KU LEUVEN

Enhanced temporal coding in cochlear implants

Raphael Koning and Jan Wouters

ExpORL, University of Leuven, Belgium

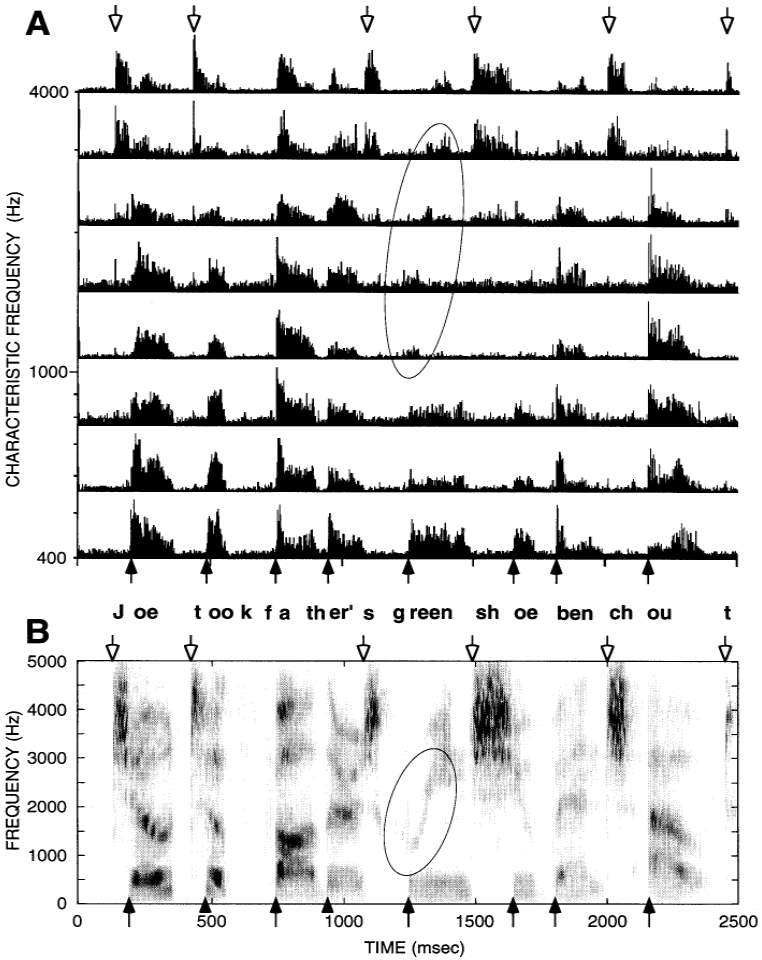
Speech in Noise 2014, Marseille, France

Outline

Challenges when listening with CIs

- Speech intelligibility (SI) in quiet can be very good for cochlear implant (CI) users.
- CI users suffer from a huge decrease of speech understanding in adverse listening conditions like
  - interfering background sounds
  - reverberant scenarios
  - cocktail party scenarios.
- Development of speech enhancement strategies to increase speech perception in adverse listening conditions.
- Two broad categories of speech enhancement algorithms
  - Speech envelope enhancement algorithms
  - Noise reduction algorithms.

Rationale: envelope enhancement (1)



Rapid adaptation effect of the auditory nerve:

- Pointing to spectra-temporal regions that are rich in phonetic information.
- Increasing the temporal precision of the onsets.
- Enhancing spectral contrast between successive segments.
- Encoding phonetic contrasts based on the characteristics of the amplitude envelope.

Figure: Post-stimulus-time histogram of the activity of the cat auditory nerve in response to "Joe took father's green shoe bench out" and the respective spectrogram (Delgutte, 1999).

Rationale: envelope enhancement (2)

- Short-term adaptation and onset effect has its origin at the auditory nerve synapse and is not represented in the signal processing of a CI.
- Envelope enhancement in CI:
  - enhanced envelope continuous interleaved sampling (EECIS, Geurts and Wouters (1999))
  - transient emphasis spectral maxima (TESM, Vandali (2001))
- Phoneme transitions and rapid changes in temporal and spectral content (transients, onsets, offsets) contribute most to speech intelligibility (Chen *et al.* (2012), Stilp & Kluender (2010), Lewicki (2010), Lee & Kewley-Port (2009)).
- Re-investigation in adverse listening conditions and new signal processing strategy.

Outline

General CI processing steps

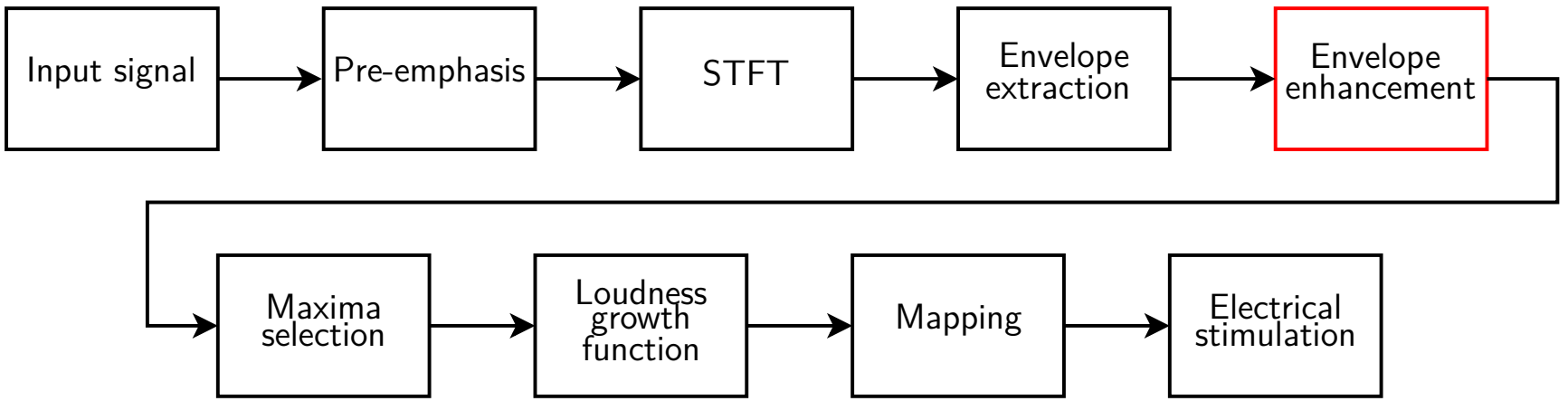


Figure: Processing steps of a CI with the envelope enhancement block.

