Workshop on Speech in Noise: Intelligibility and Quality, Amsterdam on the 7th and 8th of January, 2010

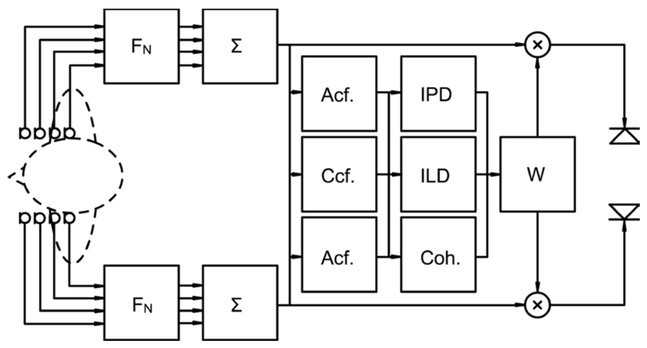
Speech intelligibility assessment in binaural and nonlinear hearing aids

Anton Schlesinger and Marinus M. Boone
Acoustical Imaging and Sound Control, TU Delft

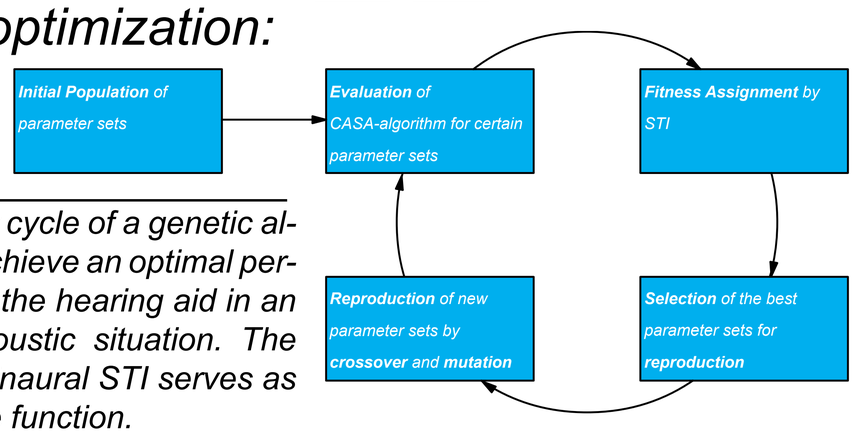
Motivation:

Binaural and nonlinear hearing aid:

The combined processing scheme of a fixed binaural beamformer and a binaural CASA processor. The performance depends on a set of algorithmic parameters that have to be optimized.



And its optimization:

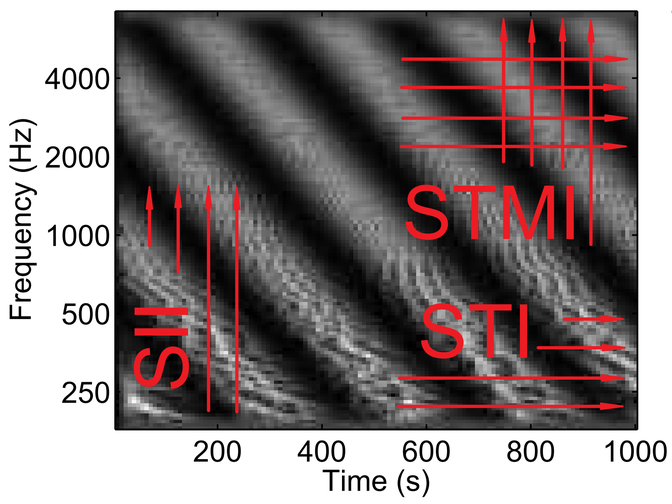


Evolutionary cycle of a genetic algorithm to achieve an optimal performance of the hearing aid in an arbitrary acoustic situation. The proposed binaural STI serves as the objective function.

Problem: the classical speech based STI does not account for the binaural advantage and the nonlinear processing of the hearing aid on speech intelligibility.

