

Some Notes on the Psychoacoustics and Signal Processing of RASTA-PLP Analysis of Speech

Jörg Buchholz

- Introduction
- Reflection Masking Model
- RASTA-PLP Processing
- RASTA applied to RIR filtered speech

Comments on Perceptual Modelling

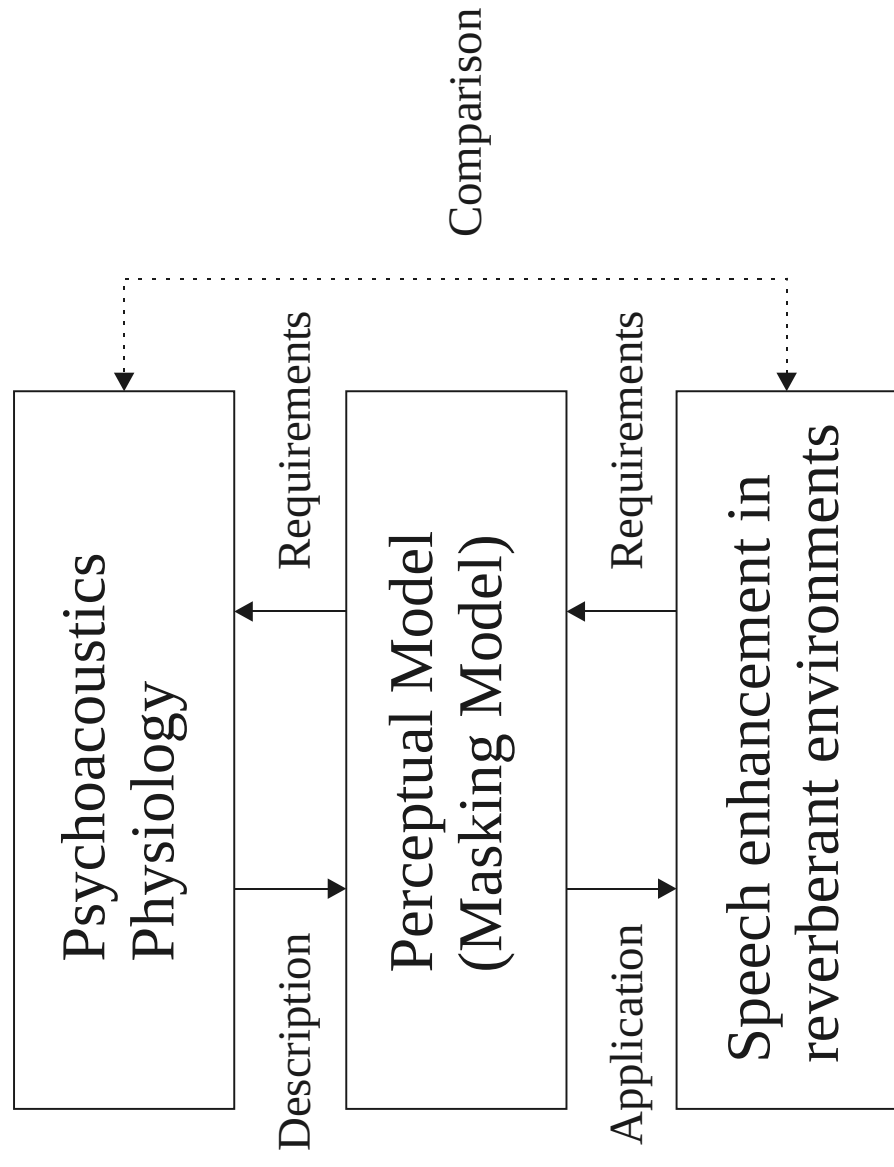
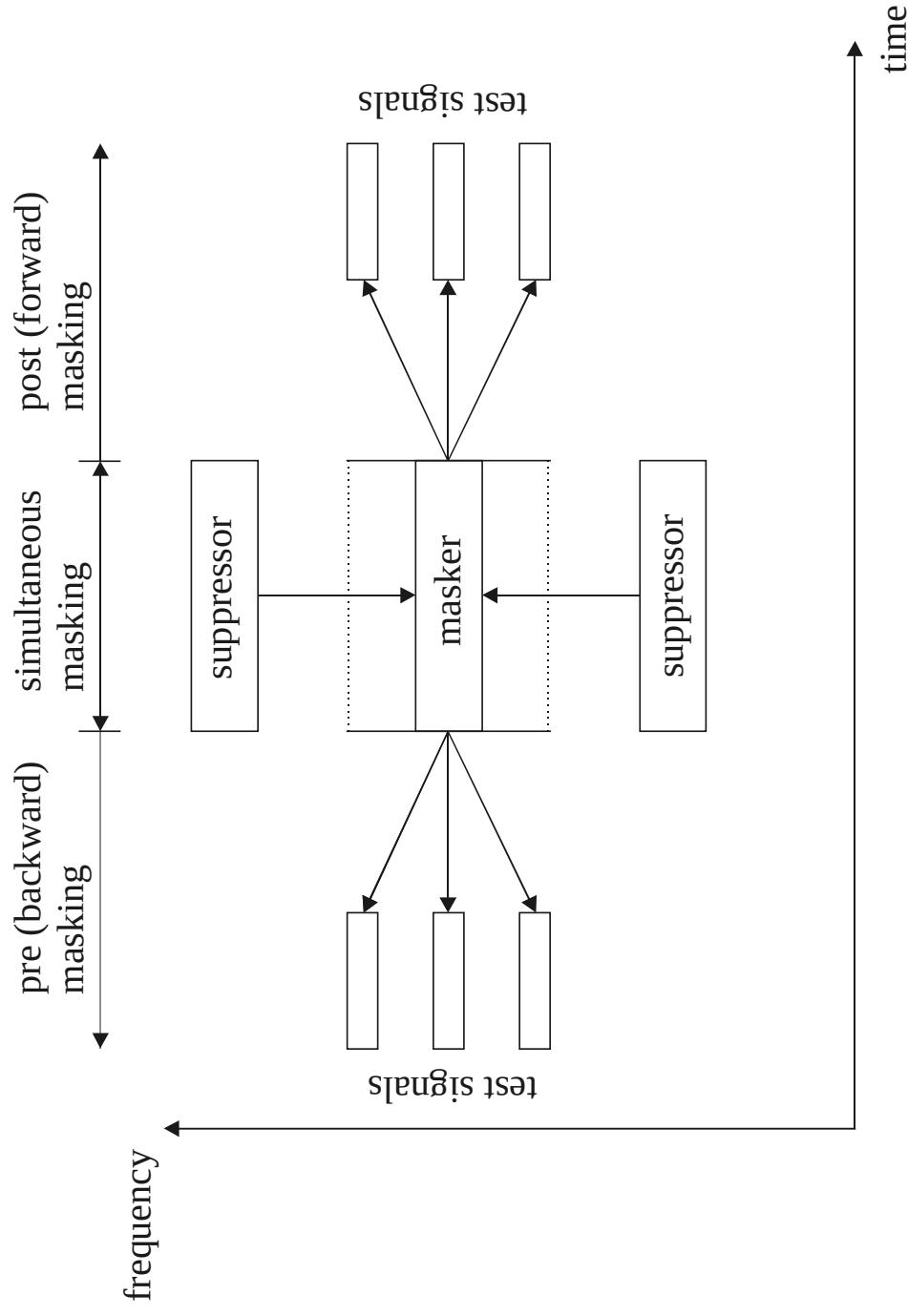
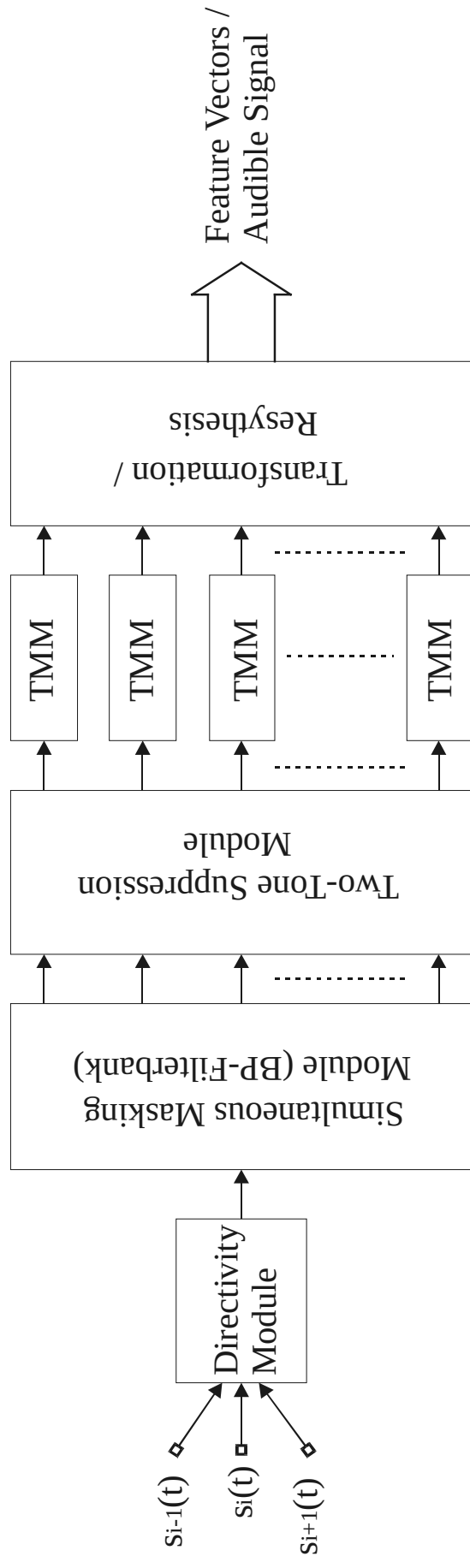


