IMPROVING CHANNEL SELECTION OF SOUND CODING ALGORITHMS IN COCHLEAR IMPLANTS

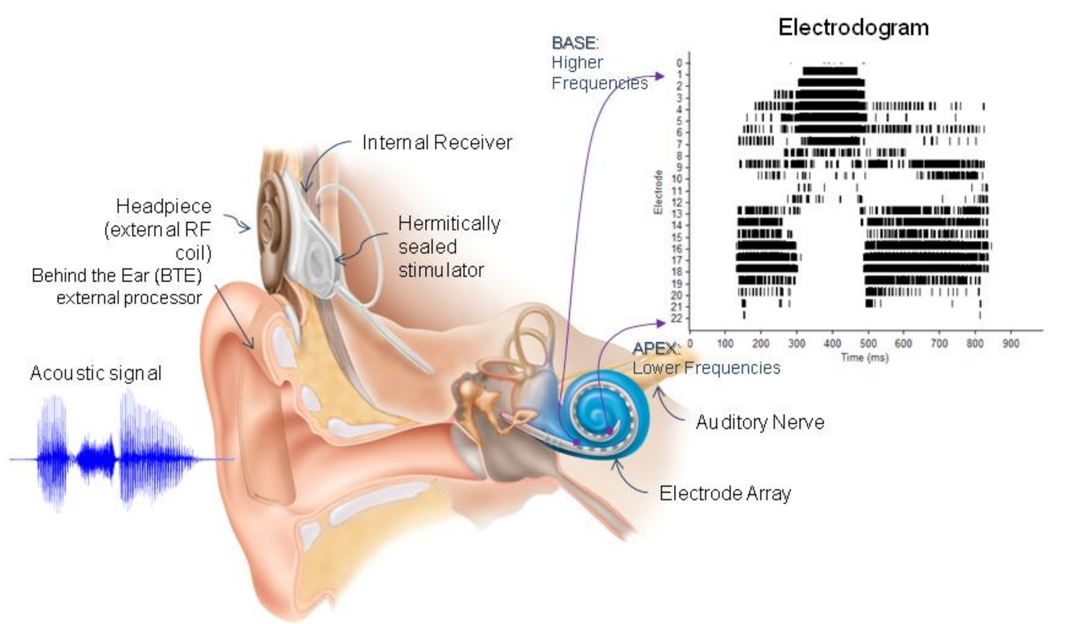
Hussnain Ali, Feng Hong, John H. L. Hansen, and Emily Tobey

hussnain.ali, feng.hong, john.hansen, etobey @utdallas.edu

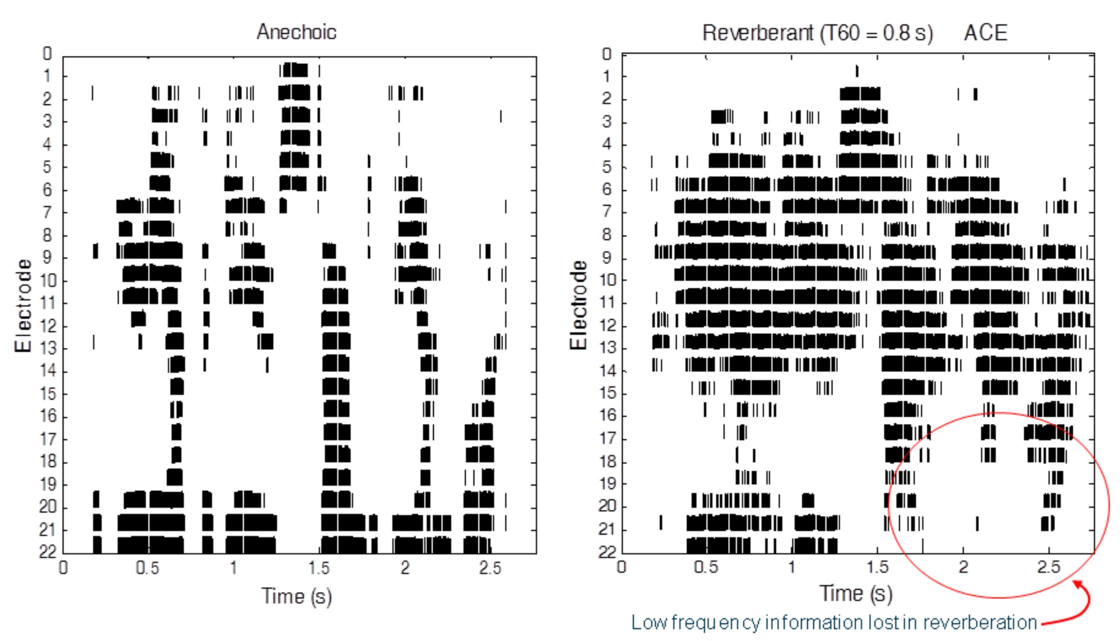
Center for Robust Speech Systems (CRSS)
Erik Jonsson School of Engineering & Computer Science
The University of Texas at Dallas
Richardson, Texas 75080-3021, U.S.A.