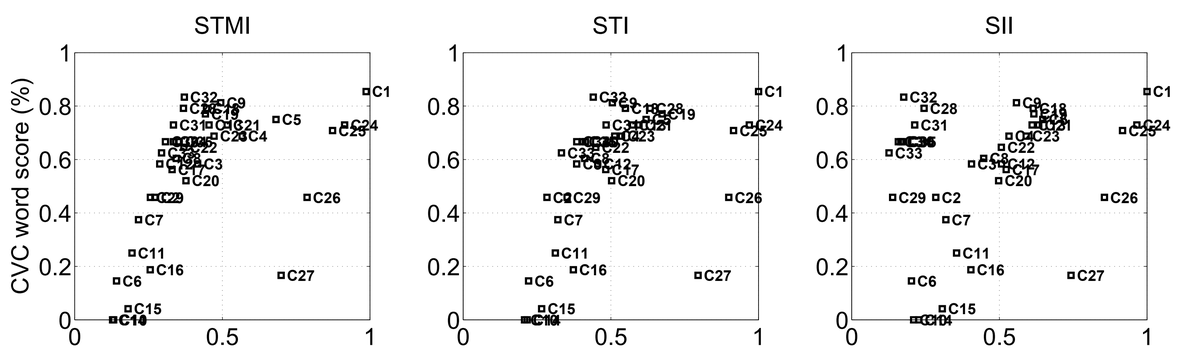
Nonlinearly processed speech:

Three speech based measures of SI:



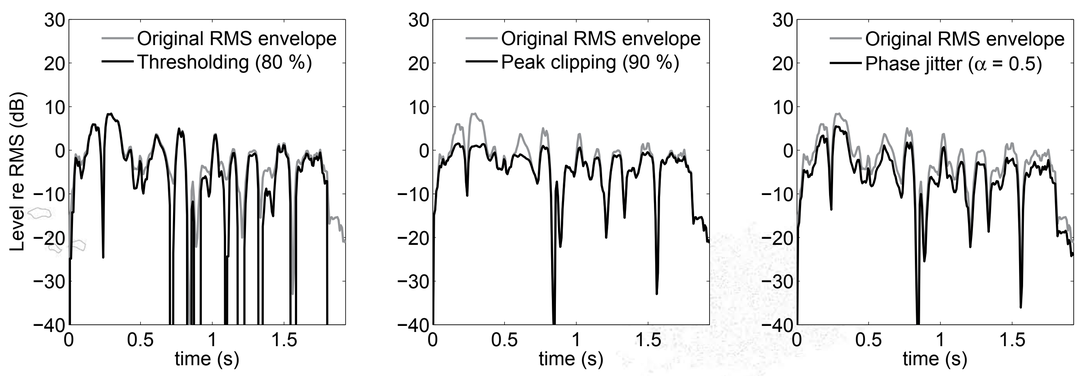
Cochleogram of a rippled envelope and the dominant directions of three speech based measures: the SII (Kates and Arehart, JASA 2005), the STI (Goldsworthy, JASA 2004) and the STMI (Chi et al., JASA 1999).

None of these speech based measures is per se able to map nonlinear conditions correctly.



