Gerhard Schmidt / Tim Haulick
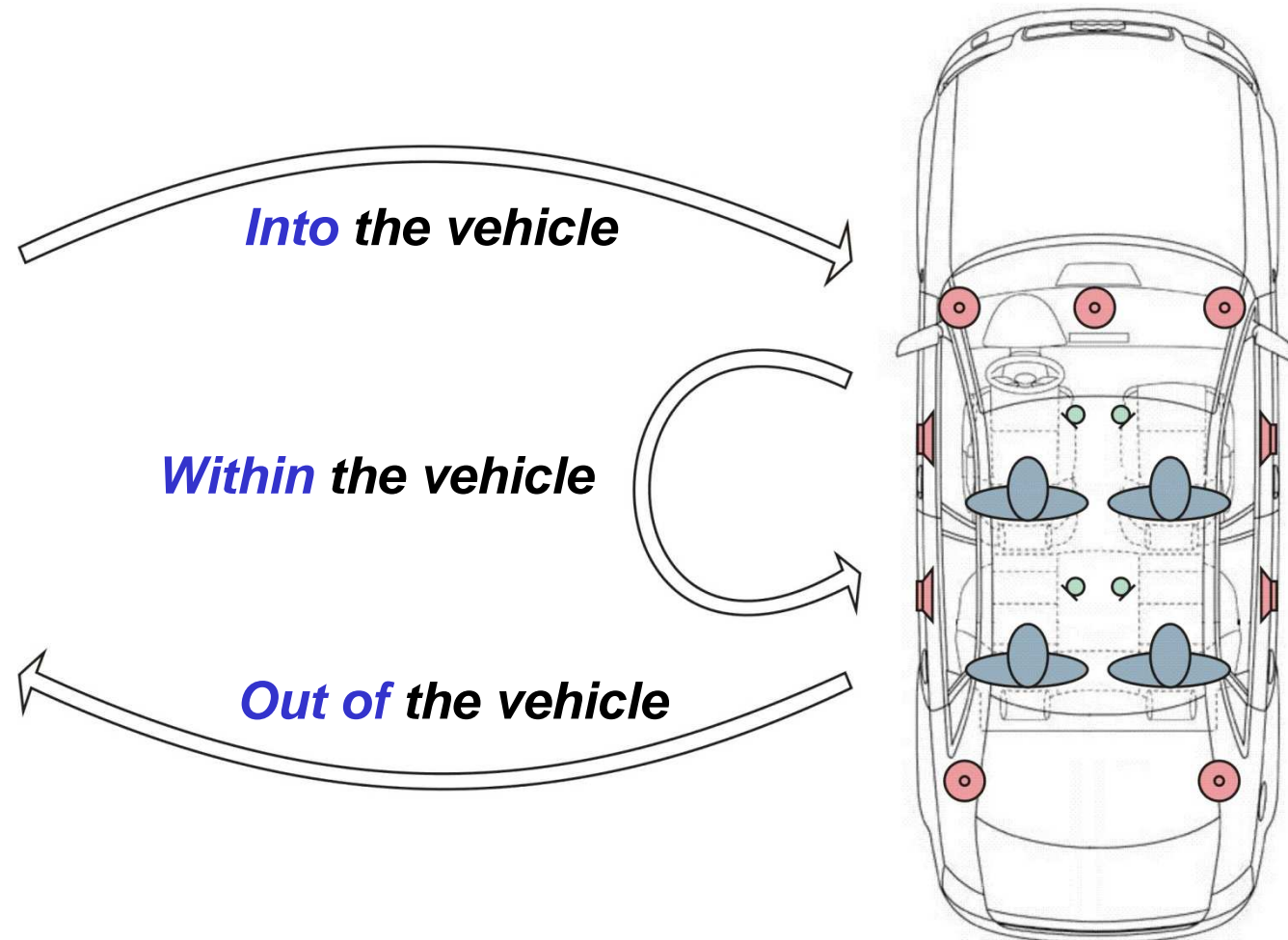
Recent Tends for Improving Automotive Speech Enhancement Systems
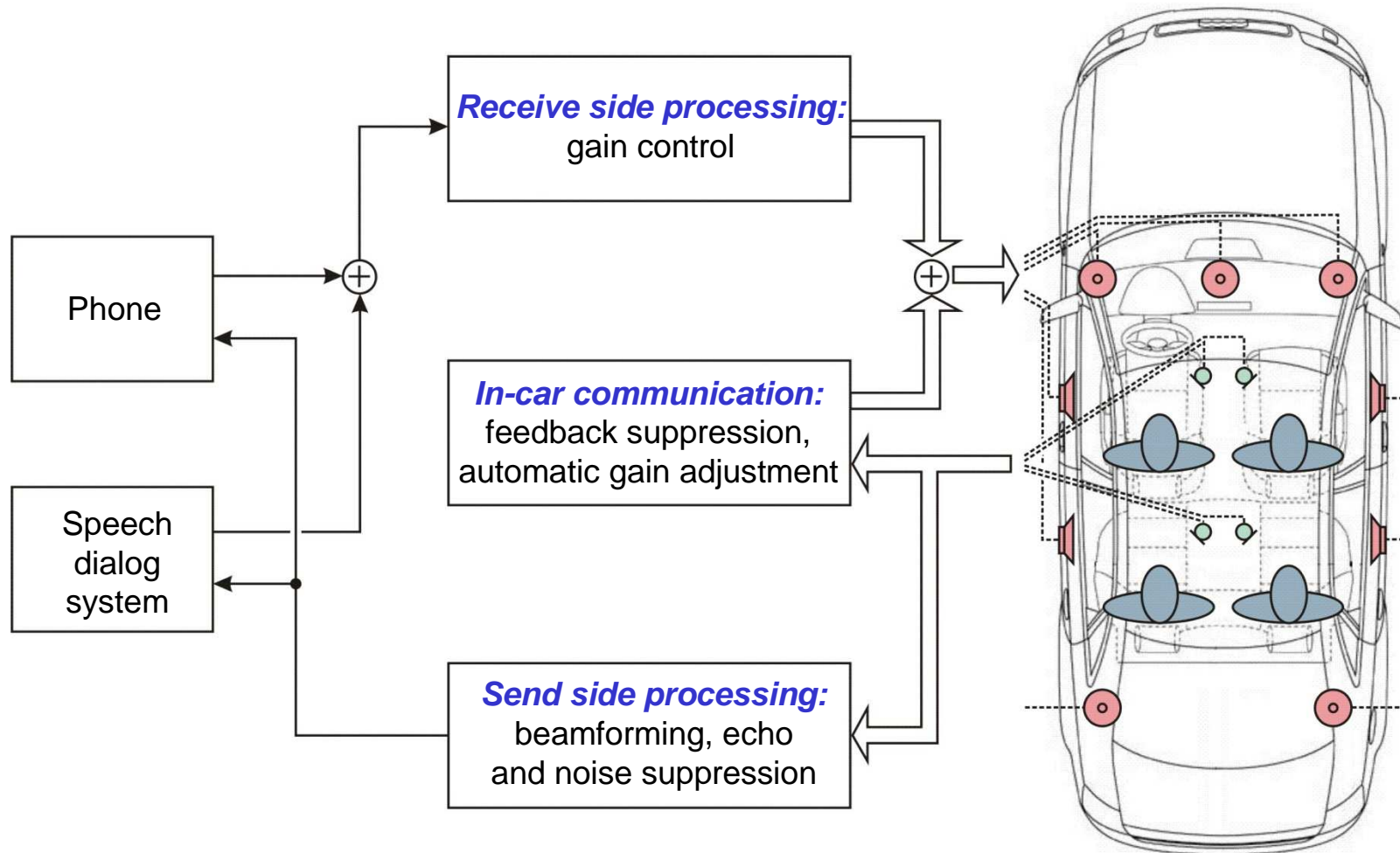
Geneva, 5-7 March 2008

