Simulation data and a model approach for speech perception with electric-acoustic stimulation (EAS)

Tobias Rader[1], Youssef Adel[1], Uwe Baumann[1], Hugo Fastl[2]
[1] Audiological Acoustics, ENT Department, University Hospital Frankfurt, 60590 Frankfurt am Main, Germany
[2] AG Technische Akustik, MMK, Technische Universität München, 80333 Munich, Germany
Corresponding author: Tobias.Rader@kgu.de

Introduction

Electric-acoustic stimulation (EAS) denotes simultaneous stimulation of high-frequency hearing by means of a cochlear implant (CI) and of residual low-frequency hearing up to 500 Hz by acoustic stimulation in the same ear. Patients implanted and fitted according to the EAS concept show significantly higher speech intelligibility in complex noise environments compared to bilaterally implanted CI patients.

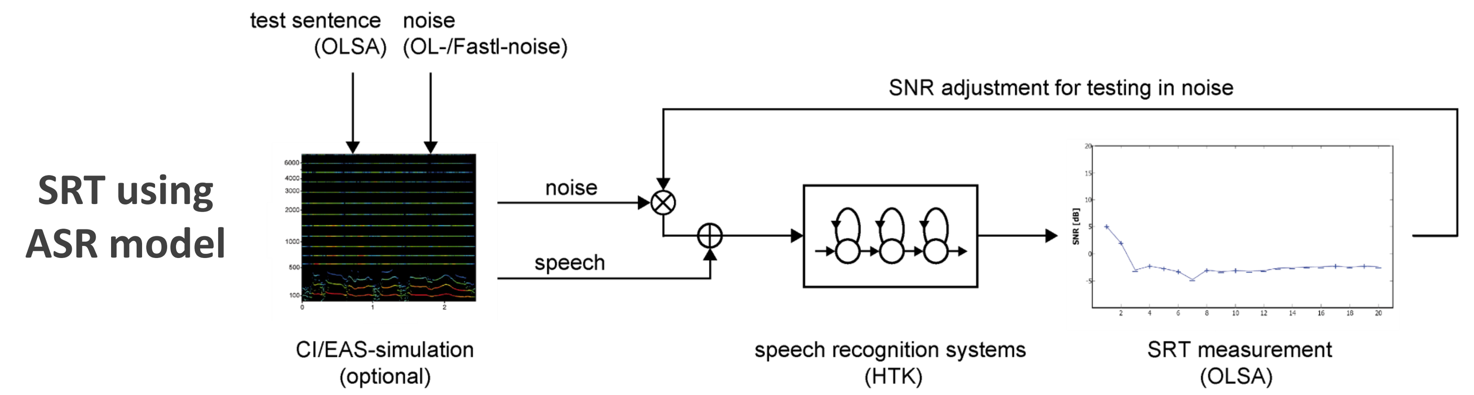
To investigate the effect of EAS on speech perception in noise we developed a simulation to mimic electric-acoustic stimulation using recordings of the German Oldenburg sentence test (**OLSA**) and two types of competing noise: (1) pseudo-continuous **OL-noise** and (2) amplitude-modulated **Fastl-noise**. Speech perception scores were obtained using a specialized automatic speech recognition (ASR) model to determine characteristic parameters of the synergic EAS results.

Materials and Methods

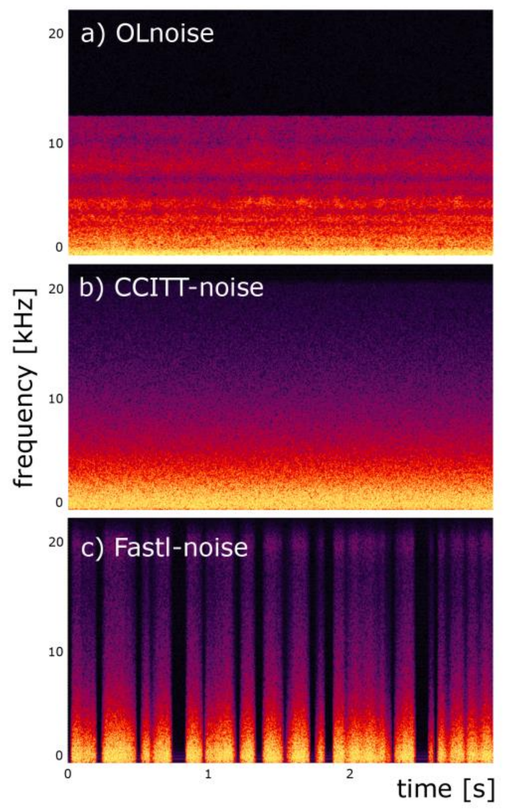
The acoustic time signal of OLSA was transformed by means of an ear-related spectral transformation with subsequent peak-picking into a stream called "part-tone time pattern" (PTTP). Part-tone frequency was reordered following the 12 center frequencies of MED-EL DUET CI speech processors and re-synthesized employing a 12-channel sinusoidal vocoder to simulate electrical stimulation. The acoustic component of EAS was simulated by low-pass filtering the input signal with cut-off frequencies 200 Hz to 500 Hz.

