A Novel CI Coding Strategy Based on a Cochlear Model and Deep Learning

Maryam Hosseini[1,2], Tim Brochier[2], Zachary Smith[2], Brett Swanson[2], Andrew Vandali[2], Alan Kan[1], Fadwa Alnafjan[1], Kat Fernandez[1], Richard F. Lyon[3]

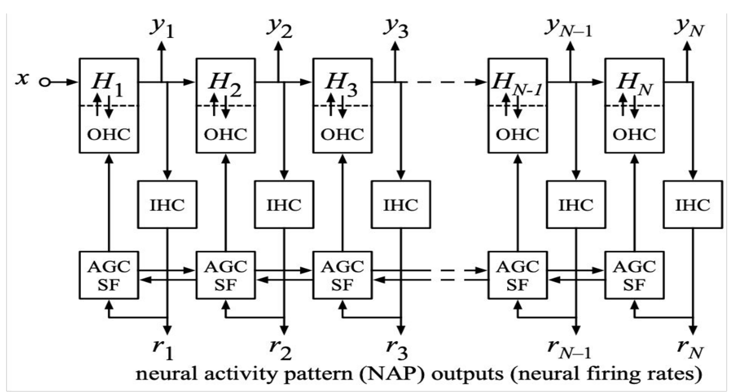
1. Department of Linguistics, Macquarie University 2. Cochlear Ltd 3. Google Research Australia

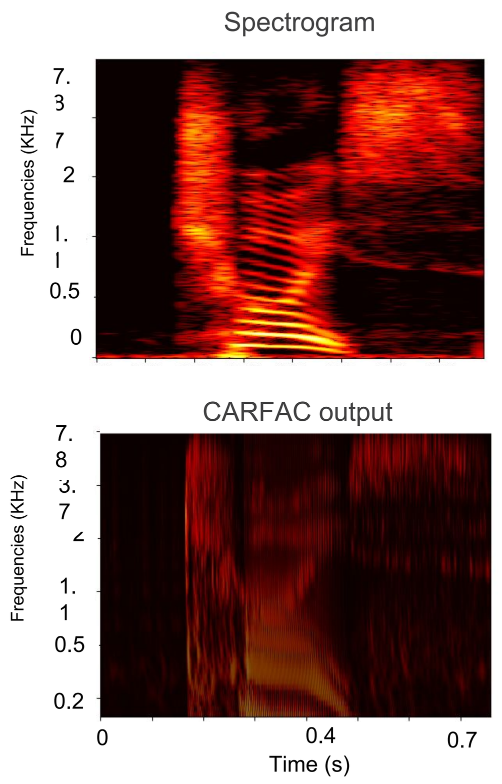
Introduction

- Models of the cochlea generate potentially useful features for speech processing such as **harmonic modulations**, **onset enhancement**, and **travelling wave delay**.
- Our goal is to design a **biomimetic sound coding strategy,** based on the CARFAC model of the cochlea, to generate CI neural activity patterns that closely resemble normal hearing neural activity.