Advanced combination encoder (ACE) strategy:

- Clinical speech processing strategy of Cochlear, Ltd.
- Envelope enhancement block is bypassed.
- Serves as the reference strategy.

EE strategy: Peak signal extraction

Addition of a peak signal at the onsets in each frequency band to the bandpass filtered envelope

$$E_{\mathrm{EE}}(\lambda, k) = E(\lambda, k) + E_{\mathrm{peak}}(\lambda, k).$$

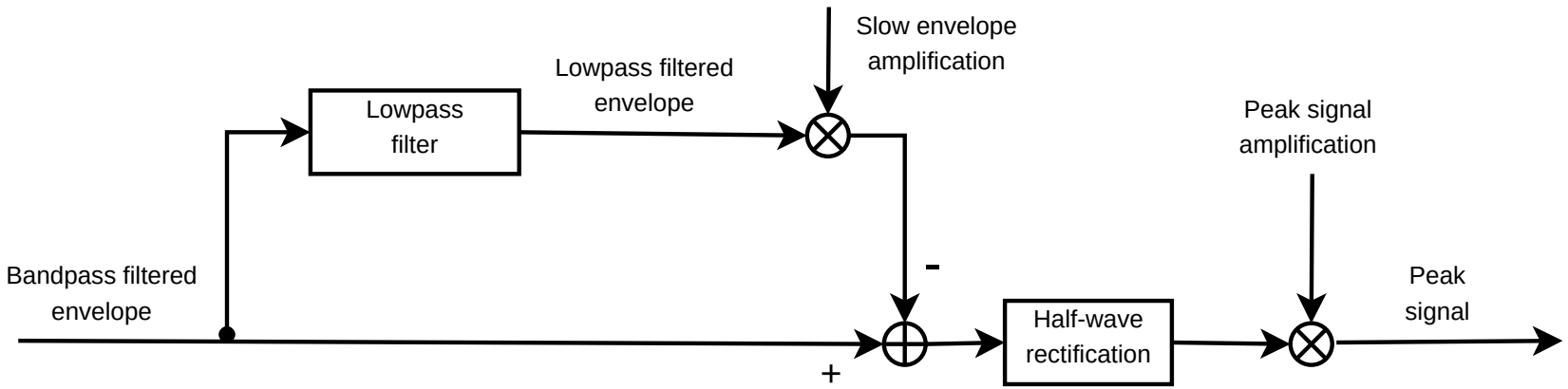


Figure: Onset extraction by a comparison between an input envelope and a (delayed) lowpass filtered version of the input envelope in each frequency band (Koning & Wouters, 2012).

$$E_{\mathrm{peak}}(\lambda, k) = A_{\mathrm{peak}} \max\left[E_{\mathrm{in}}(\lambda, k) - A_{\mathrm{slow}} E_{\mathrm{slow}}(\lambda, k), 0\right]$$

Example: Enhanced envelopes in 8 channels

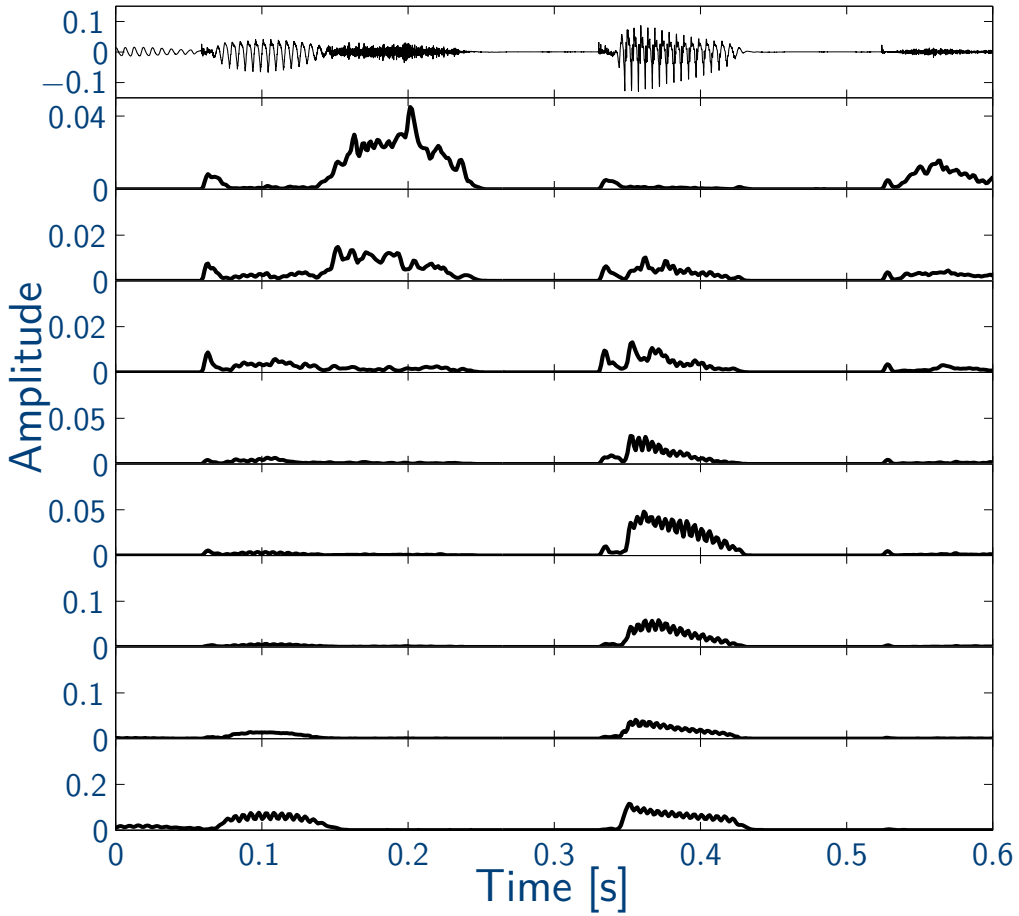


Figure: Eight channel processing of the Dutch words "de stad" for the reference CIS (black) processing and the EE strategy (red).

