

W9: LOMBARD EFFECT PERTURBATION PRE-PROCESSING STRATEGY FOR COCHLEAR IMPLANT USERS

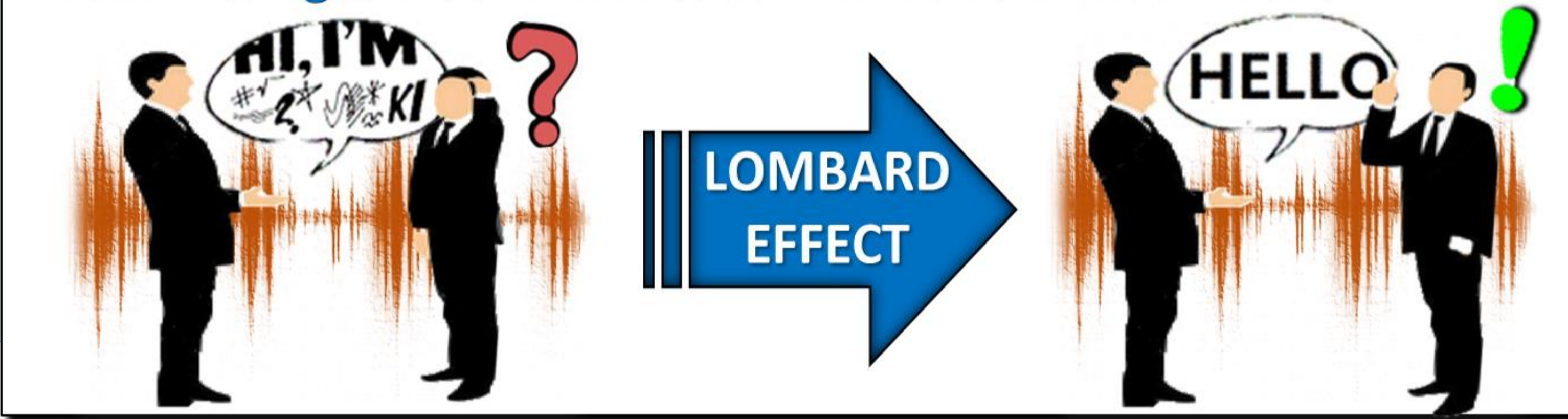
Jaewook Lee, Hussnain Ali, John H.L. Hansen

Center for Robust Speech Systems - Cochlear Implant Processing Lab (CRSS-CIL)
 Department of Electrical & Computer Engineering, The University of Texas at Dallas, Richardson, TX, US
 (jaewook, hussnain.ali, john.hansen)@utdallas.edu

1. Introduction

- ◆ **CI under Noise** - Speech intelligibility of cochlear implant (CI) users degrades significantly in noise.
- ◆ **Lombard Effect** - a type of stressed speech introduced by human exposure to noise; causes an increase in vocal intensity, F0, spectral slope, speech rate, etc. (Lombard,1991; Hansen,1996).
- ◆ **Previous Study** – Lombard effect was found in the speech of CI users during voice communication in noisy listening environments (Lee *et al.*, 2015).

FIG1: Intelligible Communication in Noise Environment



- ◆ **Goal 1:** Develop a speech enhancement algorithm based on Lombard effect for CI users.
- ◆ **Goal 2:** Examine how CI users perceive the neutral-to-Lombard processed speech in noisy environments.