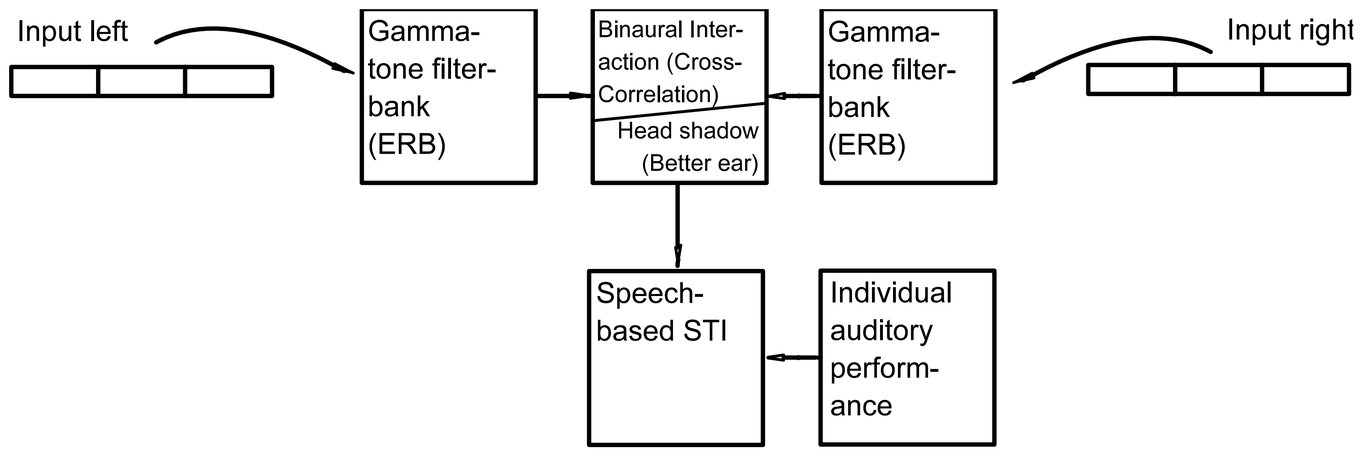
Monaural test of additive noise, reverberation, echo and envelope thresholding (note for the SII, in addition to nonlinear conditions (24-27) the linear reverberant and echo conditions are not correctly assessed).

RMS time course of nonlinearly processed speech

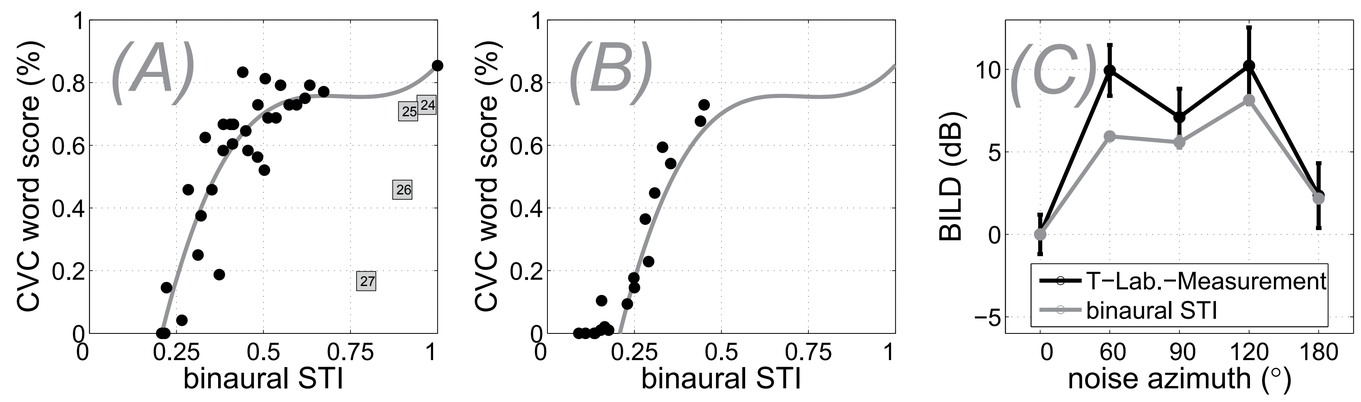


Envelope for a CVC sentence showing the level in dB for three distortions.

