peripheral auditory model, binaural processor, auditory scene processor and automatic speech recognition system. The binaural model is based on EC processing, with performance limited by internal noise. The auditory scene processor groups speech harmonics by common F0 and identifies 'glimpses' in which the signal-to-noise ratio is favorable for speech. The performance of human listeners and the computational model are compared on the same speech intelligibility test (Edmonds & Culling, 2005, JASA 117(5), 3069-3078). The BILD of human listeners can be replicated by adjusting parameters that determine the internal noise in the EC model; however, the speech reception threshold of the model is lower than that of human listeners.