Into *the vehicle*

Within *the vehicle*

Out of *the vehicle*

Receive side processing:
gain control

Phone

In-car communication:
feedback suppression,
automatic gain adjustment

Speech
dialog
system

Send side processing:
beamforming, echo
and noise suppression

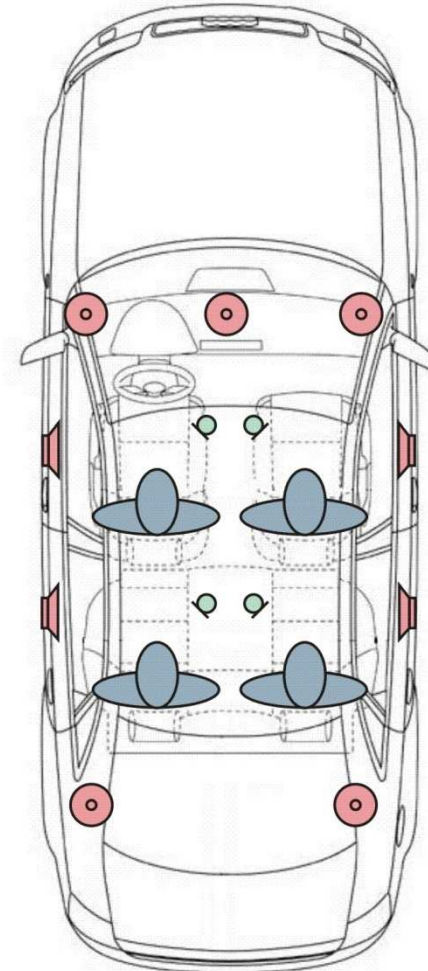*Often the car is owned by the driver or the passenger:*

The communication channels "within the vehicle" and "into the vehicle" should be in focus for improving the perceived system quality.