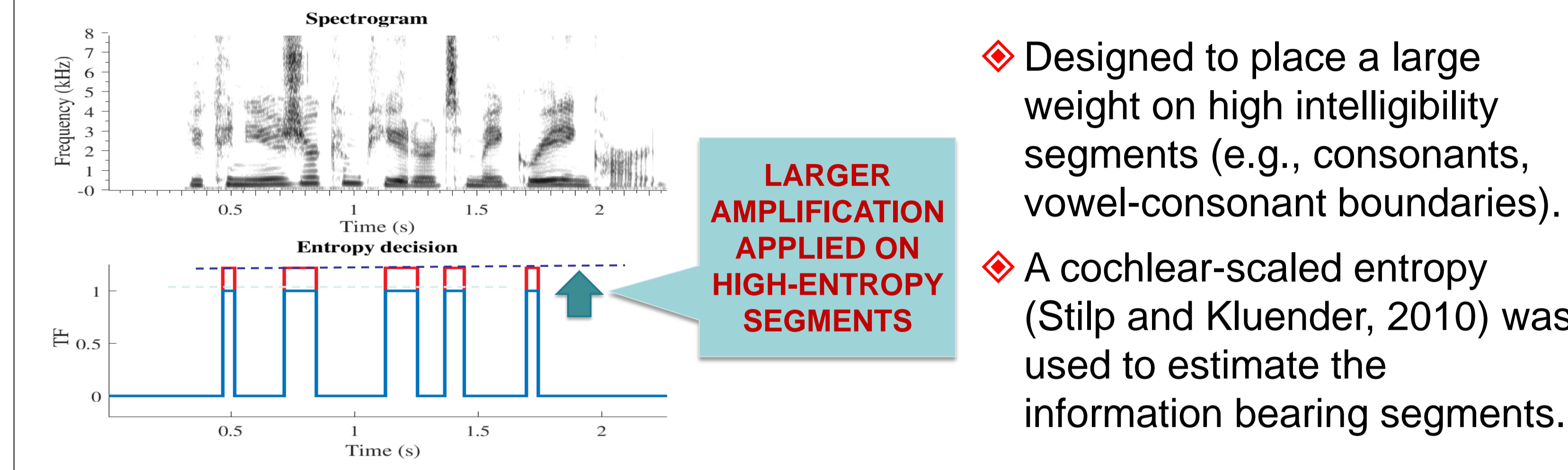
2. Methods

Dataset for Parameter Modeling

- ◆ 2 normal hearing (NH) speakers produced AzBio sentences (Spar *et al.*, 2012) in 2-way conversation.
- ◆ Large-crowd noise at 90 dB SPL presented via open-air headphones.