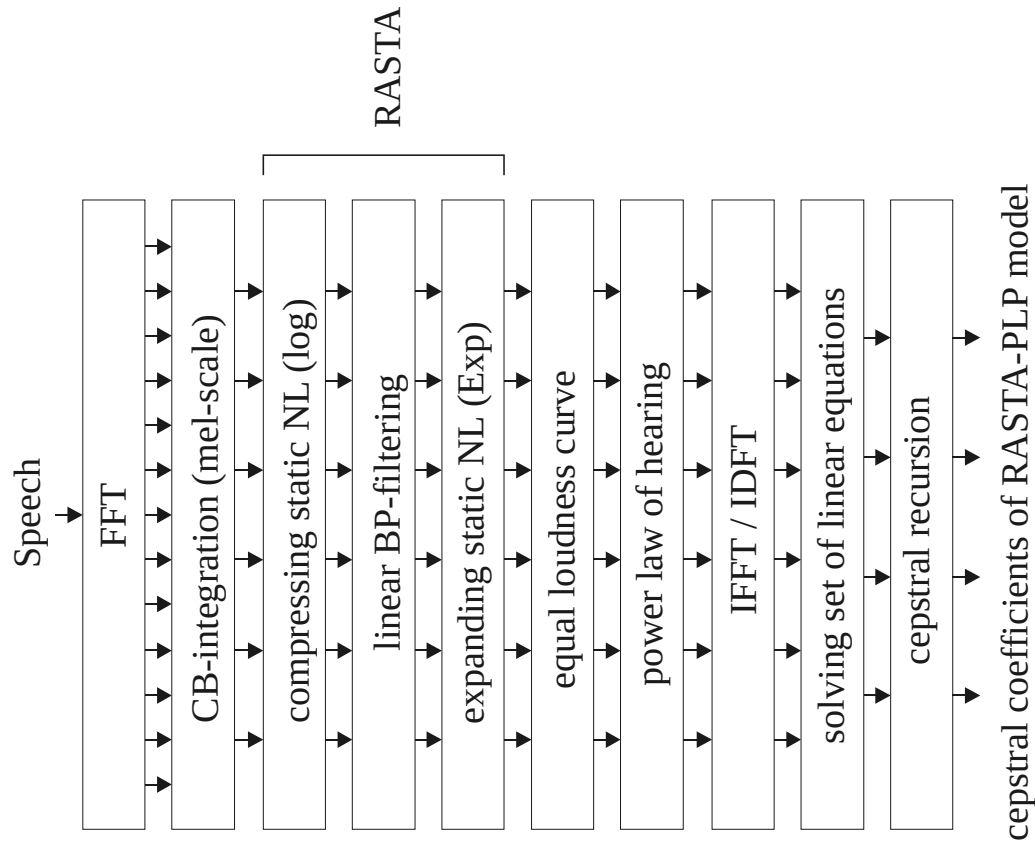
Illustration of Masking and Suppression



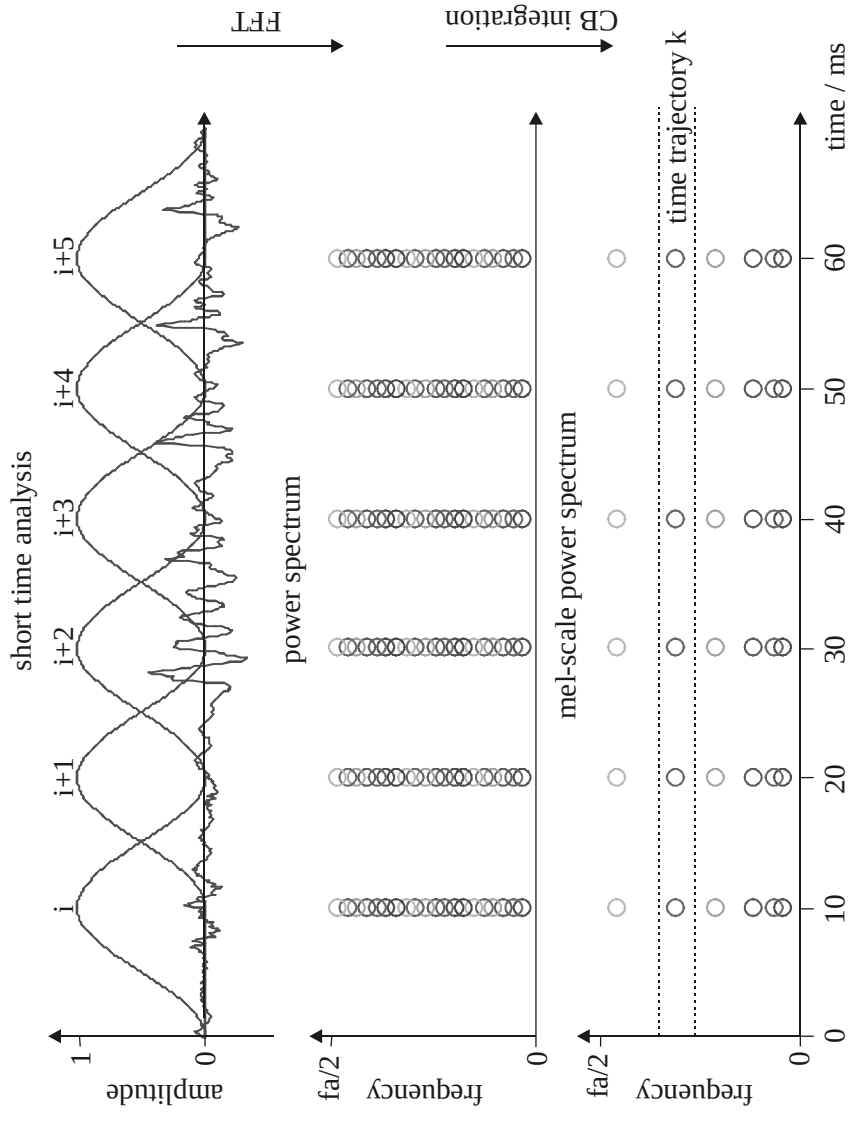
Structure of the Masking Module



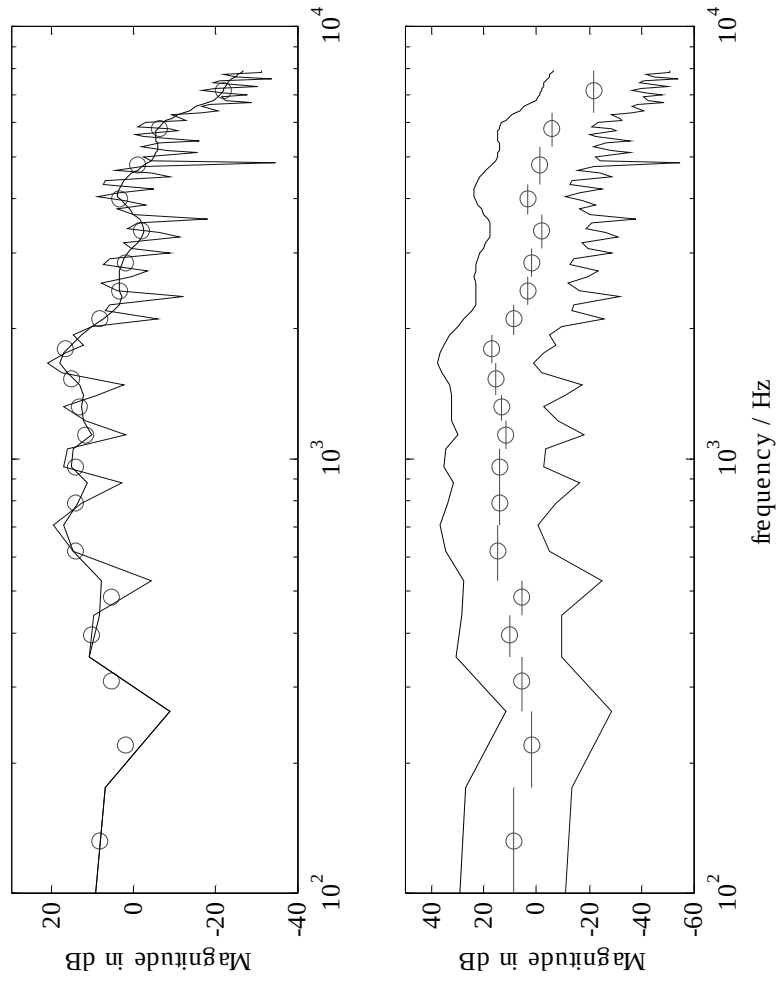