Eight channel processing of the Dutch words "de stad" (the city)

- CIS: black solid line

Example: Enhanced envelopes in 8 channels

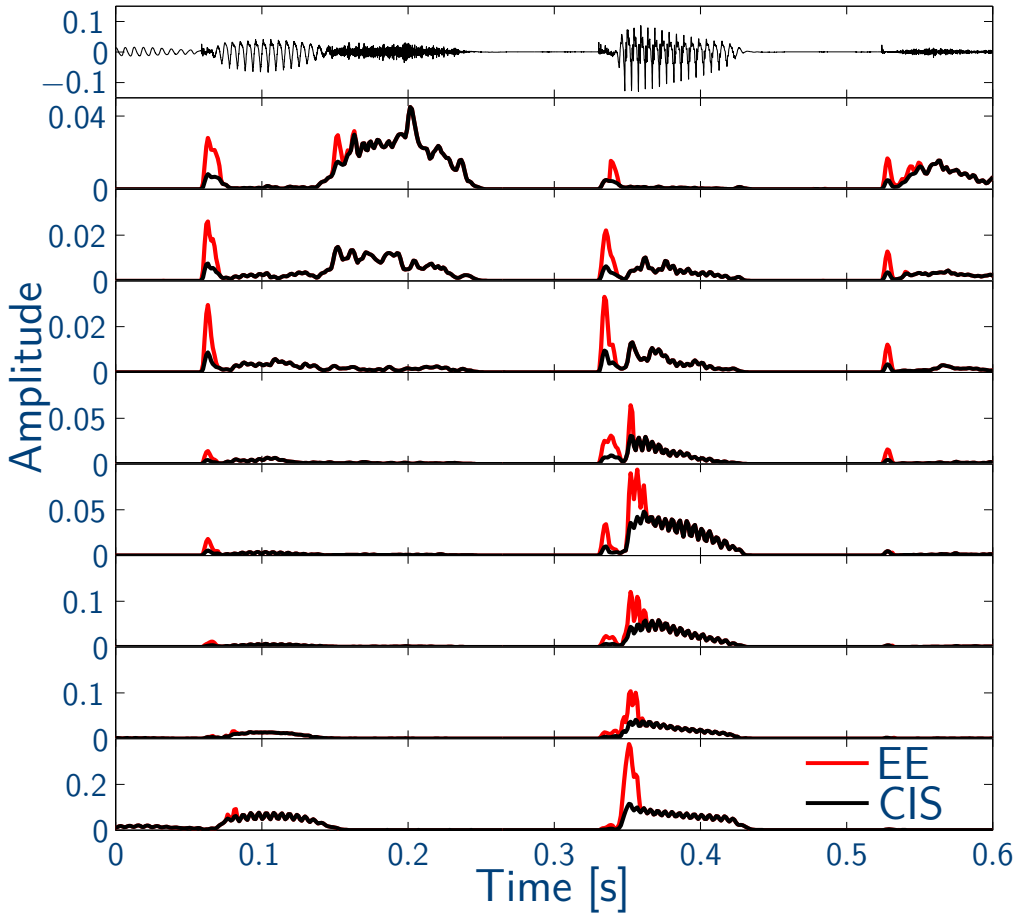


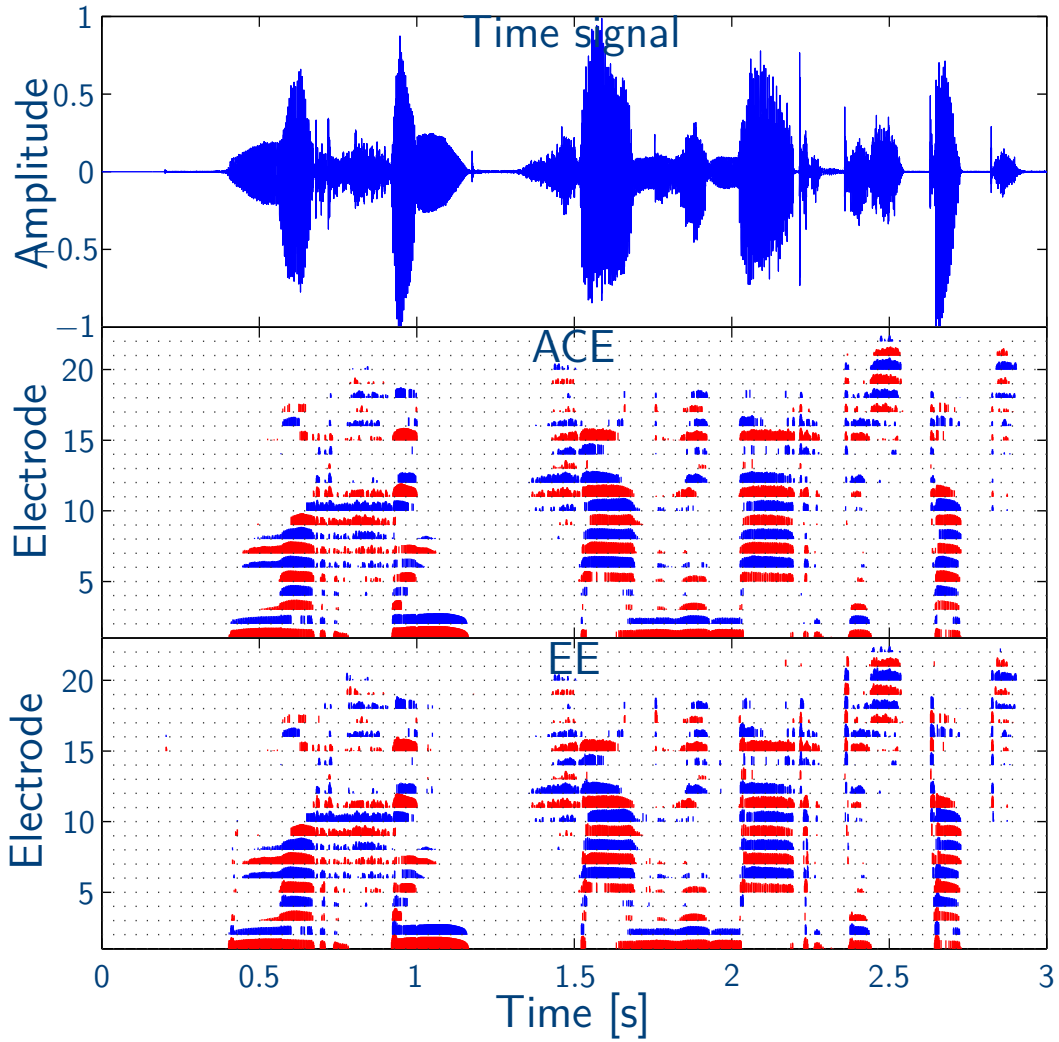
Figure: Eight channel processing of the Dutch words "de stad" for the reference CIS (black) processing and the EE strategy (red).