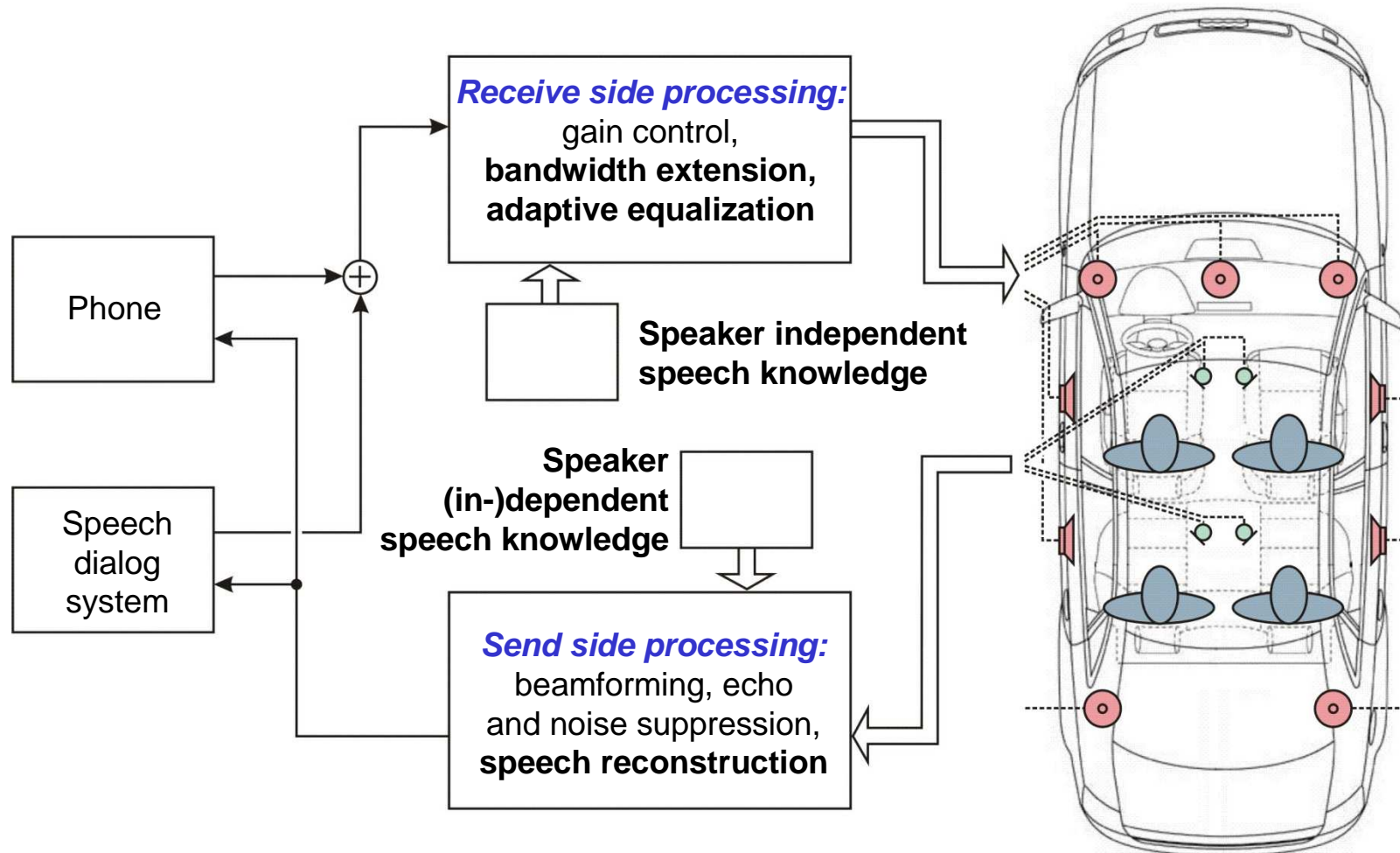*Often a car is used only by a few (3 … 5) people:*

Speaker dependent speech processing might be a good choice.

Often the speech signals of passengers who are using the car can be recorded in periods with high SNR:

Speaker dependant speech models might be extracted in periods of high SNR and used in periods of low SNR.
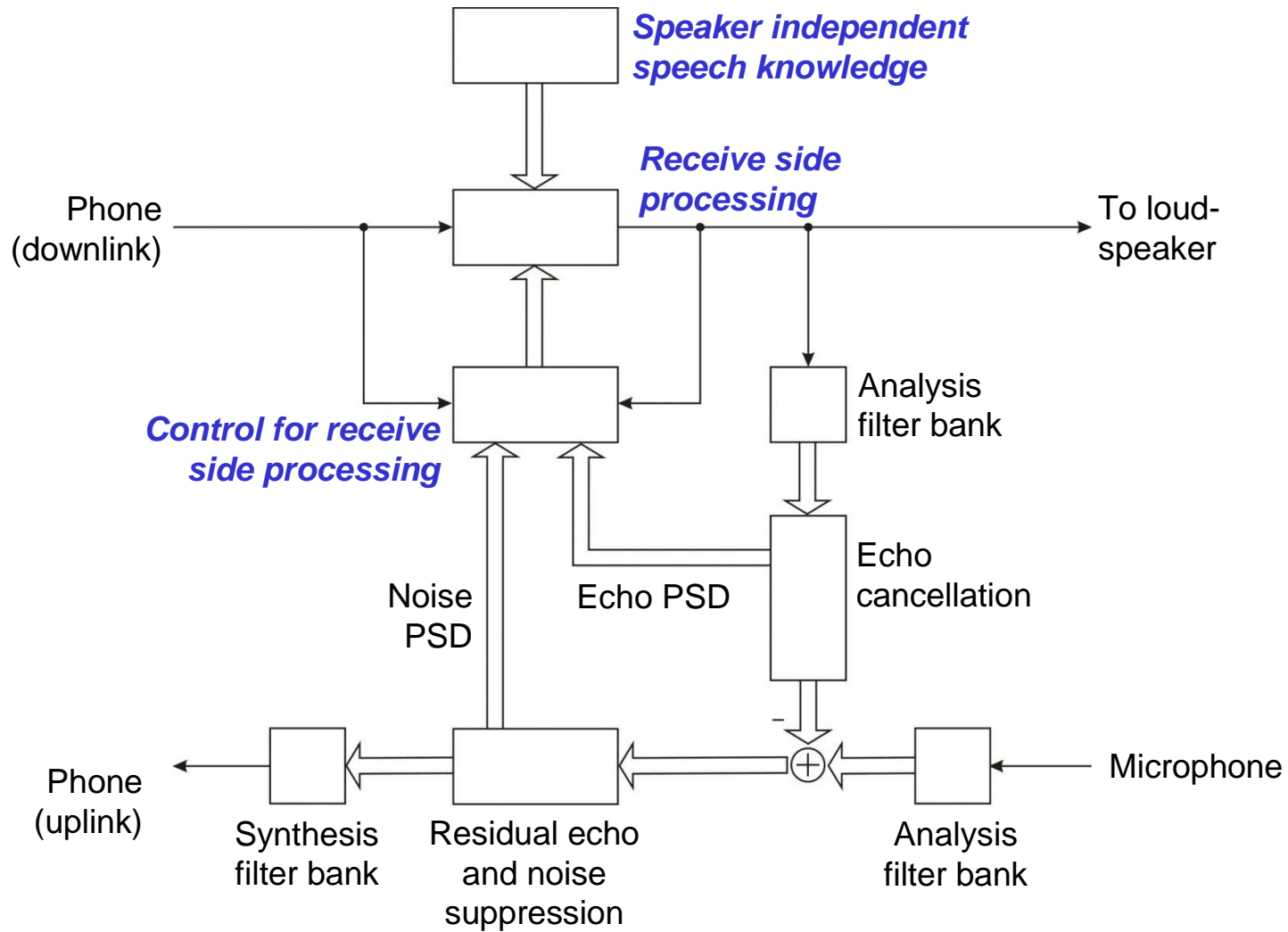
Extended Speech Enhancement

Receive side processing:
gain control,
**bandwidth extension,
adaptive equalization**

Phone

**Speaker independent
speech knowledge**

Speech
dialog
system

**Speaker
(in-)dependent
speech knowledge**

Send side processing:
beamforming, echo
and noise suppression,
speech reconstruction

Speech Enhancement in the Receiving Path (1)

Speaker independent speech knowledge

Receive side processing

Phone (downlink)

To loud-speaker

Control for receive side processing

Analysis filter bank

Noise PSD

Echo PSD

Echo cancellation

Phone (uplink)

Synthesis filter bank

Residual echo and noise suppression

Analysis filter bank

Microphone

Packet loss correction — Noise suppression — Bandwidth extension — Adaptive equalization — Adaptive limiter

Phone (downlink) → To loud-speaker

Control for receive side processing

Speaker independent speech knowledge

Noise PSD   Echo PSD

Input signal

Output signal

Low-frequency extension

Lowpass filter

Speaker independent speech knowledge

High-frequency extension

Highpass filter

Narrowband connection (current standard):