Block Diagramm of the RASTA-PLP Method



Time / Frequency Analysis



CB-integration (mel-scale)



Overview of some RASTA Methods

- Additive Noise (uncorrelated) $\rightarrow$ Lin-RASTA

$$s(t) + n(t) \xrightarrow{FT} S(\omega) + N(\omega)$$

- Convolutional Noise $\rightarrow$ Log-RASTA

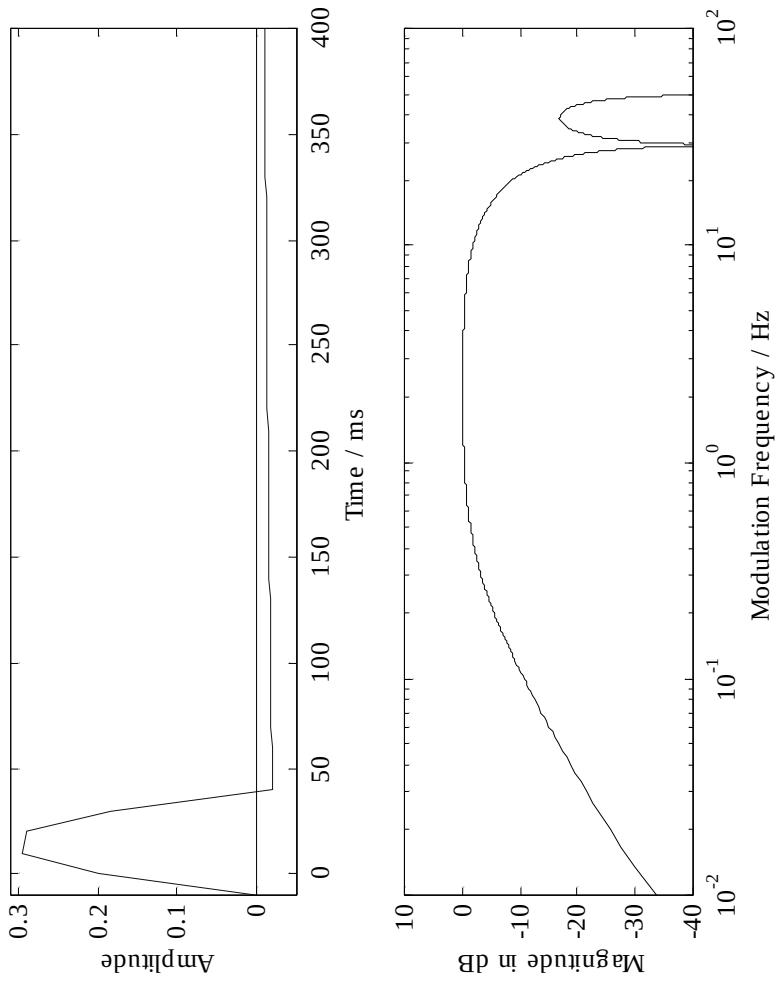
$$s(t) \otimes h(t) \xrightarrow{FT} S(\omega) \cdot H(\omega) \xrightarrow{\log} \log(S(\omega)) + \log(H(\omega))$$