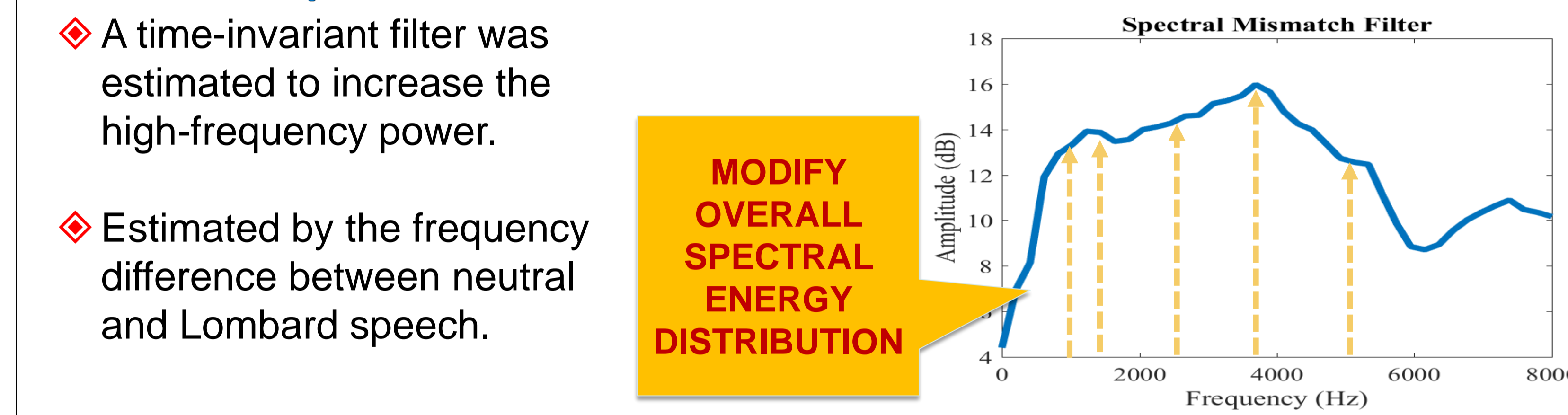


Feature 1: Entropy-Based Temporal Amplification



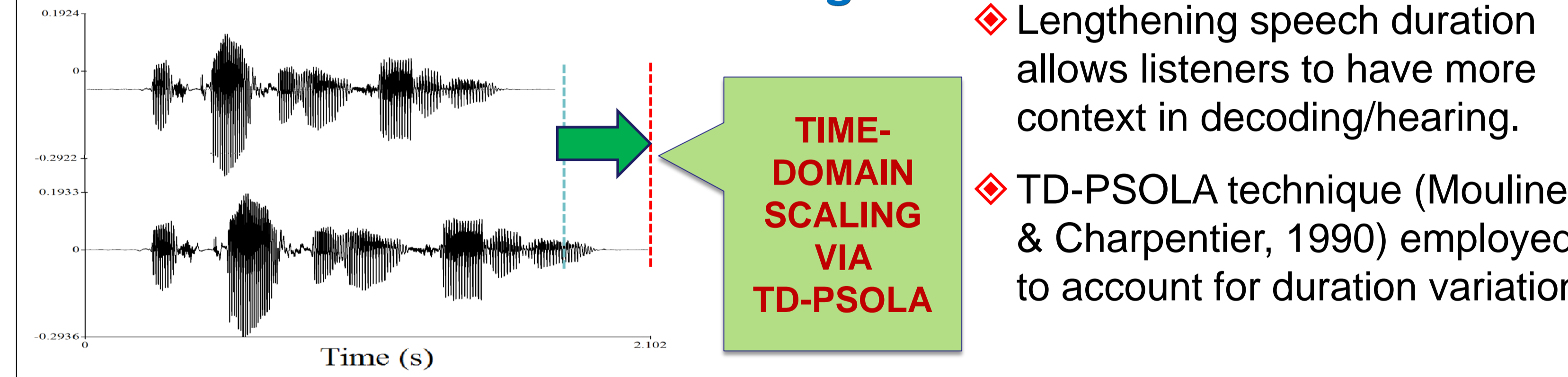
- ◆ Designed to place a large weight on high intelligibility segments (e.g., consonants, vowel-consonant boundaries).
- ◆ A cochlear-scaled entropy (Stilp and Kluender, 2010) was used to estimate the information bearing segments.

Feature 2: Spectral Contour Transformation



- ◆ A time-invariant filter was estimated to increase the high-frequency power.
- ◆ Estimated by the frequency difference between neutral and Lombard speech.