Bandwidth-extension for narrowband speech signals (bandwidth 3.4 …3.8 kHz) – extension of low frequency components and extension of high frequency components up to 5.5 or 8 kHz.
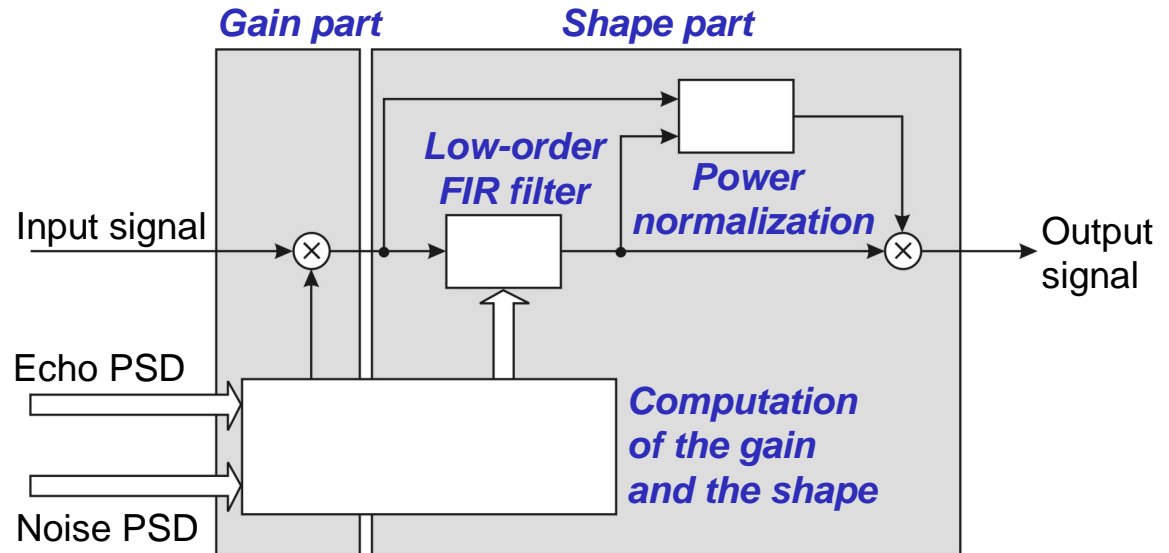
Wideband connection:

Bandwidth-extension for wideband speech signals (bandwidth 7 kHz, e.g. AMR wideband codec G.722.2) – extension of high frequency components up to 11kHz.
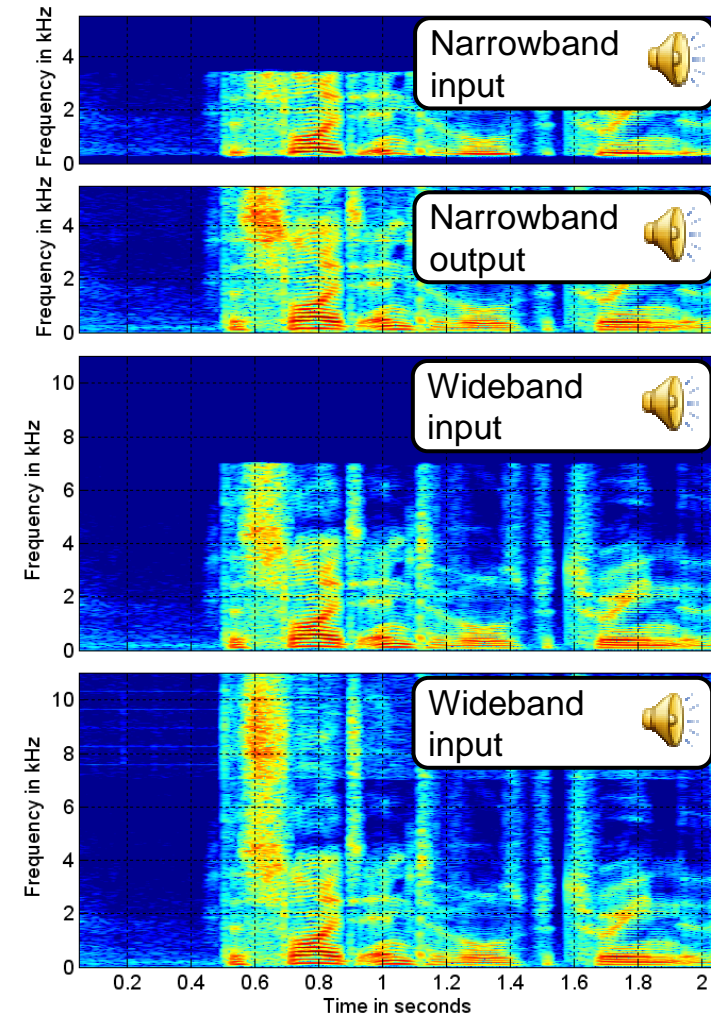
Gain adjustment:

The echo and the background noise power are analyzed and a gain correction is computed in order to achieve a predefined SNR in the passenger compartment.

Gain part   **Shape part**

Input signal

Low-order FIR filter

Power normalization

Output signal

Echo PSD

Noise PSD

*Computation of the gain and the shape*

Shape adjustment:

In addition to the short-term powers also the short-term spectra of the noise and the echo are analyzed and a correction filter is designed in order to boost frequencies with low SNR while slightly attenuating those with good SNR. The design process is computed 10 to 20 time per seconds. An improvement of the speech intelligibility can be achieved while maintaining the loudness of the output signal.