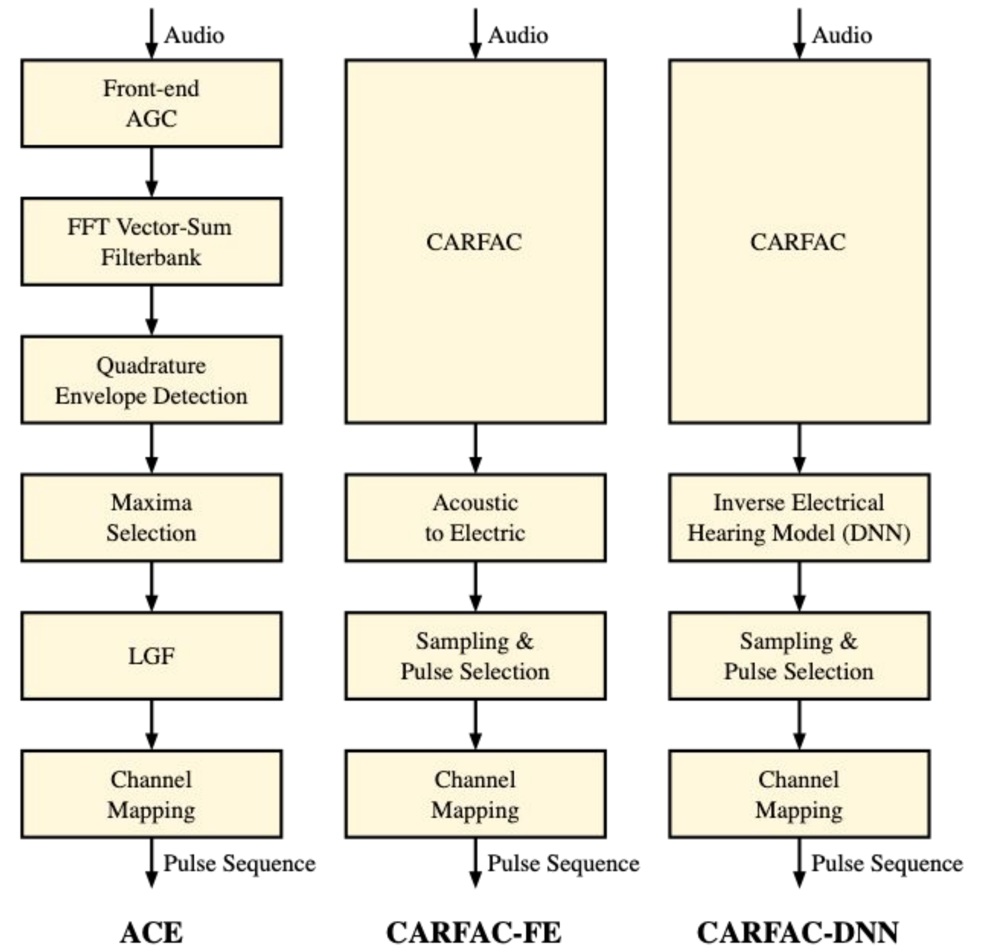
Methods

Cochlear model : CARFAC

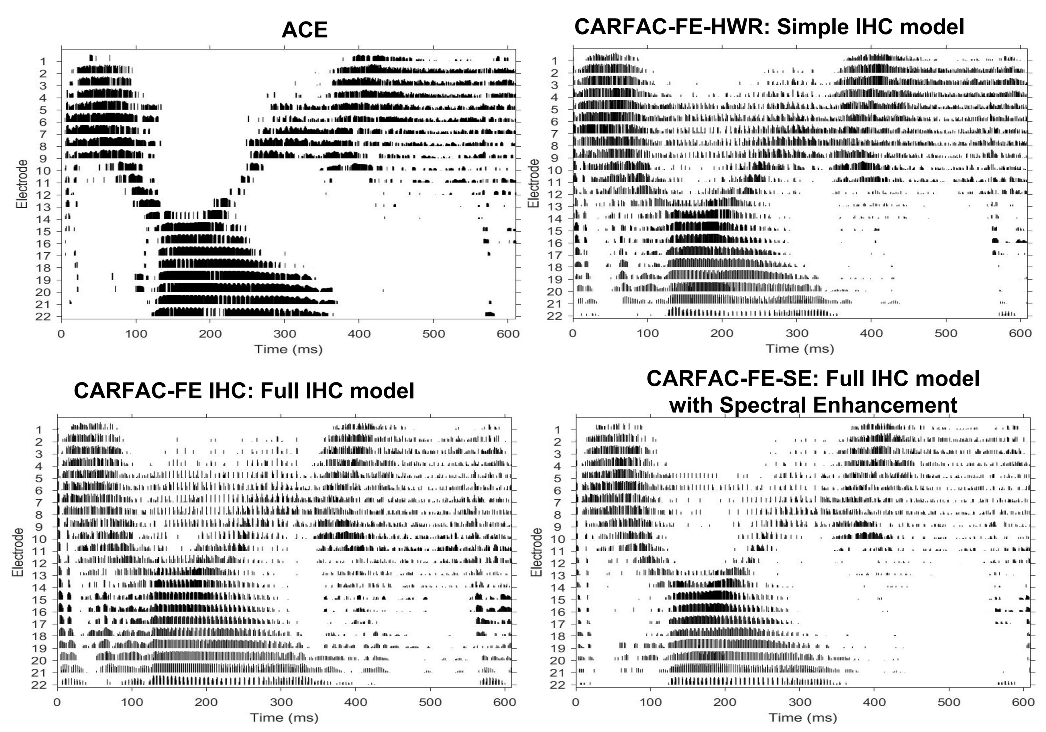




Sound processing



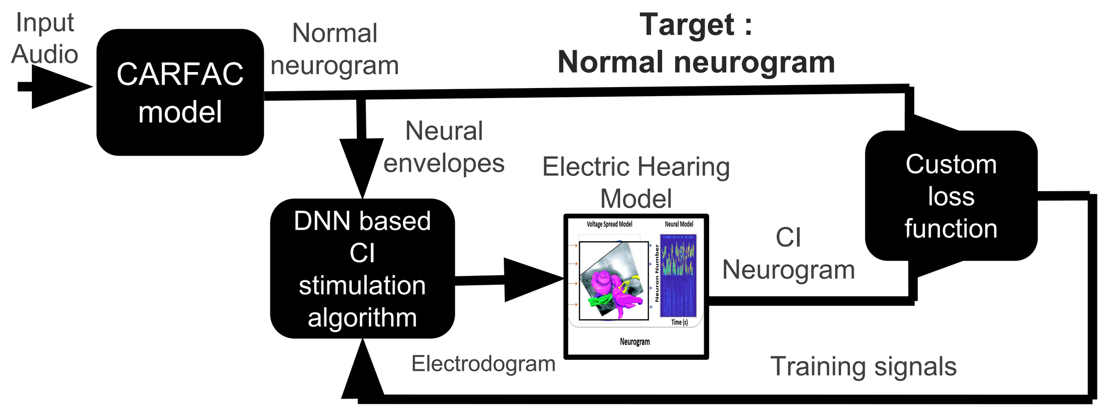