Eight channel processing of the Dutch words "de stad" (the city)

- CIS: black solid line
- EE: red solid line

The stationary part of the signal is not affected by the envelope enhancement.

EE in quiet

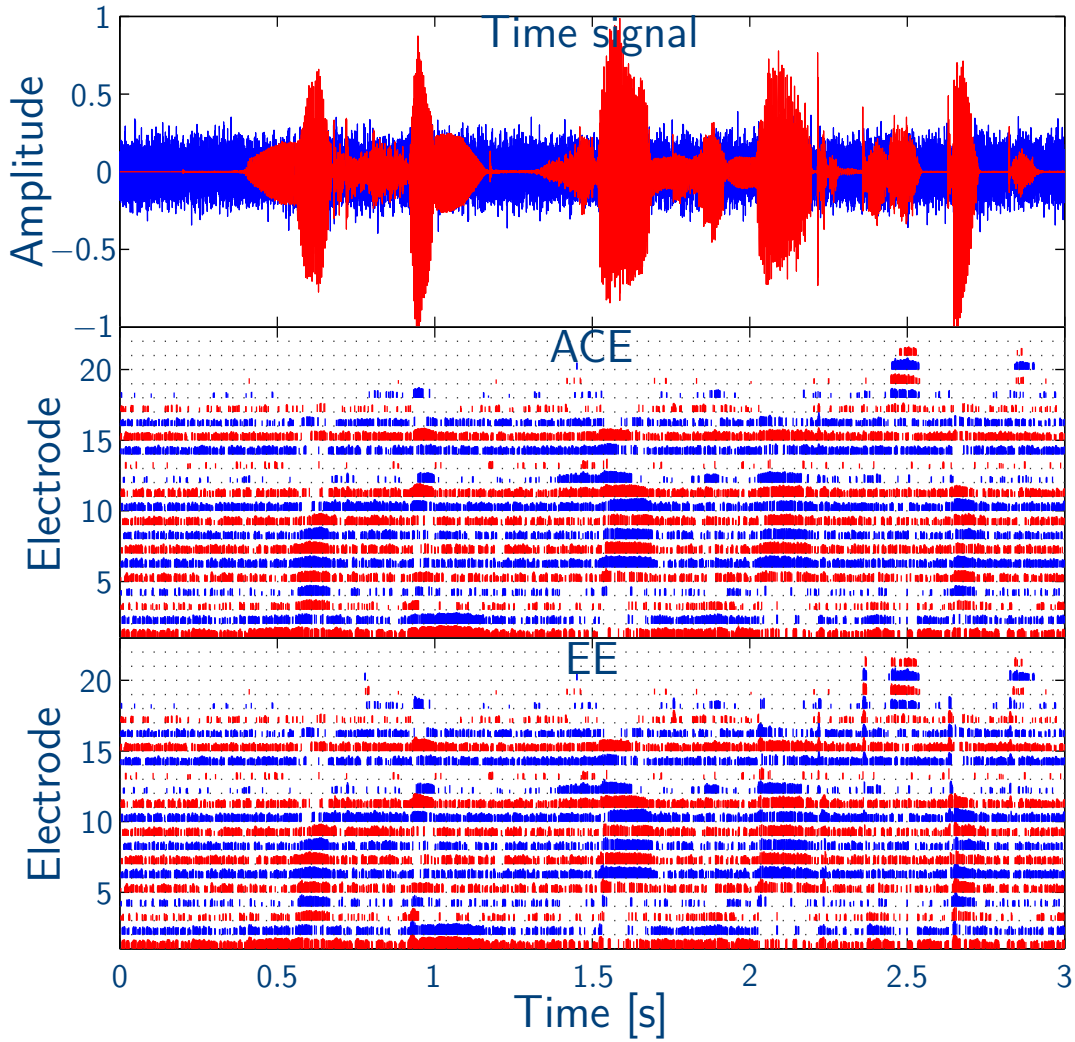


"Morgen gaan wij naar de stad" (Tomorrow, we are going to the city)

- Time signal (upper)
- ACE strategy (middle)
- EE strategy (bottom)

Figure: Elecrodogram of the ACE and the EE strategy.