ISO   IEC   ITU International Telecommunication Union

Measurement:

Signal was recorded in an accelerating vehicle, while entering a motorway.
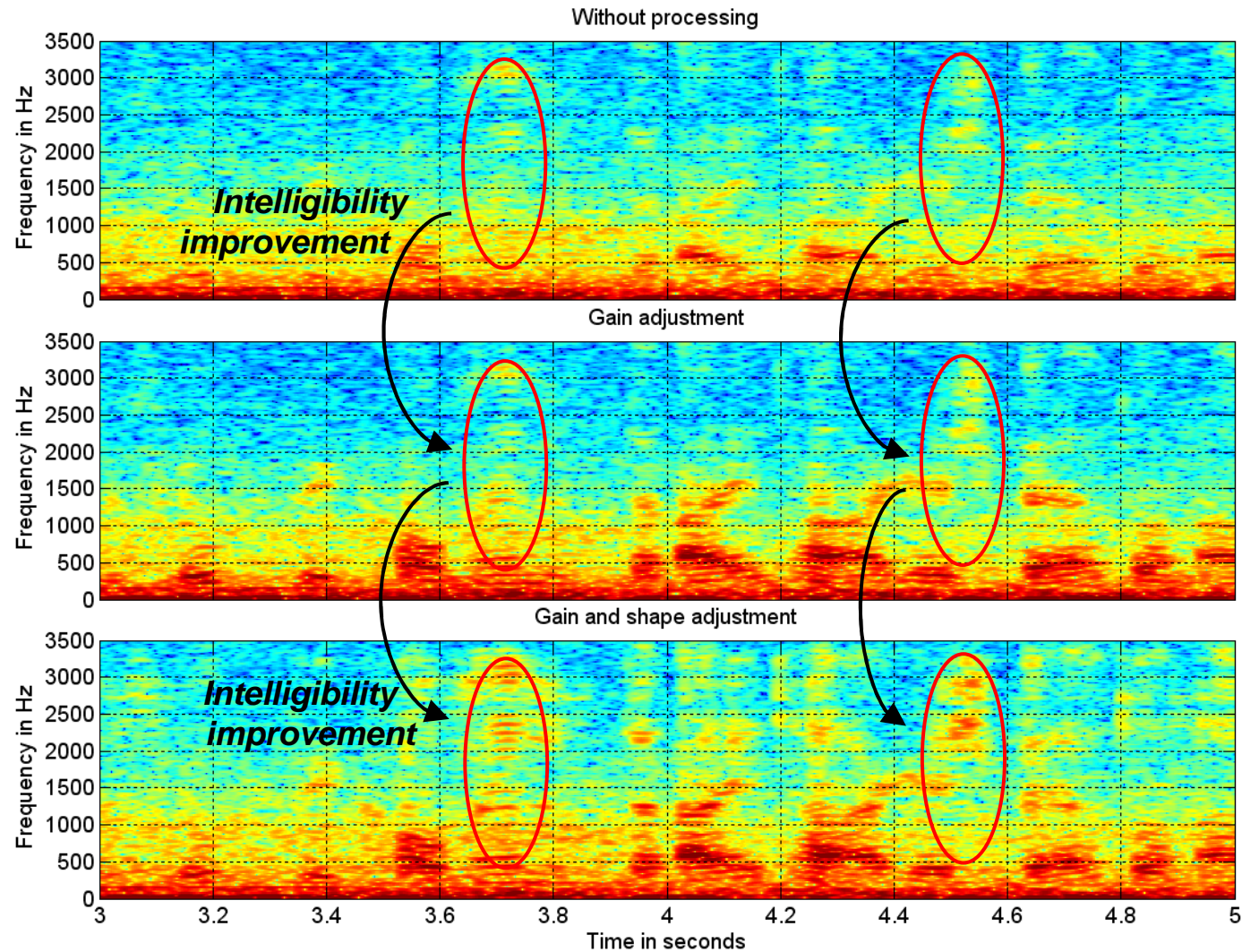
Motivation:

At medium and high speed the SNR often drops at low frequencies below the 0 dB threshold. Thus, standard noise suppression schemes perform only an attenuation at these frequencies.

For further improvement of the speech quality a reconstruction approach is an alternative.