Replacing ACE front-end filterbank with a Cochlear model – CARFAC Front-end strategies



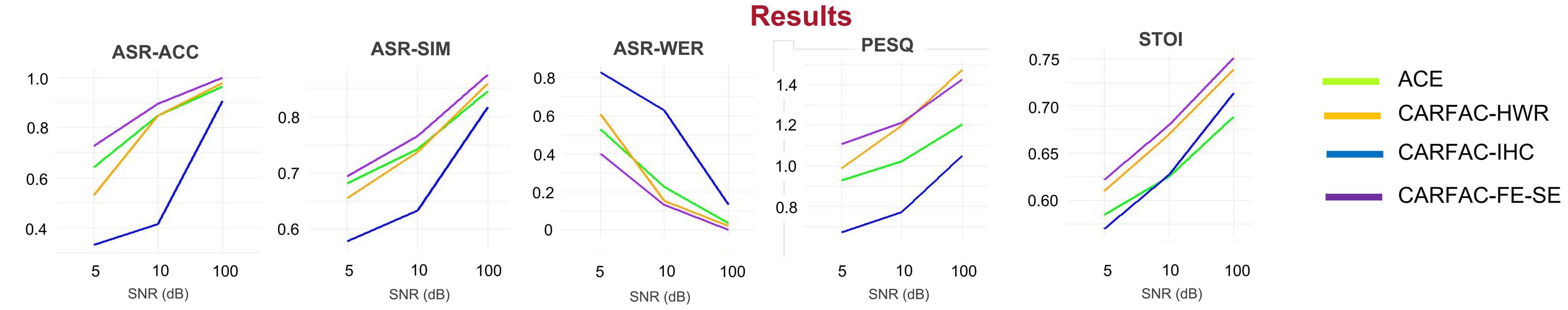
Electrodograms of the word “Choice” for experimental processing strategies