ICASSP 2014 IEEE International Conference on Acoustics, Speech and Signal Processing
Florence, Italy, May 4-9, 2014
Firenze Fiera / Fortezza da Basso

1 Cochlear Implant – Functional Diagram

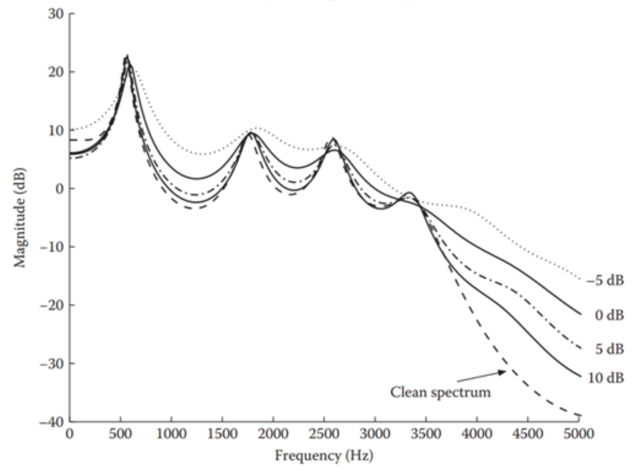


2 Sound Processing: *n-of-m* Strategies Clean Signal *vs* Distorted Signal



3 Channel Selection of Formants in Noise

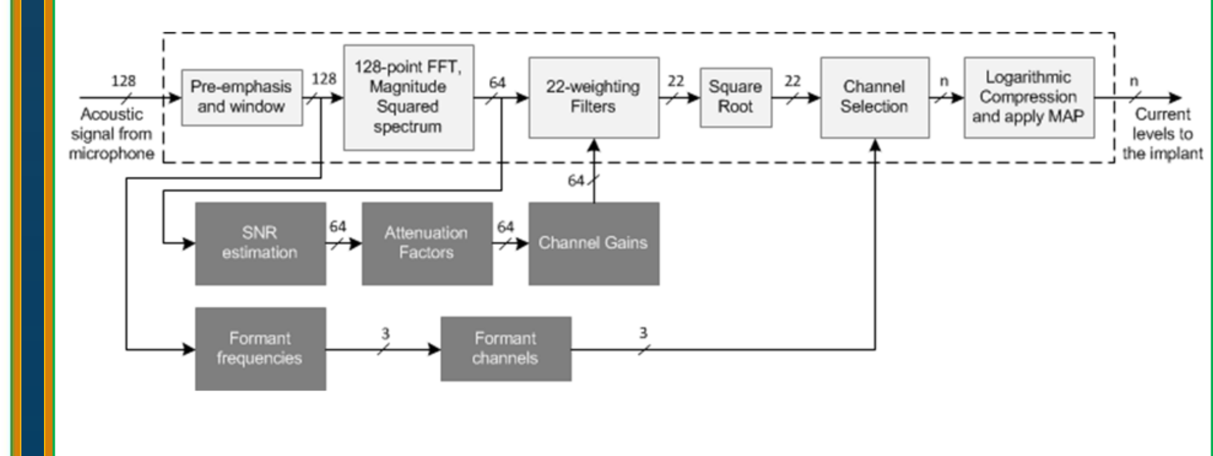
- Channel selection in ACE is based on the largest filter amplitudes.
- Several maxima may come from a single peak.
- Spectral smearing and reduced spectral contrast in noise.
- Noise dominant channels may be given preference in noise.



LPC spectra of the vowel /eh/ corrupted by multitalker babble at -5 to 10 dB SNR

Ref. Philipos C. Loizou, Speech Enhancement – Theory and Practice, CRC Press, 2013