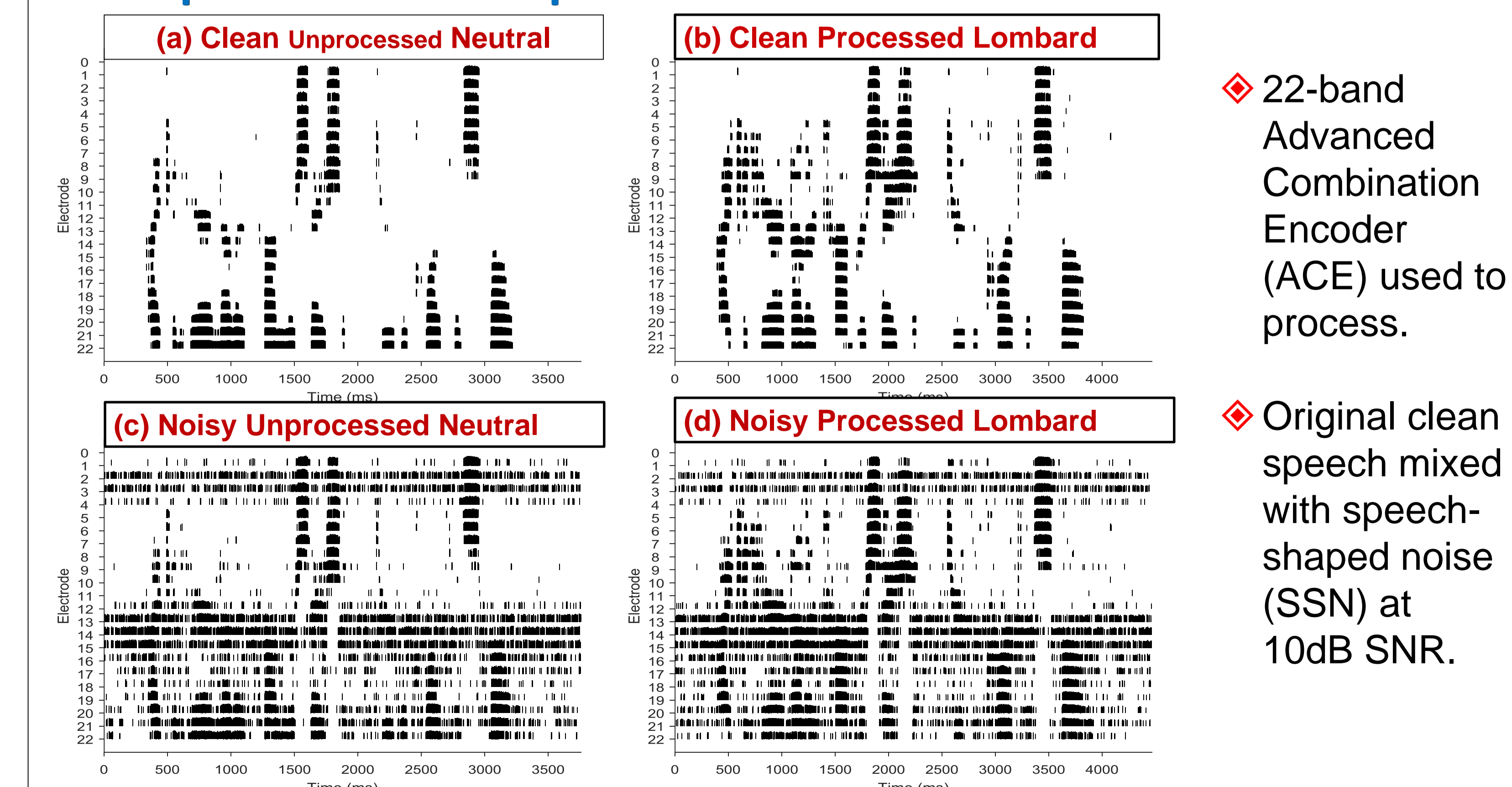
Feature 3: Uniform Time-Stretching



- ◆ Lengthening speech duration allows listeners to have more context in decoding/hearing.
- ◆ TD-PSOLA technique (Moulines & Charpentier, 1990) employed to account for duration variation.