DNN-based Sound Processing

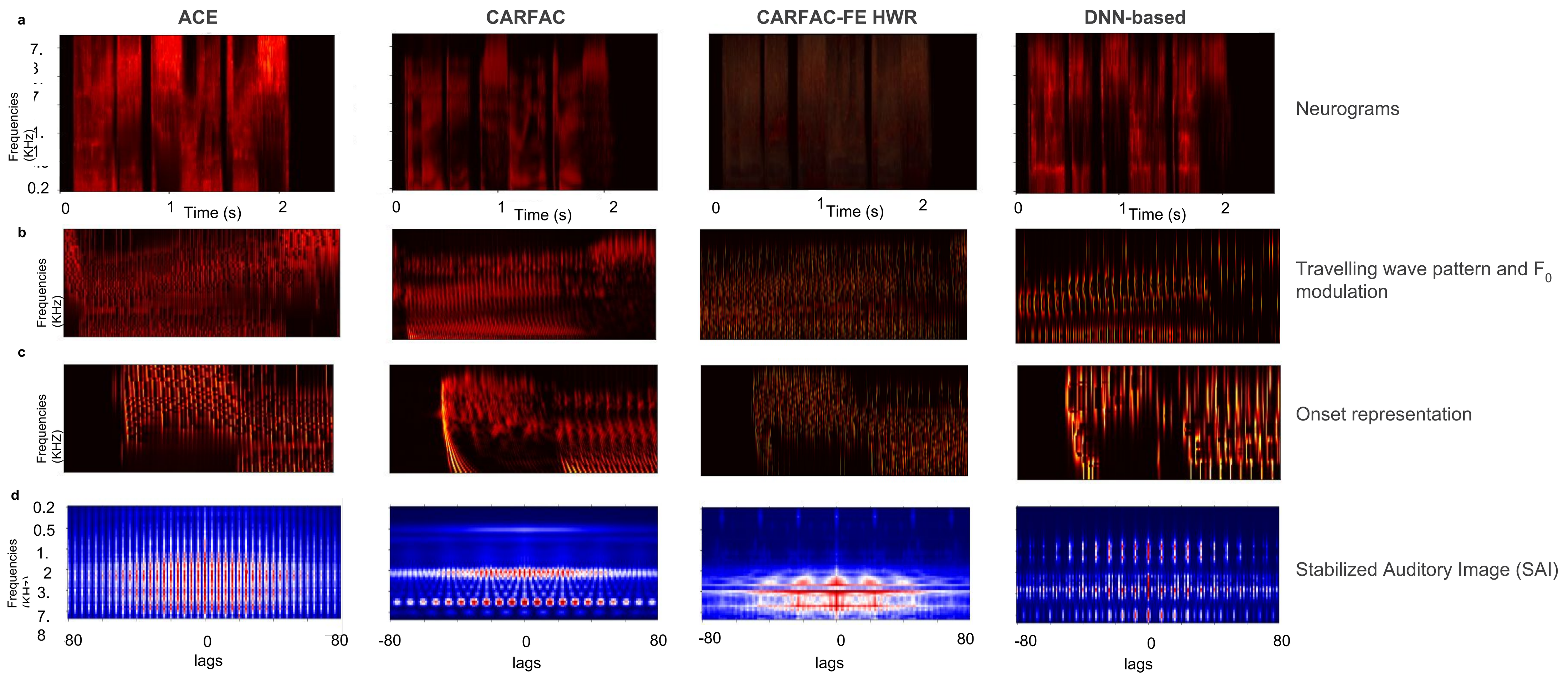
- The loss function for training the DNN used a combination of mean squared error (**MSE**) between neurograms, MSE of stabilized auditory images (**SAI**) of neurograms and **structural similarity index (SSIM)** of SAIs.
- Electric Hearing model needed to be **fast** *and* **differentiable** to incorporate in loss function.
- Model was trained on 34 hours of clean speech, noisy speech and noise.