4 ACE Coding Algorithm with Proposed Channel Selection Enhancement



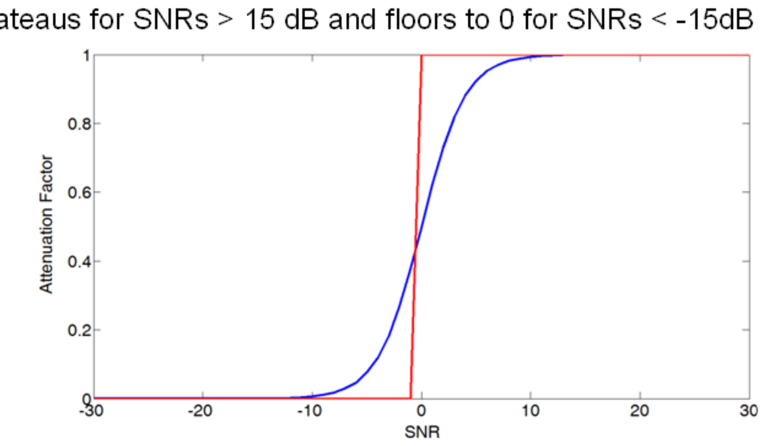
Signal flow in the standard ACE strategy (shown inside the dotted block). Processing blocks for the proposed technique are shown in the darker tone. Numbers on connecting arrows represent the frame size in number of samples at each step.

5 SNR Estimation

- ACE strategy: For every 8 ms analysis window, 128-point FFT is computed thereby generating 64 frequency bins.
- Estimate instantaneous SNR of each time frequency (TF) unit $X(i,j)$, *where X is magnitude squared spectrum of the i^{th} analysis frame and j^{th} frequency bin.*
- SNR estimation using improved minimum controlled recursive average (IMCRA) algorithm.
- Generate attenuation factor based on the estimated SNR using binary or soft masking.
- Assign weights to each TF unit every stimulation cycle.

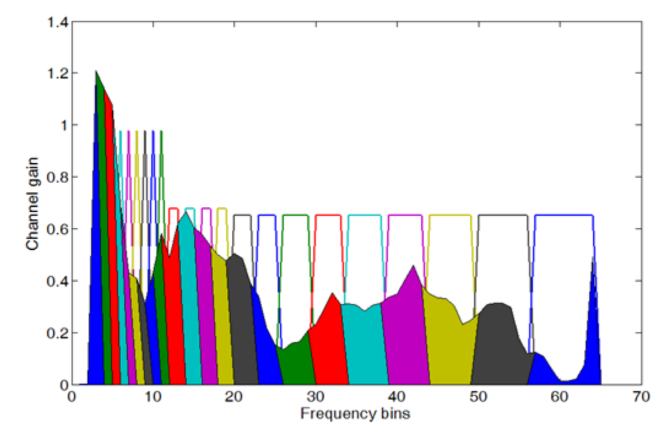
6 Masking Function

- Binary weighted
  - Step function
  - Assign a weight of 0 if $SNR(i,j)$ < 0 dB, for the rest (i.e., $SNR(i,j) \geq 0$ dB), assign a binary value of 1.
- Soft masking
  - Sigmoidal-shaped function
  - plateaus for SNRs > 15 dB and floors to 0 for SNRs < -15dB



7 Proposed Technique

- **Assign priority to channels** corresponding to the formant frequencies
  - Continuously track the location of formant frequencies
  - (F1, F2, and F3) peaks computed by solving roots of LPC
- **Shape the weighting functions** according to the estimated SNR
  - Each TF unit is assigned a weight based on a binary or soft masking attenuation function



8 Experimental Protocol– Experiment 1

- Evaluation of speech intelligibility and speech perception quality
- **Subjects:** 3 post-lingually deaf adult CI subjects with Nucleus 24 device from Cochlear Corp.
- **Speech Material:** IEEE sentences.
- Experiments implemented offline in MATLAB and stimuli presented via UT Dallas's PDA-based research platform.

Experiment 1: Effectiveness of assigning priority to formant channels

- speech in quiet,
- speech in speech shaped noise (SSN), SNR = 10 dB,
- speech in SSN, SNR = 5 dB,
- speech in white Gaussian noise (WGN), SNR = 10 dB,
- speech in reverberation with reverberation time T_{60} = 600 ms,
- speech in reverberation (T_{60} = 600 ms) and SSN, SNR = 10dB.