3. Results

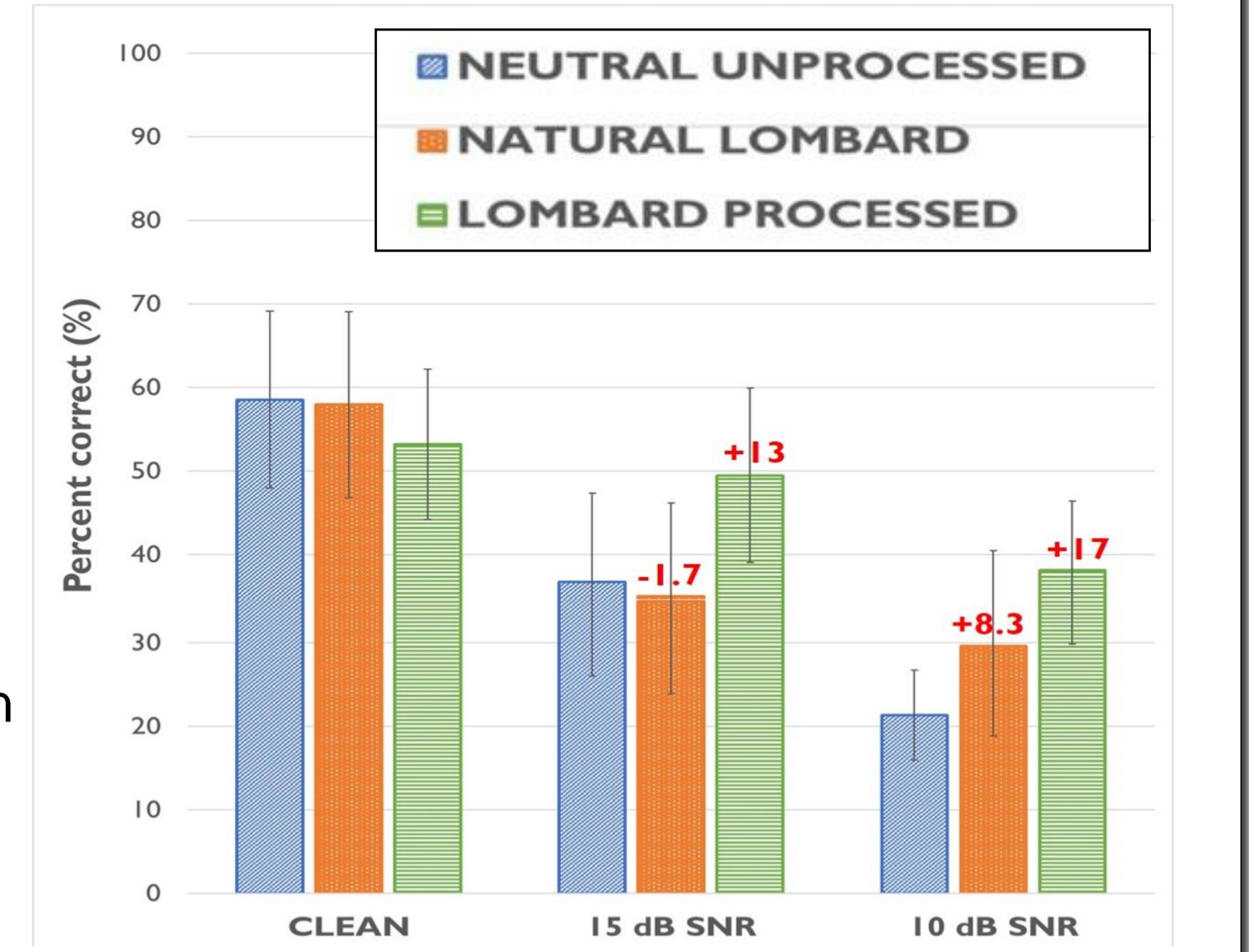
Example Stimulus Output Patterns



- ◆ 22-band Advanced Combination Encoder (ACE) used to process.
- ◆ Original clean speech mixed with speech-shaped noise (SSN) at 10dB SNR.

Average Word Recognition Score with 5 CI users

- ◆ Five CI subjects (mean age of 64) listened to original stimuli in quiet and corrupted by large-crowd noise.
- ◆ Speech recognition scores were calculated based on number of words correctly identified.



4. Summary & Discussion

- ◆ A new speech modification criterion based on Lombard effect is proposed.
- ◆ Perceptual experiment of Lombard processed speech was performed with five CI users.
- ◆ Improvement in intelligibility was found with Lombard processed speech, particularly in noisy environments.
- ◆ The modification of speech using the proposed algorithm might contribute to higher intelligibility.
- ◆ The results suggest potential of the Lombard effect-based speech enhancement algorithms for CI users.

5. References

- ◆ Lombard (1911) "Le signe de l'elevation de la voix [the sign of voice raising]", *Annals des Maladies de l'Oreille et du Larynx*, pp. 101-119
- ◆ Hansen (1996) "Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition", *Speech Communication* 20, 151-173
- ◆ Lee *et al.* (2015) "Analysis of speech language communication for cochlear implant users in noisy Lombard conditions", in *Proc. IEEE ICASSP*, Brisbane, Australia