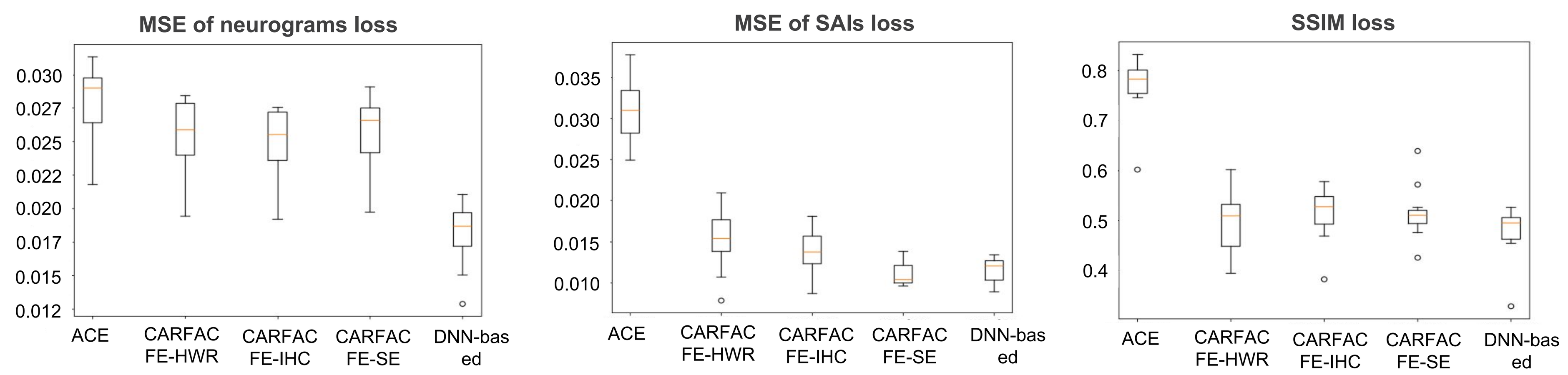
Results



Objective measures on vocoded strategies. For PESQ, STOI, ASR Accuracy (ASR_ACC) and ASR Similarity (ASR_SIM), higher is better. For ASR Word Error Rate (ASR_WER), lower is better