- Convolutional and Additive Noise $\rightarrow$ J-RASTA (Lin/Log)

$$y = \ln(1 + J \cdot x) \quad x = \frac{e^y - 1}{J} \approx \frac{e^y}{J}$$

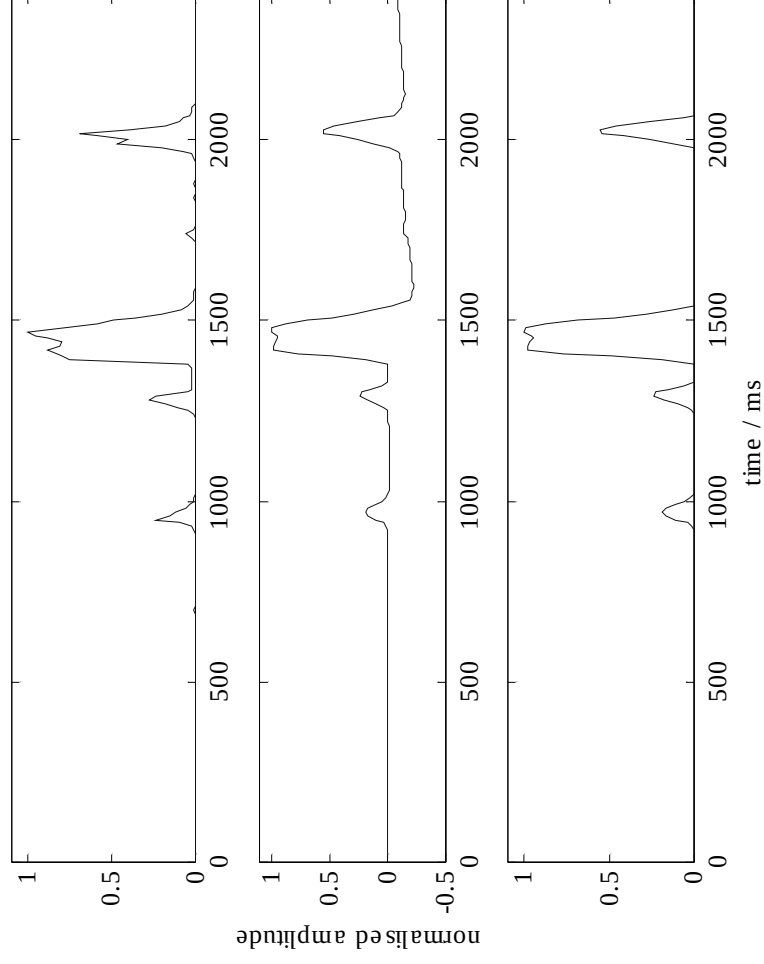
Optimal J depending on the noise power!