A binaural speech based STI:

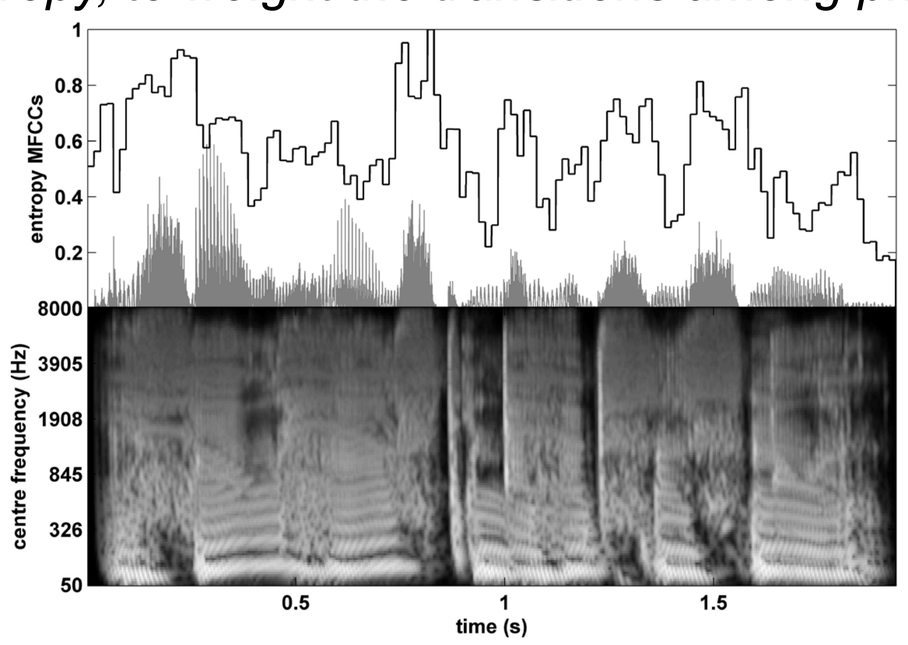


The clean and degraded signals are used in the envelope regression method of the STI. The signals are expanded in a Gammatone filterbank of 30 bands with centre frequencies of 0.1 to 8 kHz. In the range of 0.5 to 2 kHz the binaural interaction process is performed with a cross-correlation stage which refers to the coincidence or correlation model. In the remaining spectrum a 'better ear' approach is used. The threshold of the perceptibility of directional changes is 250 ms. The increment of the time-shift in the cross-correlation stage is 0.1 ms and in the range of -0.8 to 0.8 ms of interaural time differences.



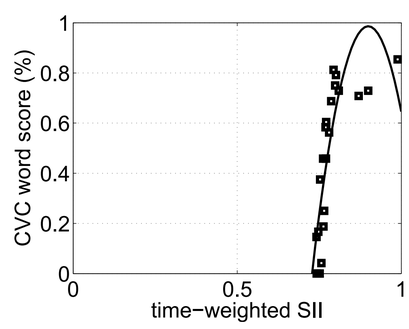
*Qualification of the binaural STI: (**A**, 4 subjects 3 trials) diotic test of 35 conditions. No. 24 to 27 are nonlinear envelope threshold conditions that increase the modulation depth and that are not correctly assessed by the proposed binaural STI. The solid line is a third order polynomial fit through the conditions, excluding the nonlinear samples (standard dev. 12 %). (**B**, 8 subjects 3 trials) binaural conditions at fixed SNRs for N_0 to N_{120} and S_0 and the diotic curve.The standard intelligibility curve was used for the calculation of the BILD at different noise azimuths in an SRT test (**C**, SRT test 8 subjects). Test performed at T-Labs Berlin 2009.*

Application of statistical properties, here Shannon's entropy, to weight the transitions among phonemes:



The upper graph shows the normalized entropy of a MFCC sentence. The positive waveform is plotted (gray) in the background. The lower graph shows the cochleogram of the sentence. - Transitions between phonemes are important for SI and exhibit high entropy.

Preliminary result - a functional relationship for nonlinear conditions.



A contrast enhancement was applied to the entropy course and fitted to the monaural listening test (without reverberation and echoes).

Outlook: combination of the binaural STI and the phoneme course weighting.

TU Delft