

# Effective speech intelligibility enhancement system based on Spectral Shaping and Dynamic Range Compression

Tudor-Cătălin Zorilă, Varvara Kandia, Yannis Stylianou, *Senior Member, IEEE*

## Abstract

This paper introduces a non-parametric approach to improve speech intelligibility in noisy. The speech signal is enhanced before presented in the acoustic environment under the constraint of equal global signal power before and after modifications. Two systems are combined in a cascade form to enhance the quality of the signal, first in frequency (referred to as Spectral Shaping, SS) and then in time (referred to as Dynamic Range Compression, DRC). The purpose of SS system is to increase the “crisp” and “clean” quality of the speech signal, and therefore improve the intelligibility of speech even in clear (not-noisy) conditions. DRC combines a downward compression to decrease the loudness of the most sonorant parts of speech with an upward compression to increase the loudness of the less sonorant parts of speech like vocalic onsets and offsets, nasal, stops. To avoid introducing artifacts in the modified signal, Spectral Shaping is adaptive to the probability of voicing while DRC is adaptive to the power of the noise. The combined system, SSDRC, has been applied to natural as well to synthetic speech. Experiments with speech shaped noise and competing speaker noise at various SNRs, indicate that the suggested system outperforms state-of-the-art methods both objectively, in terms of the Speech Intelligibility Index (SII) scores, and subjectively. Extensive formal listening tests show that SSDRC may provide a Signal to Noise Ratio (SNR) gain up to 5dB, especially in low SNR conditions.

## Index Terms

Speech-in-noise enhancement, speech intelligibility, spectral shaping, dynamic range compression.