EE in stationary speech shaped noise

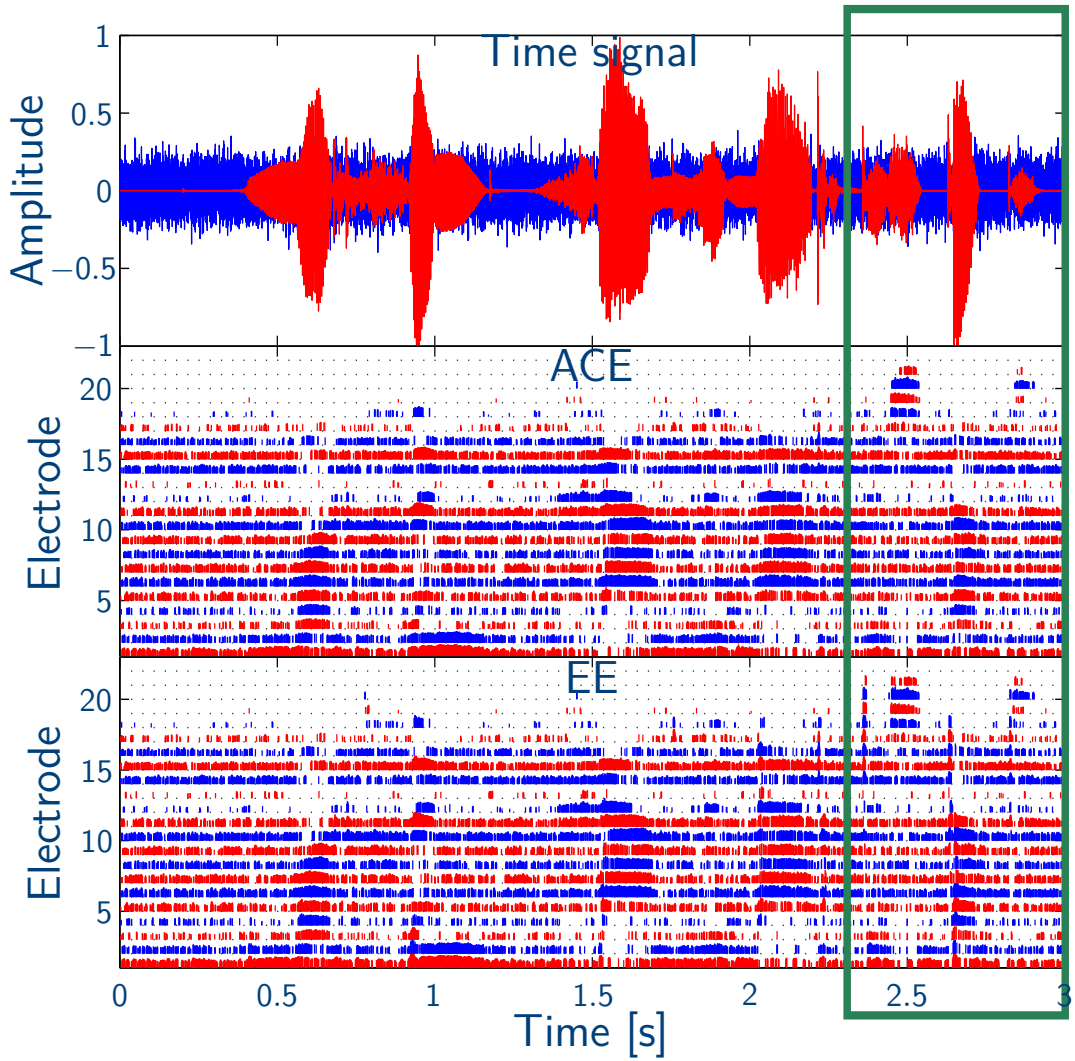


"Morgen gaan wij naar de stad" (Tomorrow, we are going to the city)

- 3 dB SNR
- Time signal (upper)
- ACE strategy (middle)
- EE strategy (bottom)

Figure: Elecrodogram of the ACE and the EE strategy.

EE in stationary speech shaped noise

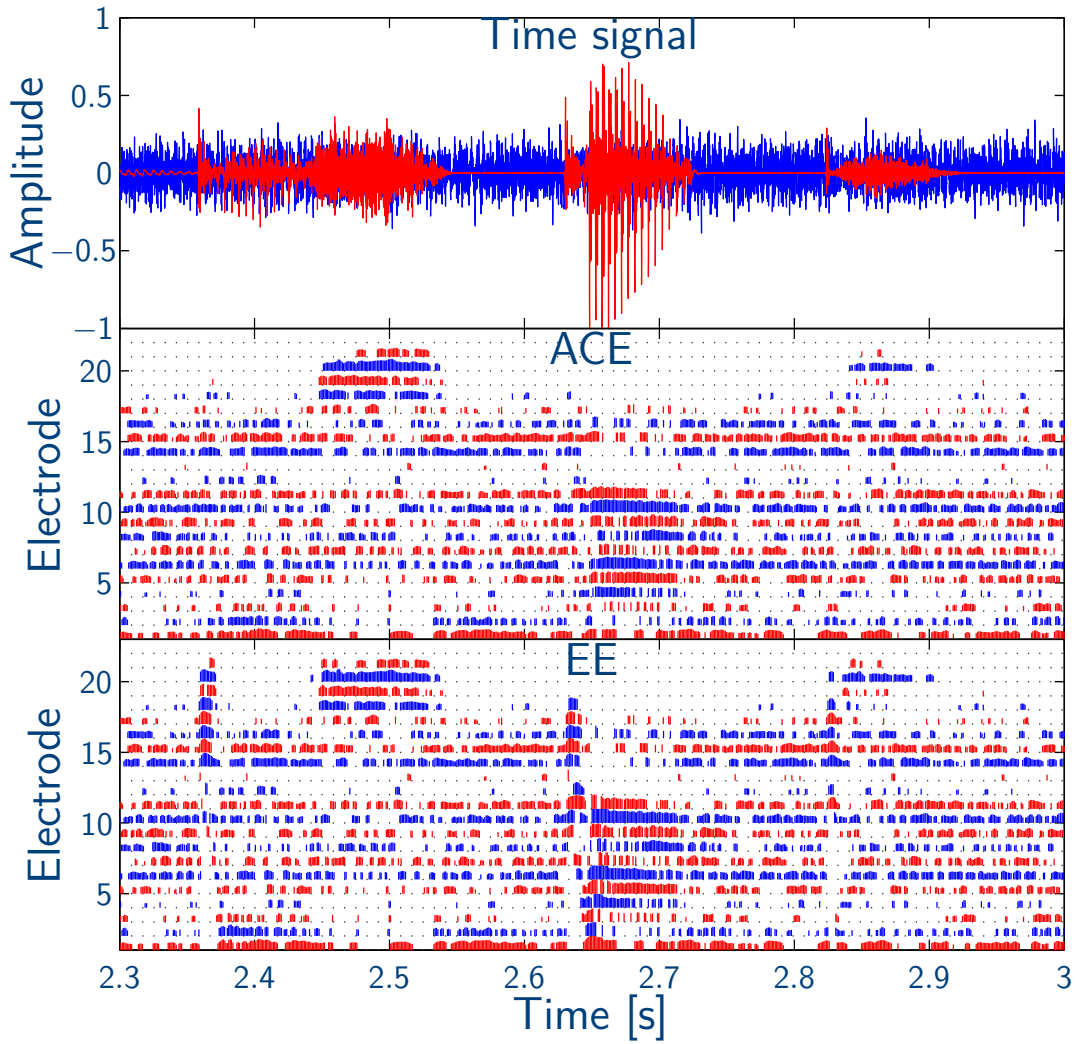


"Morgen gaan wij naar de stad" (Tomorrow, we are going to the city)

- 3 dB SNR
- Time signal (upper)
- ACE strategy (middle)
- EE strategy (bottom)

Figure: Elecrodogram of the ACE and the EE strategy.

EE in stationary speech shaped noise



"de stad" (the city)

- 3 dB SNR
- Time signal (upper)
- ACE strategy (middle)
- EE strategy (bottom)

Figure: Elecrodogram of the ACE and the EE strategy.