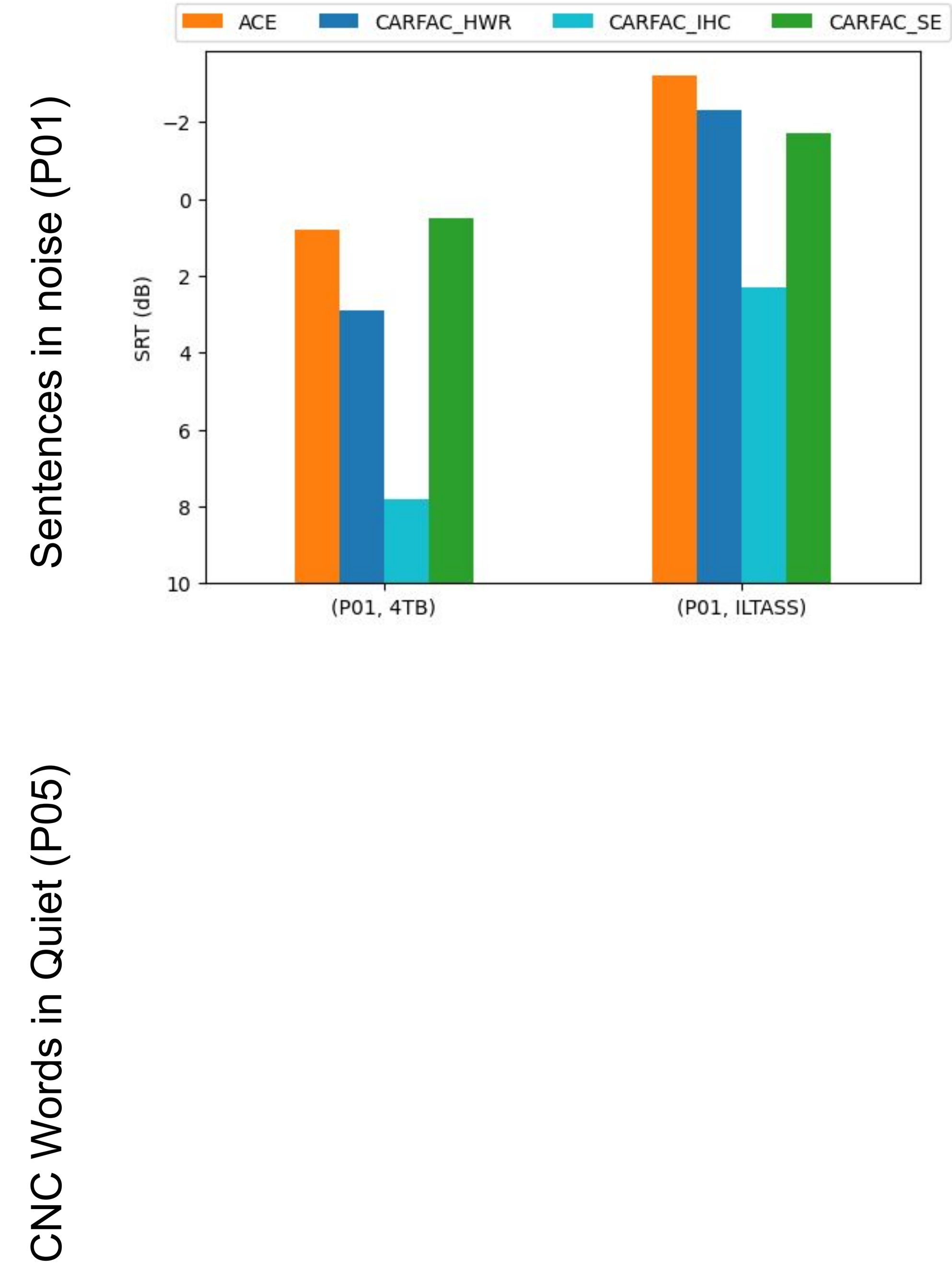


An example of the results for ACE strategy (left), CARFAC output (middle left), CARFAC-FE-HWR (middle right) and DNN-based method (right). A) NAPs for the sentence “The lady is sewing a dress”. B) 125ms segment of (a), showing the travelling wave pattern and F0 modulation. C) 125ms segment of (a) showing the onset enhancement. D) a 40ms SAI frame showing better reconstruction of F0 modulation for DNN-based method and CARFAC-FE HWR compared to ACE.



Recipient Feasibility Study – Preliminary Results

- We have begun a CI recipient study, with results for two subjects (P01, P05).
- Processing of wav files was performed off-line in MATLAB, and stimuli were delivered using NIC4. No live-microphone listening was possible. Recipients listened to a five-minute audio book for acclimatisation.
- We tested three variants of CARFAC front-end:
  - HWR: Simple half-wave rectification inner hair cell model
  - IHC: Non-linear inner hair cell model
  - SE: IHC model with spectral enhancement post-processing
- CARFAC-HWR and -SE provided similar speech intelligibility to ACE.



CONCLUSIONS

- Our CARFAC-based biomimetic sound coding strategies can produce neurograms similar to the neurograms obtained with a **normal hearing model.**
- By using these novel biomimetic sound coding strategy , we can represent the **fundamental frequency, onset enhancement and travelling wave delay** better compared to ACE strategy. It has been shown that better **onset** enhancement or F_0 modulation could lead to improvement is speech understanding in CI recipients [3].
- PESQ, STOI, and ASR-based analysis suggest that some of these experimental strategies could improve speech perception in noise compared to ACE.
- A CI recipient study has commenced. Preliminary results are encouraging.

REFERENCES

1. Lyon et al. (2024), The CARFAC v2 Cochlear Model in Matlab, NumPy, and JAX.
2. Brochier et al.(2022), From microphone to phoneme: an end-to-end computational neural model for predicting speech perception with cochlear implants. *IEEE Transactions on Biomedical Engineering*, *69*(11), 3300-3312.
3. Vandali et al. (2019), Evaluation of the Optimized Pitch and Language Strategy in Cochlear Implant Recipients. *Ear and Hearing* 40(3):p 555-567