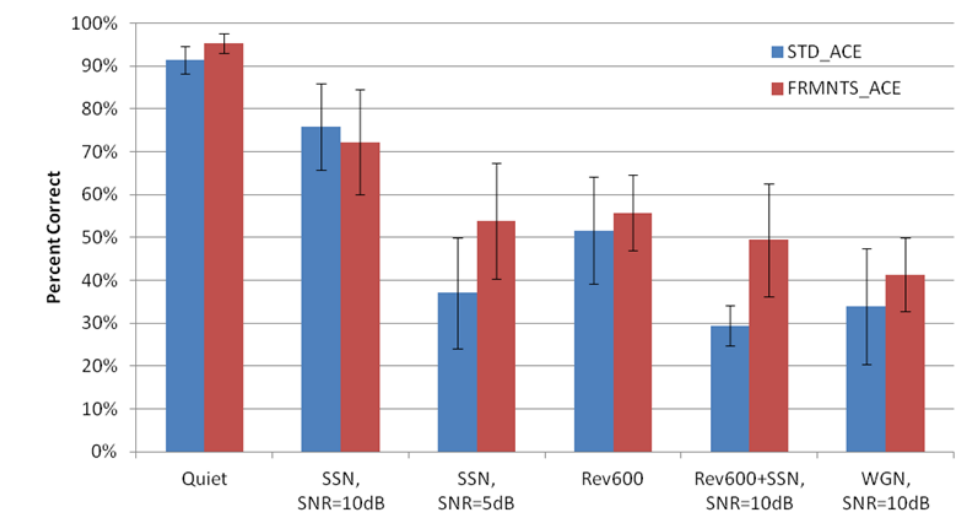
9 Experimental Protocol – Experiment 2

- Evaluation of speech intelligibility and speech perception quality

Experiment 2: Assessment of assigning weights to each TF unit of the weighting filters.

- Techniques:
  - Ideal-Binary (IdBinary),
  - Ideal-Soft (IdSoft),
  - Estimated-binary (EsBinary),
  - Estimated-soft (EsSoft).
- Conditions:
  - SSN, SNR = 10 dB,
  - SSN, SNR = 5 dB.

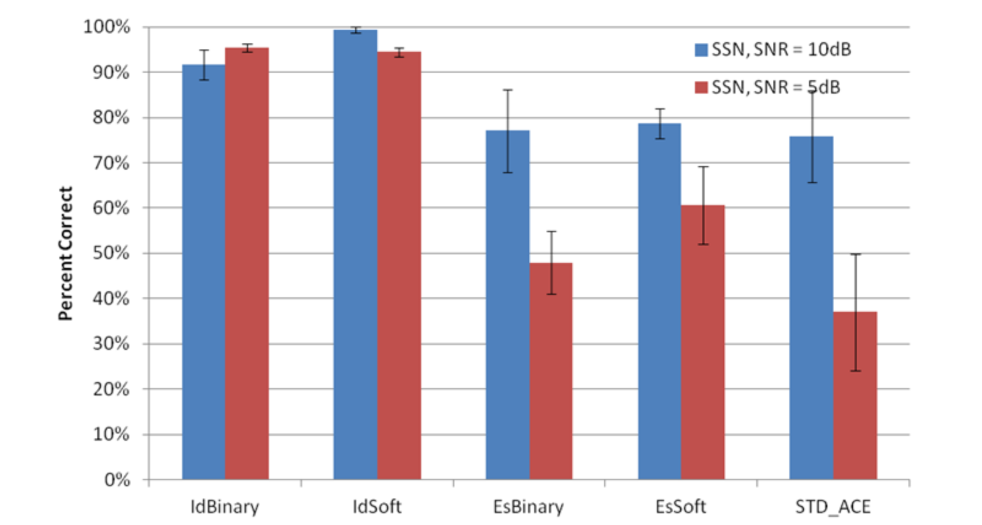
10 Results – Experiment 1



Mean speech intelligibility scores of 3 CI users in Experiment 1. Error bars represent standard error of the mean.

➡ The proposed FRMNTS_ACE technique on average provides **+8%** improvement in speech intelligibility as compared to the STD_ACE.

11 Results – Experiment 2



. Comparison of speech intelligibility scores from the proposed technique using IdBinary, IdSoft, EsBinary, and EsSoft approaches with STD_ACE strategy. Error bars represent SEM.

➡ The proposed gain shaping technique provides **+19%** (Id, SNR = 10dB), **+58%** (Id, SNR = 5dB), **+2%** (Es, SNR=10dB), and **+17%** (Es, SNR = 5dB) improvement in speech intelligibility as compared to the STD_ACE program.

12 Conclusions

- Two approaches considered to improve channel selection process of spectral maxima sound coding algorithms for cochlear implants
  1. Assign priority to channels corresponding to formant frequencies
  2. Adaptively assign weights to each time-frequency unit based on the estimated SNR
- **Experiment 1:**
  - Speech Intelligibility: Significant improvement with FRMNTS_ACE approach as compared to the STD_ACE at low SNR levels.
  - Perception Quality: Slight preference to FRMNTS_ACE over STD_ACE.
- **Experiment 2:**
  - Speech Intelligibility: Significant improvement at low SNR levels on shaping the gains of weighting filters based on the estimated noise.
  - Perception Quality: High preference to noise-estimation and soft masking approach over STD_ACE.

This work was supported by the grant R01 DC010494-01A from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

ICASSP 2014 IEEE International Conference on Acoustics, Speech and Signal Processing

Florence, Italy, May 4 – 9, 2014