The ASR model was built using the HTK Toolkit (Cambridge University) and trained using 20 OLSA lists with 30 sentences each. Coding was done using a filter bank analysis in 24 critical bands between 0 and 8 kHz. In this regard, the ASR model was specialized to achieve perfect score in the OLSA test.

Initial model evaluation was carried out using a fixed signal-to-noise ratio (SNR) for different signal simulations to determine principal speech recognition rates. Subsequently, speech reception thresholds (SRT) were measured using an adaptive procedure and compared to results with CI and normal-hearing subjects [Rader et al., 2013].

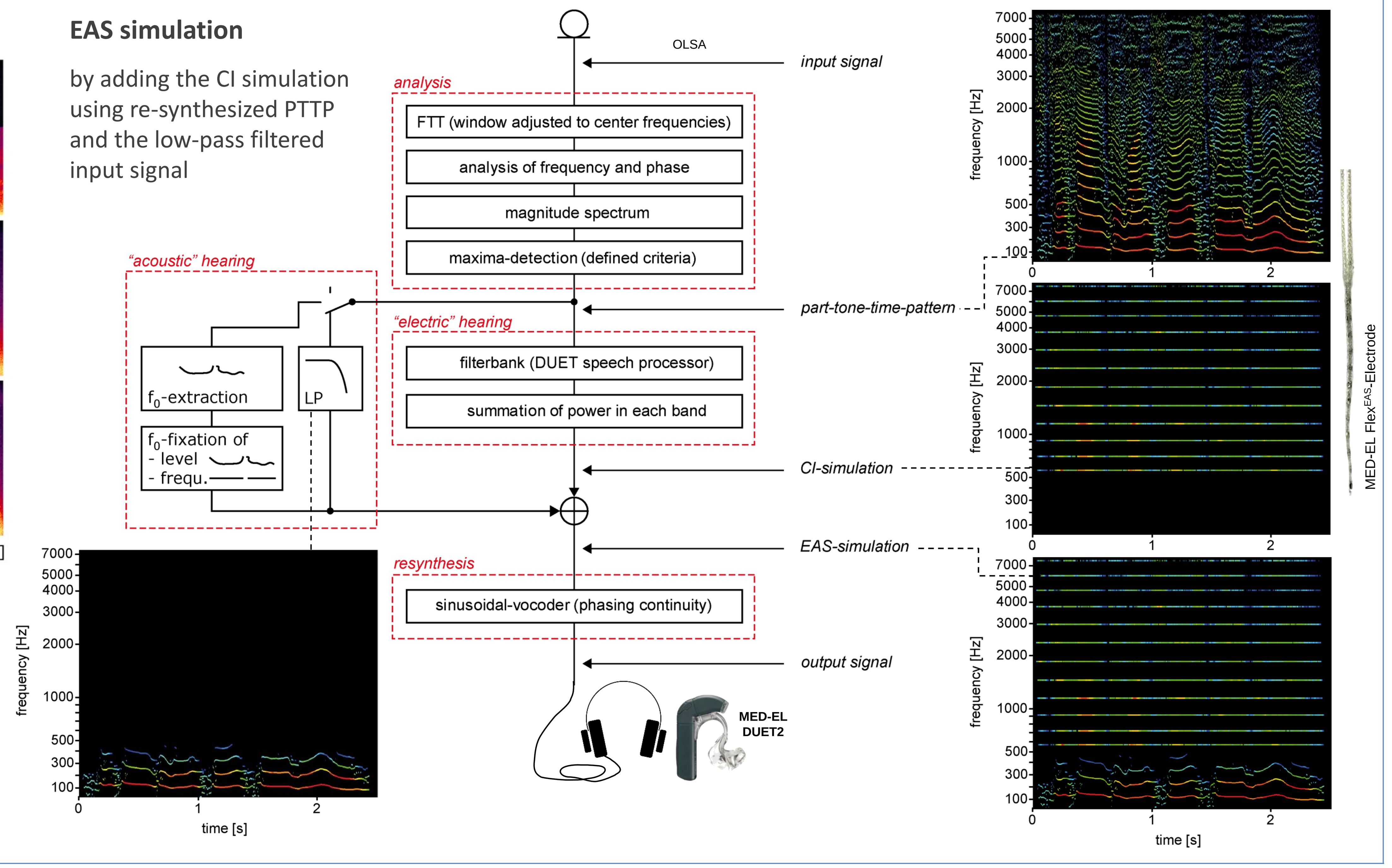


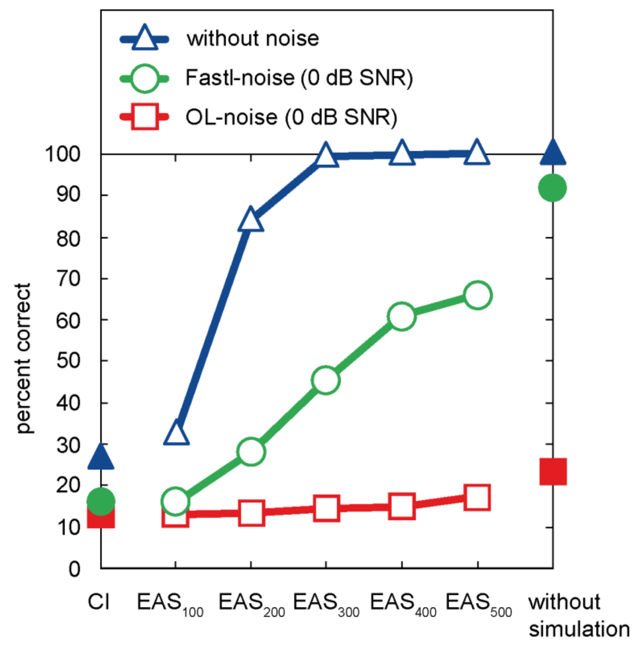
noise signals

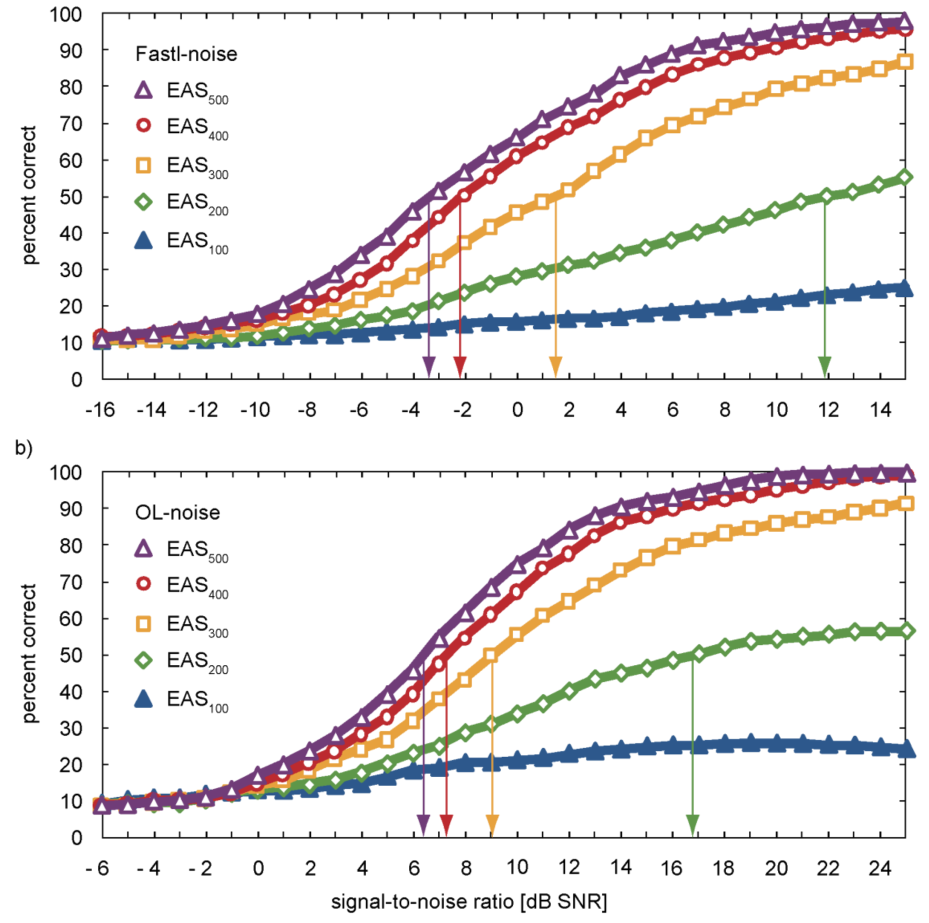


EAS simulation

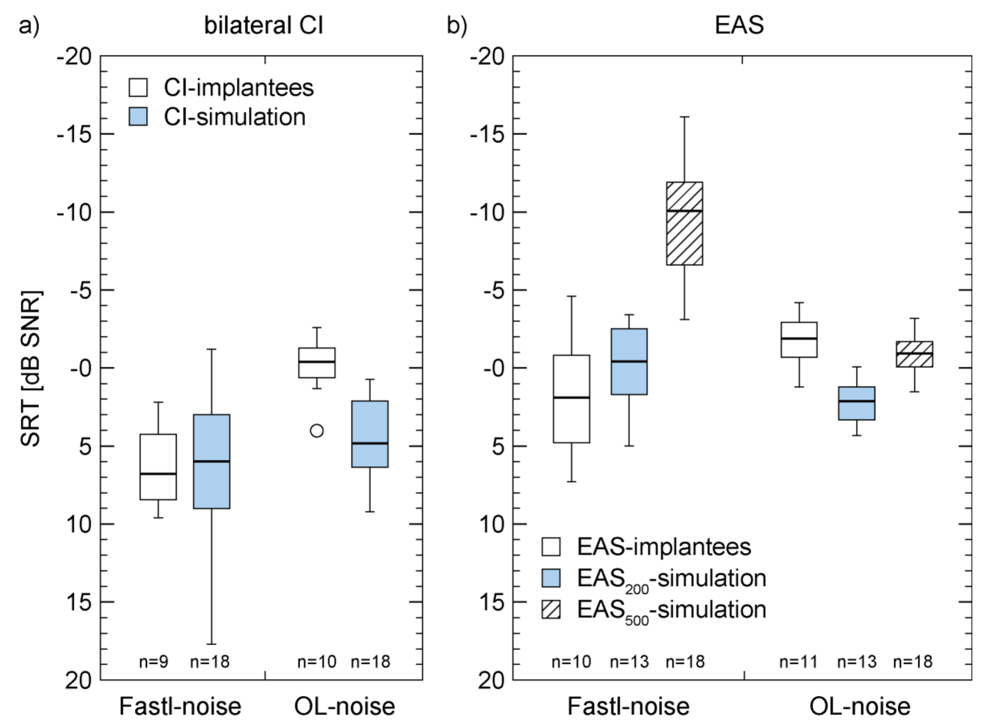
by adding the CI simulation using re-synthesized PTTP and the low-pass filtered input signal







SRT [dB SNR]	without sim.	EAS_{500}	EAS_{400}	EAS_{300}	EAS_{200}	EAS_{100}
Fastl-noise	-7.5	-3.4	-2.2	1.4	11.9	n/a
OL-noise	3.4	6.3	7.2	9.0	16.8	n/a



Results