Outline

SI of CI users with EE vs. ACE in SSN

Methods

- 6 CI users
- Leuven Intelligibility Sentence Test (LIST) sentences (Van Wieringen & Wouters, 2008): Male and female sentences
- Stationary speech shaped noise (SSN)
- Peak extraction from the clean speech signal

Procedure

- Speech level constant at 65 dB SPL
- Sentences in quiet (Q) and at 10, 6, 2, and -2 dB SNR
- Direct presentation (L34 research device by Cochlear, Ltd.)
- Keyword percent correct scores

Results (1)

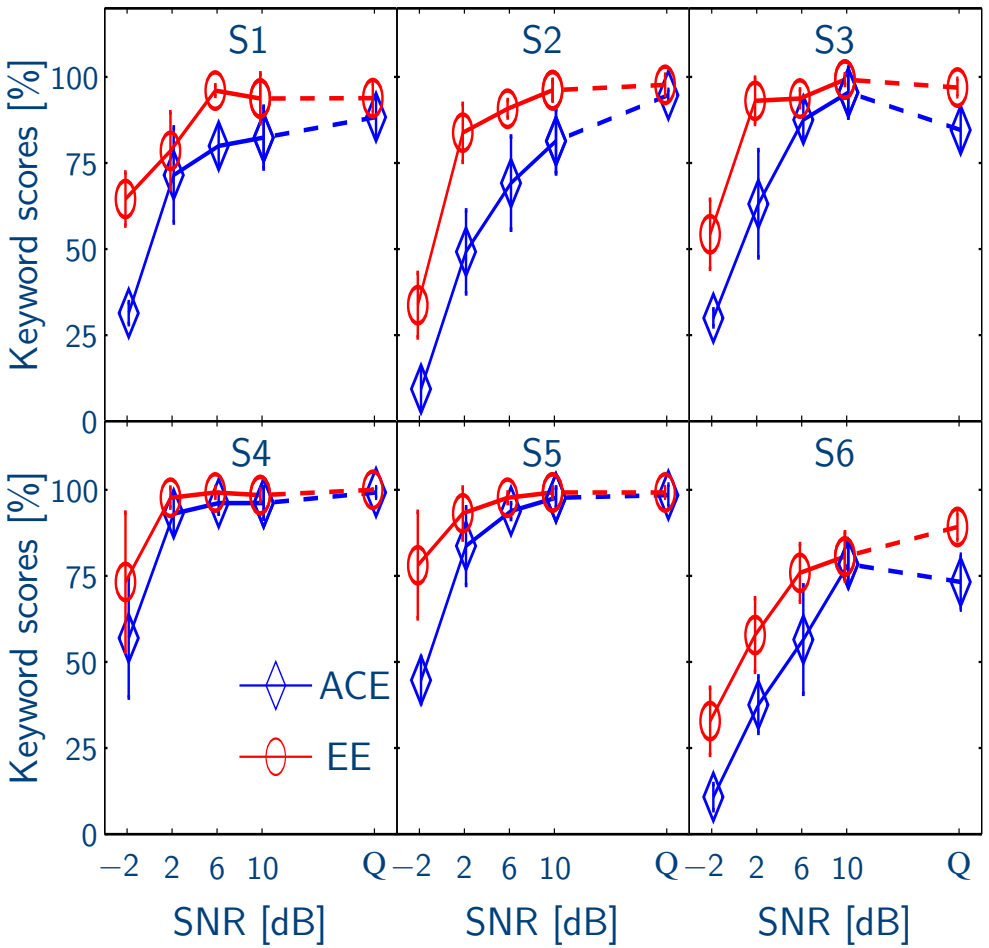


Figure: Results of the speech recognition task in SSN for all 6 CI user. Error bars depict the standard deviation.

- Transformation of the scores to rationalized arcsine units before repeated measures ANOVA analysis.
- EE performed significantly better than ACE.
- CI users showed immediate benefit of onset enhancement.

Results (2)

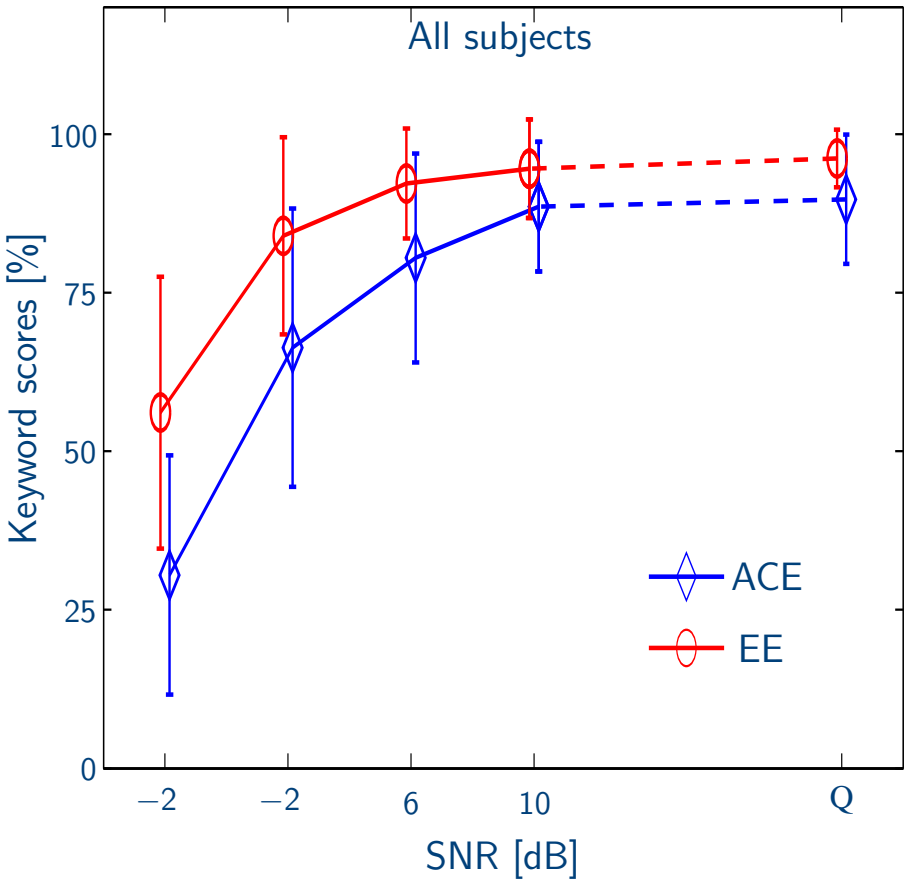


Figure: Mean results of all CI users of the speech recognition task in SSN. Error bars depict the standard deviation.

