Effects of Speech-to-Noise Ratio on Cognitive Performance: Assessing Attention to Speech with Natural Background Sounds

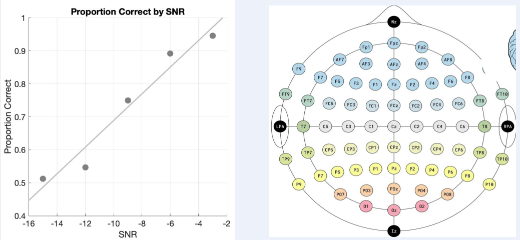
Anushka Molugu[1], Tylor Harlow[1], Laila Almotwaly[1], Heather Read[1,2]
Physiological Sciences[1] and Biomedical Engineering[2] , University of Connecticut, Storrs, CT

INTRODUCTION

Distinct brain activity patterns are engaged when listening to foreground sounds versus ignoring background sounds. Attention to speech sounds drives cortical electroencephalogram (EEG) responses that are correlated with temporal modulations in the speech envelope (1-3). Similarly non-speech sound are classified via temporal spiking patterns in neurons located in low level auditory midbrain and cortex (4). This study varies the signal-to-noise (SNR) ratios for foreground speech versus background natural sounds to make it challenging for subjects to correctly identify spoken words (numbers 1-9). The goal of this study is to understand the brain mechanisms underlying the perceptual ability to recognize and attend to speech versus background natural sounds.

Figure 1

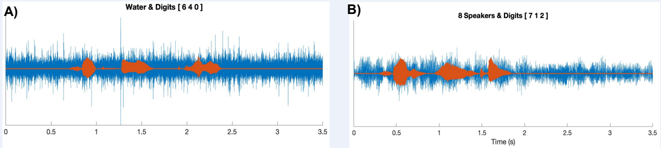
Fig 1. Psychophysics and EEG recording.



The behavioral task had subjects attend sequences of speech (i.e. digits 1-9) combined with natural background sounds and after a short delay they had to report what "digits" they heard on each trial using a computer keyboard. During task performance, 64-channel EEG signal recordings were obtained. Psychophysics and EEG data were acquired using MATLAB, Psych Toolbox and BrainVision..

Figure 2

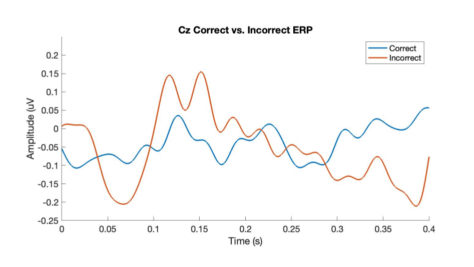
Fig. 2. Speech-in-Noise Sound Waveforms



Relatively loud natural background sounds (blue) makes attention to speech (orange) challenging. **A)** Example three digit sequence (orange) and natural water sound (blue). **B)** Example three digit sequence (orange) and background speech babble (8 speakers).

Figure 3

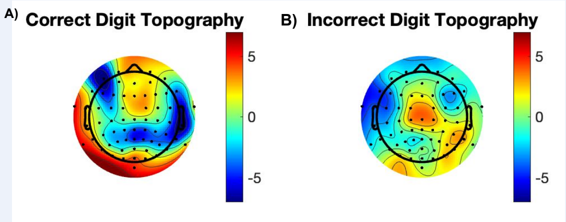
Fig 3. Example Speech Evoked Responses for Correct vs Incorrect Speech Recognition Trials



Speech-evoked responses measured at central electrode (Cz) shows higher amplitude early (0.075 s) and lower amplitude late (>0.100 s) components for trials with correct versus incorrect speech recognition.

Figure 4

Fig 4. Topographical Speech-Evoked Response Maps



Speech evoked response power in 400 ms time-window following speech onset is averaged for "Correct" (A) versus "Incorrect" (B) speech recognition trials. **A)** High responses power (red-yellow) observed in frontal and left temporal cortex with correct speech recognition. **B)** Incorrect response trials have high evoked responses localized central (Cz) a location typically picking up activity from primary auditory cortex on Heschl's gyrus.

CONCLUSION AND SIGNIFICANCE

- Correct speech recognition trials activate distinct frontal and temporal cortices indicating effective top-down executive control function.
- Incorrect speech recognition trials activate central locations that index primary auditory cortex (Heschl's gyrus) activity indicating disrupted top-down neural network activity with challenging signal-to-noise (SNR) and incorrect responses.
- Future studies will examine whether attention-related alpha oscillations are maintained in frontal & temporal cortices with correct speech recognition.

ACKNOWLEDGEMENTS & CITATIONS

Funded by Undergraduate PCLB grant, Molugu; IBACS seed grant Prof. Read
1-3) Luo, 2007; Peelle et al., 2013; Luo 2023 (4) Zhai, Read, Escabi et al., PNAS, 2020