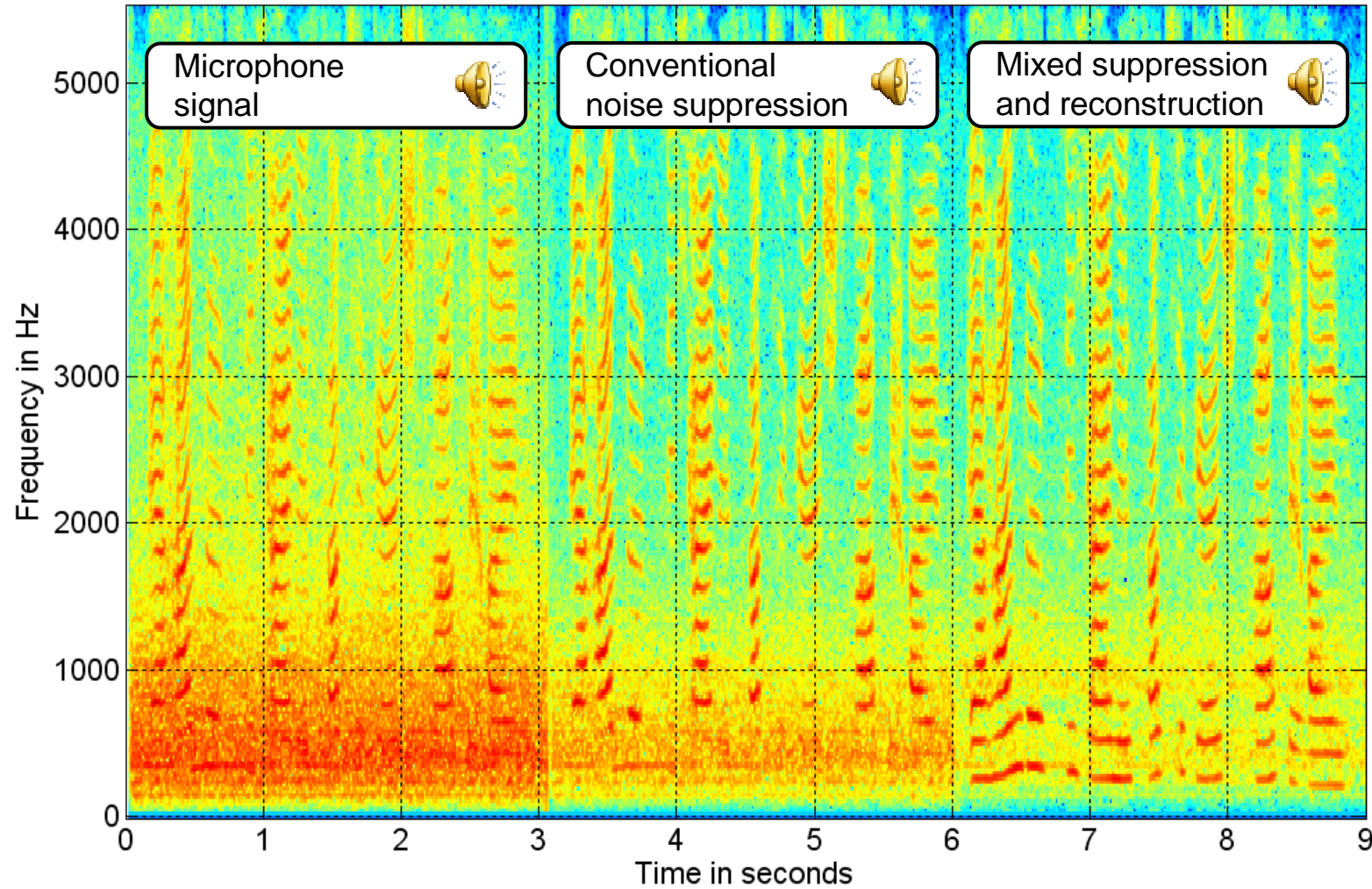However, speech recon-struction starts now where conventional noise reduction fails …