- Transformation of the scores to rationalized arcsine units before repeated measures ANOVA analysis.
- EE performed significantly better than ACE.
- CI users showed immediate benefit of onset enhancement.

SI with EE vs. ACE: Competing talker

Methods

- 4 CI users
- Target speaker (T): Female LIST sentences
- Interfering speaker (I): Male LIST sentences
- Four conditions
  - T+I: ACE (reference)
  - EE(T)+EE(I): Onsets of both speakers enhanced
  - EE(T+I): Onsets of the noisy mixture enhanced
  - EE(T)+I: Onsets of the target speaker enhanced

Procedure

- Target speaker level constant at 65 dB SPL
- Adaptive SRT determination with 2 dB SNR stepsize

Example: EE with two talkers

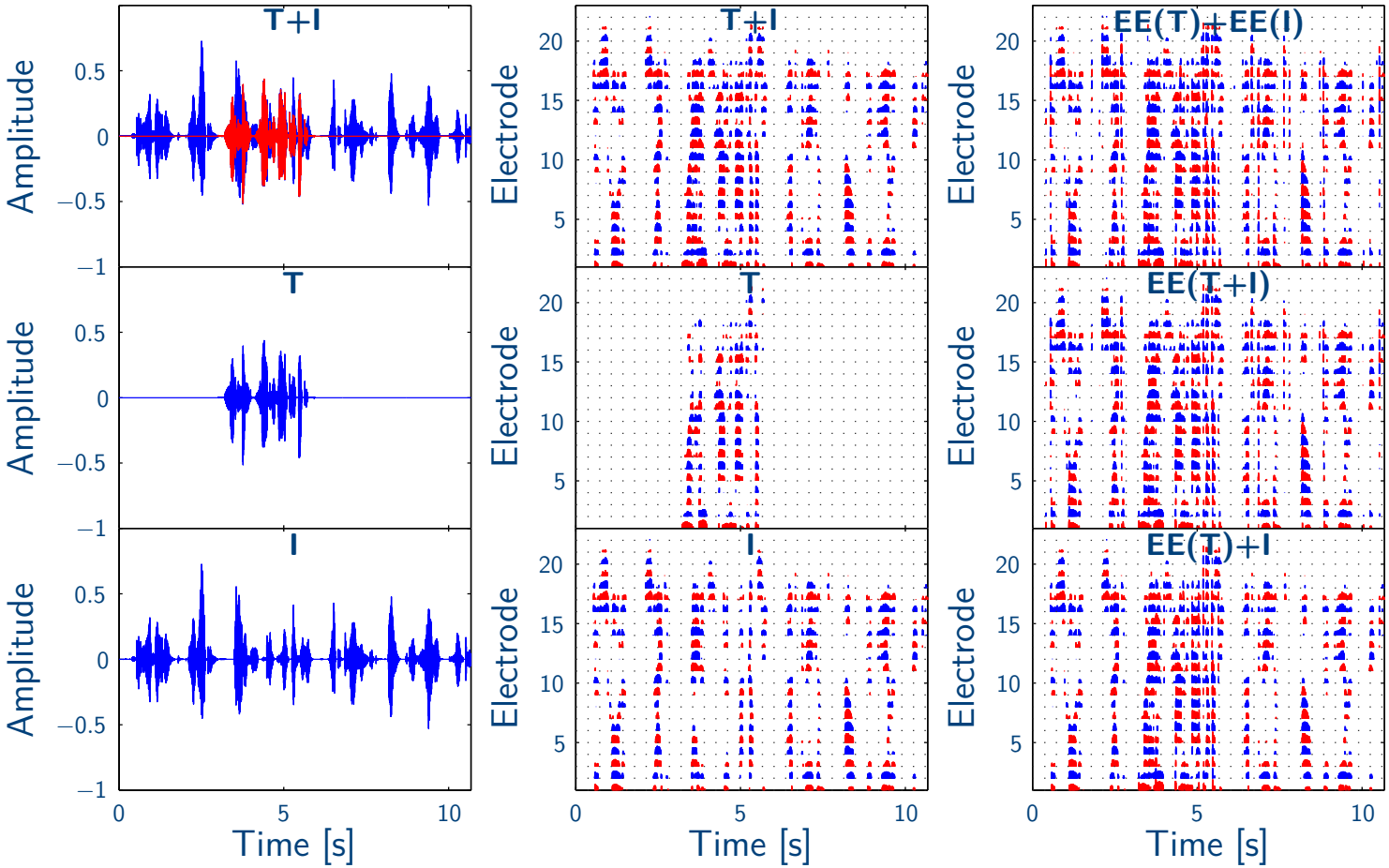
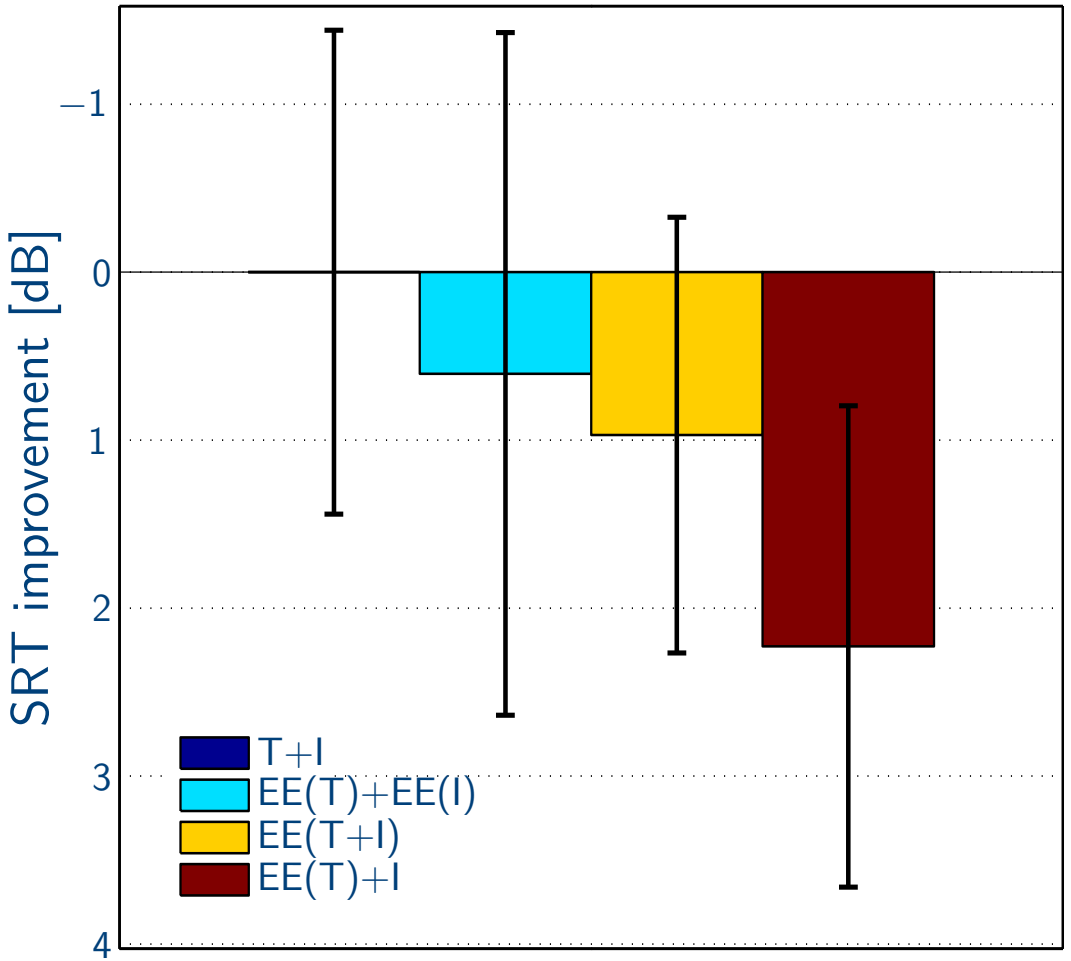