For the fixed SNR evaluation set to 0 dB SNR, CI simulation recognition rates are slightly higher than chance (10 possibilities per word, 5 words per sentence, i.e. 10%) with 16.3% for Fastl-noise and 13.5% for OL-noise. In quiet, the recognition rate of 26.3% is nearly twice as good as measured with OL-noise. After adding the low-pass filtered signal as the acoustic component of EAS simulation, increased recognition rates were observed for higher cut-off frequencies. However, this effect is only evident for the noise conditions "without noise" and "Fastl-noise". The condition "OL-noise" is limited by the low recognition rate for clean (without simulation) input signals.

Using an adaptive SNR procedure, speech reception thresholds for the ASR model were determined. Recognition rates are shown (a) from −16 to 14 dB SNR for Fastl-noise, and (b) from −6 to 24 dB SNR for OL-noise. And the determined SRT values for Fastl- and OL-noise are listed in the table below. The EAS condition EAS_{100}, i.e. with low-pass filter cut-off at 100 Hz, did not achieve 50% in the examined range, and thus no SRT can be reported.

As reference, SRTs were measured with CI and normal-hearing subjects in Fastl- and OL-noise: (a) bilateral CI subjects compared to normal-hearing subjects tested with CI simulation signals, and (b) EAS subjects compared to normal-hearing subjects tested with EAS_{200} and EAS_{500} simulation.

Discussion

Analogous SRT results were found for normal-hearing subjects and the ASR model, although shifted ca. 8 dB SNR apart. The synergic effect of EAS for different cut-off frequencies of the low-pass filtering (simulating residual hearing) was demonstrated and a significant amount of speech information was found for cut-off frequencies ≥ 300 Hz. This result could be considered when examining eligibility criteria for EAS.

SRT improvement for Fastl-noise can be traced back to gap listening for both the normal-hearing subjects as well as the ASR model. The reversed effect for CI and EAS users could motivate further investigation, whereas Rader et al. (2013) exclude the presence of gap listening for this group.

Another aspect of consideration are the center frequencies used for the re-synthesis of the CI simulation. While the inter-electrode distance is fixed, affected critical bands can be shifted based on insertion depth. Hence electric and acoustic stimulation might overlap in the lower frequencies.

References

[Rader et al., 2013] Rader, T., H. Fastl, and U. Baumann (2013): *Speech Perception With Combined Electric-Acoustic Stimulation and Bilateral Cochlear Implants in a Multisource Noise Field.* Ear and Hearing 34(3): 324-332.