Rasta BP-Filter

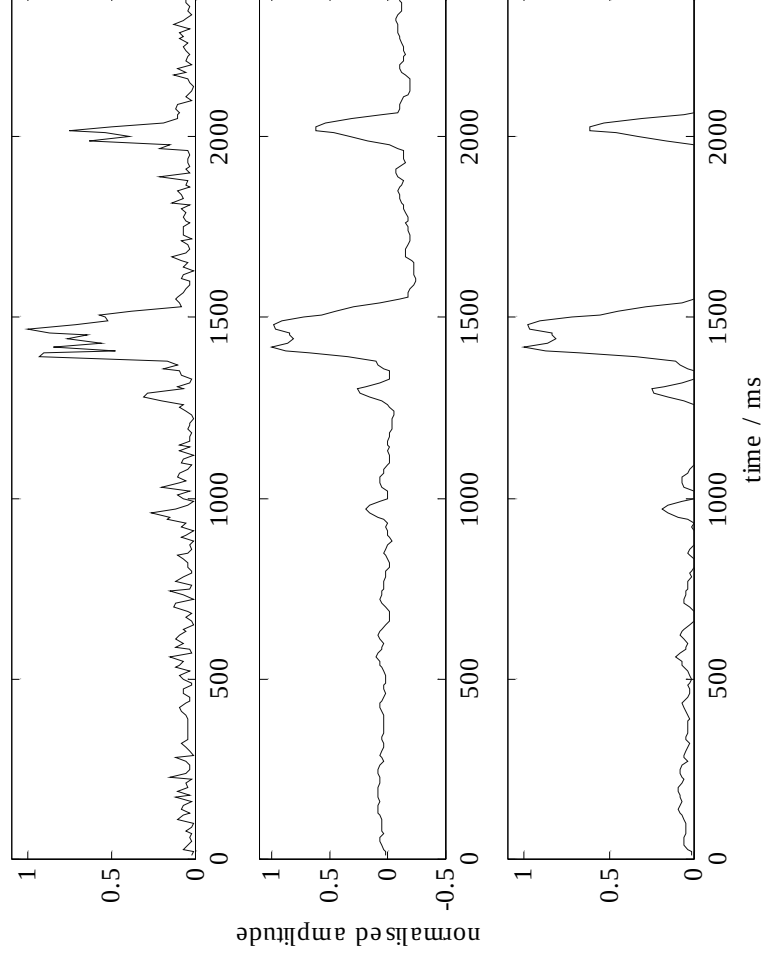


$$H(z) = 0.1 \cdot z^4 \cdot \frac{2 + z^{-1} - z^{-3} - 2z^{-4}}{1 - 0.98z^{-1}}$$

LIN-RASTA-Processing of clean speech



Lin-RASTA-Processing of noisy speech (-5 dB SNR)



Time Trajectory (1 kHz band)

BP-filtered Time Trajectory

Negative values set to zero

RASTA applied to Room Impulse Responses

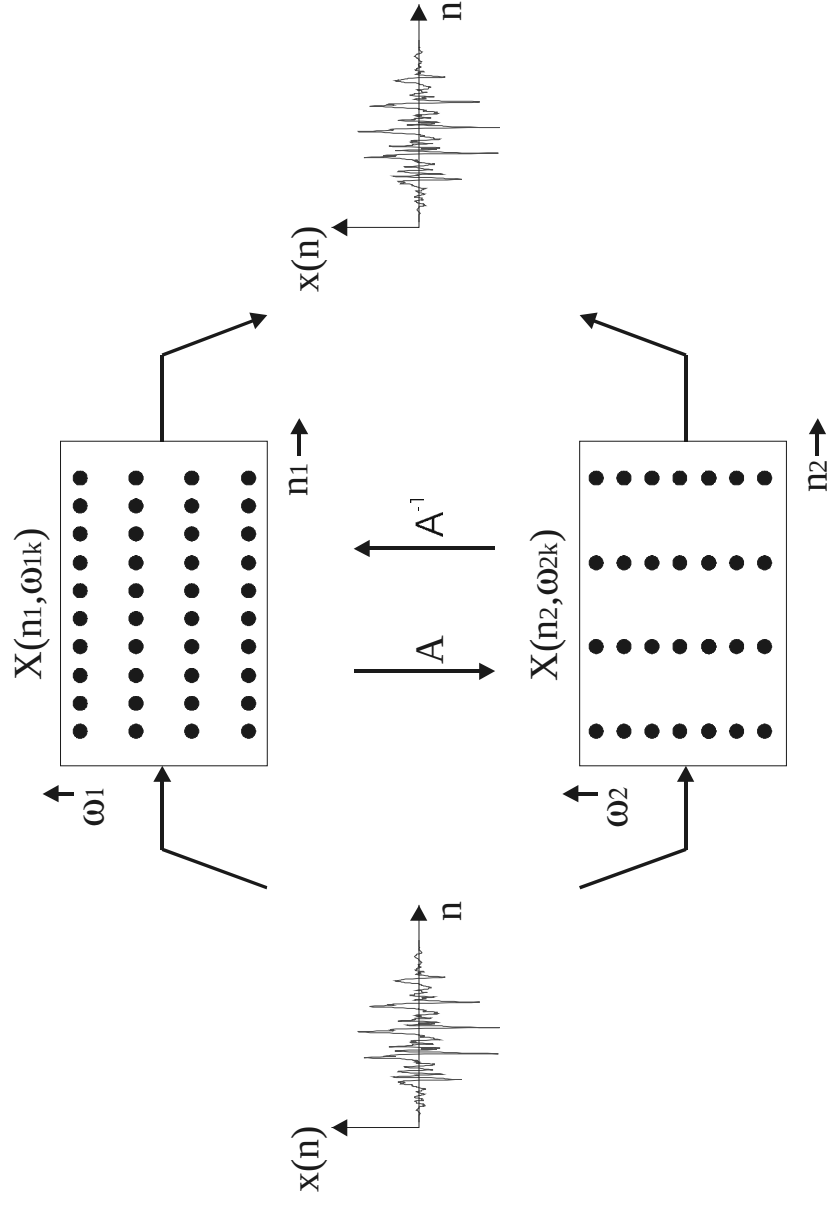
Basic Assumptions for RASTA processing:

- Analysis Window length $\gg$ Filter (RIR) length
- Filter (RIR) should be constant (slowly changing) for a duration $\gg$ Analysis Window length
- For ASR: Window length $\approx$ 20-30 ms Conflict!!