Figure: Example of the four different tested conditions in the competing talker scenario.

Results



- All 4 CI users show significant SRT improvement for EE(T)+I
- Variation across subject
- EE(T+I) can be implemented without noise reduction

Figure: Results of the adaptive SRT determination of the CI users in the competing talker scenario.

Loudness rating

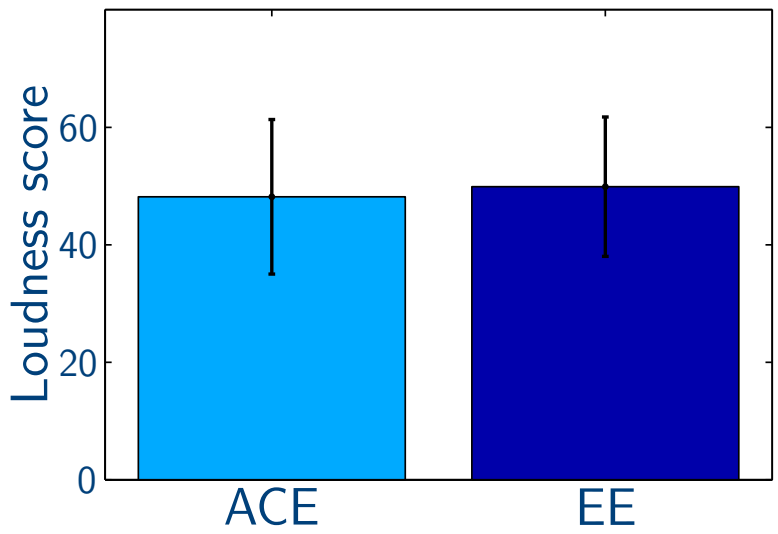


Figure: Results of the loudness rating experiment with 6 CI users of one list of the LIST sentences in quiet. The error bars depict the standard deviation of the scores.

Loudness rating of one list of the LIST sentences (10 sentences) in quiet at 65 dB SPL.

- No significant difference.
- Loudness perception not affected by the processing.

Outline

The potential of EE in auditory prostheses

- Onset enhancement can lead to better speech perception in noise and with interfering speaker.
  - Potential was demonstrated in CI users.
  - Immediate benefit for all tested subjects.
  - Results valid for different sets of target speech material.
- Results of the interfering talker scenario suggest that the "brain can do the rest" when onset enhancement is done on the noisy mixture.
- Cues to onset features help for segregation, cfr. temporal coherence across frequency

Appendix

Bibliography

Chen, F., and Loizou, P. C. (**2012**). "Contributions of cochlea-scaled entropy and consonant-vowel boundaries to prediction of speech intelligibility in noise," J. Acoust. Soc. Am. **131**, 4104–4113.

Delgutte, B. (**1999**). Chapter in *The Handbook of Phonetic Sciences* (Blackwell Publishers, Oxford), 507-538.

Geurts, L., and Wouters, J. (**1999**). "Enhancing the speech envelope of continuous interleaved sampling processors for cochlear implants," J. Acoust. Soc. Am. **105**, 2476–2484.

Koning, R., and Wouters, J. (**2012**). "The potential of onset enhancement for increased speech intelligibility in auditory prostheses," J. Acoust. Soc. Am. **132**, 2569-2581.

Lee, J. H., and Kewley-Port, D. (**2009**). "Intelligibility of interrupted sentences at subsegmental levels in young normal-hearing and elderly hearing-impaired listeners," J. Acoust. Soc. Am. **125**, 1153–1163.

Lewicki, M. S. (**2010**). "A signal take on speech," Nature **466**, 821–822.

Stilp, C. E., and Kluender, K. R. (**2010**). "Cochlea-scaled entropy, not consonants, vowels, or time, best predicts speech intelligibility," Proc. Natl. Acad. Sci. USA **107**, 12387–12392.

Vandali, A. E. (**2001**). "Emphasis of short-duration acoustic speech cues for cochlear implant users," J. Acoust. Soc. Am. **109**, 2049–2061.

Van Wieringen, A., and Wouters, L. (**2008**). "LIST and LINT: Sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands," Int. J. Audiol. **47**, 348–355.

Appendix

Acknowledgement

- All subjects
- AUDIS, Marie Curie ITN
- IWT
- Cochlear, Ltd.