Time-frequency analysis of noisy speech before and after processing
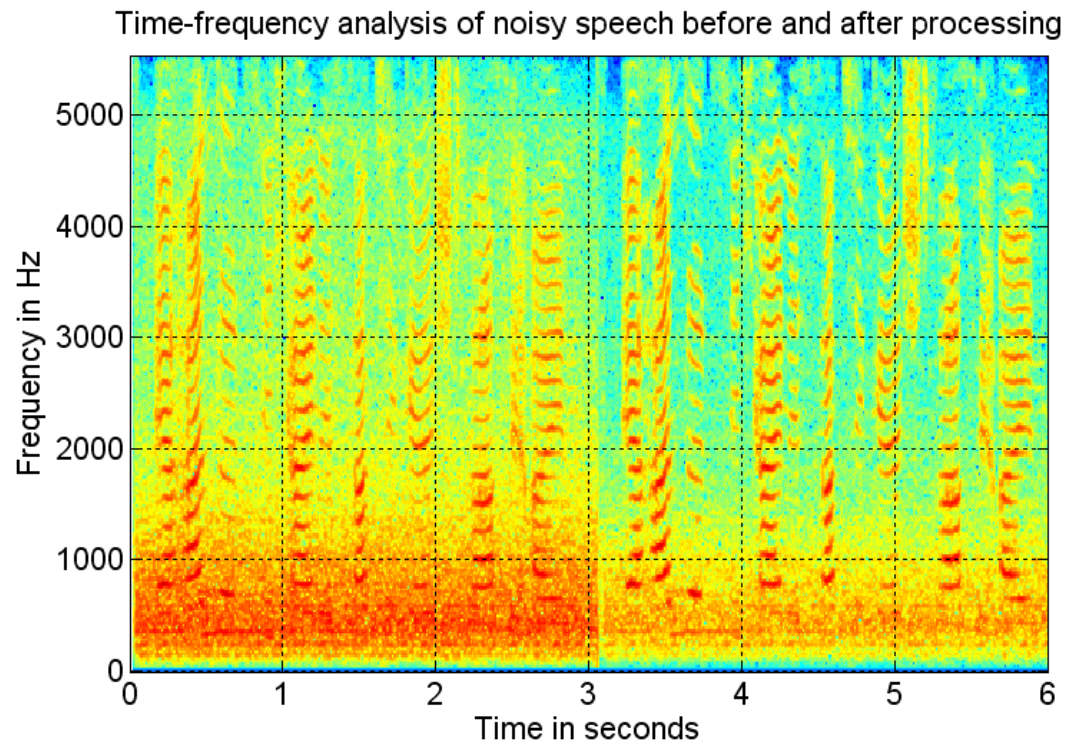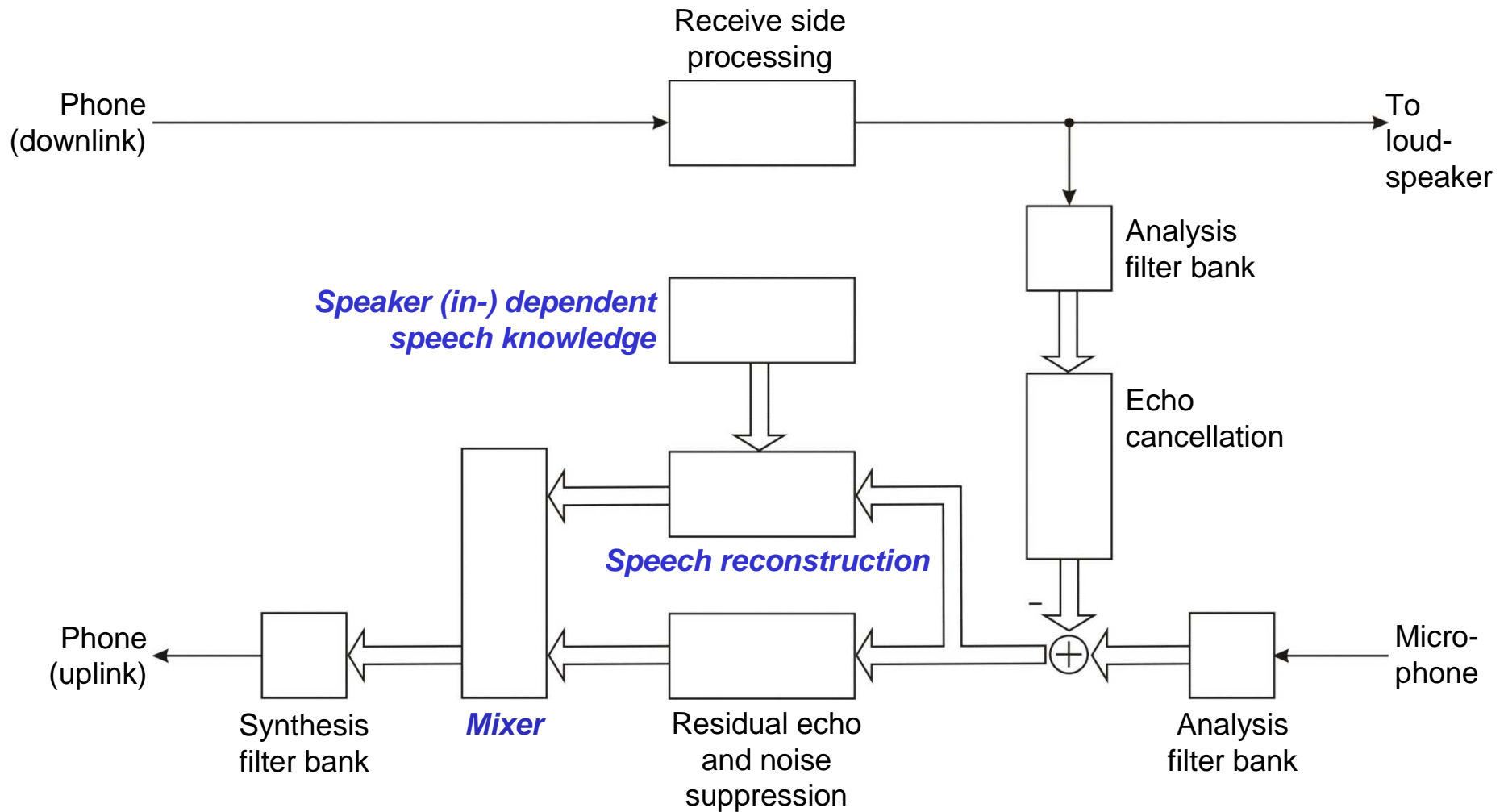
Speech Reconstruction – Audio Examples

Time-frequency analysis of noisy speech before and after processing

Additional Information:

HARMAN/BECKER
AUTOMOTIVE SYSTEMS

Harman/Becker Automotive Systems
Acoustic Signal Processing

Gerhard Schmidt   Tim Haulick
geschmidt@harmanbecker.com   thaulick@harmanbecker.com

Söflinger Str. 100
89077 Ulm, Germany