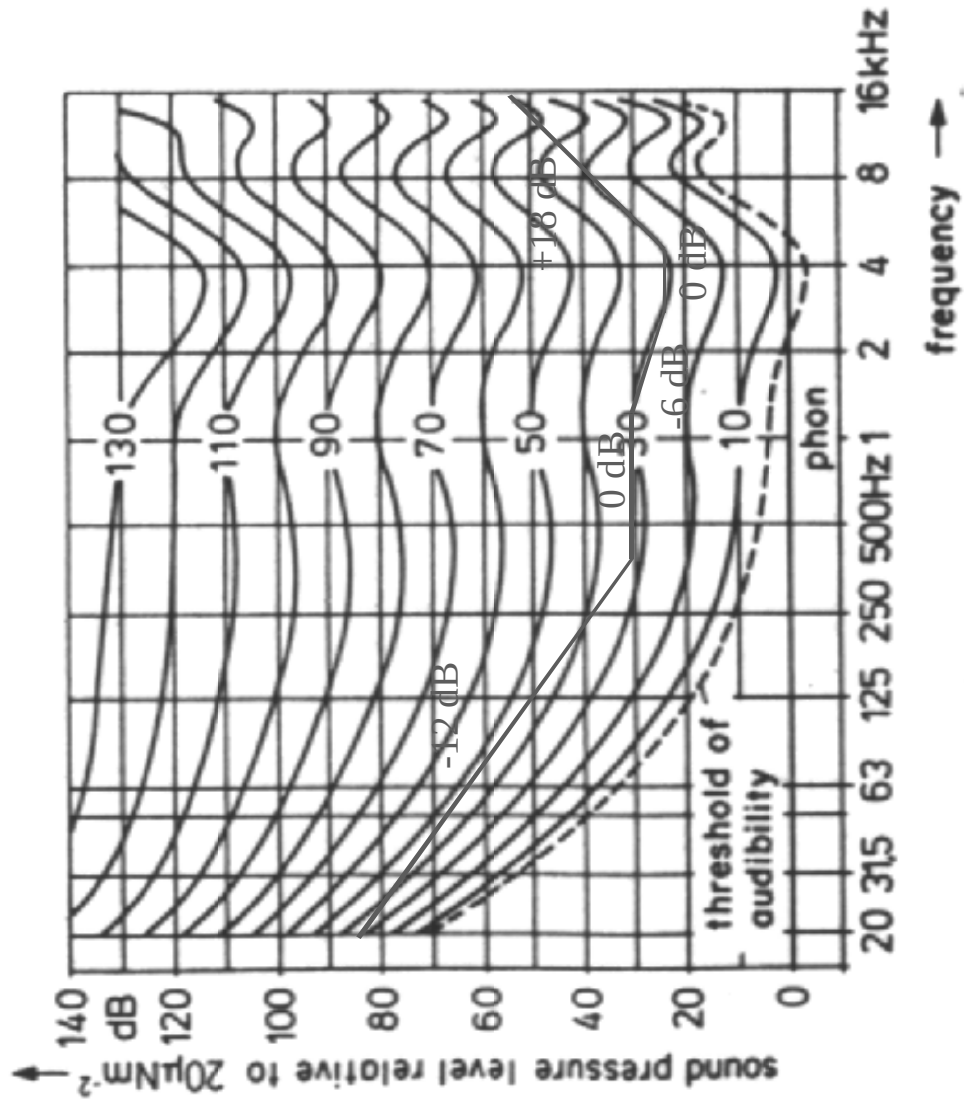
	Baseline	Multiresolution
Clean	8,6 %	13,5 %
Reverberant	34,8 %	22,8 %

Isolated Digits Word Error Rates

Multiresolution Processing Concept (Avendano)



Equal-Loudness Curves



Loudness function (Zwicker)

