

Teoh, Ahmed, F., & Lalor, E. C. (2022). Attention Differentially Affects Acoustic and Phonetic Feature Encoding in a Multispeaker Environment. *The Journal of Neuroscience*, 42(4), 682–691.  
<https://doi.org/10.1523/JNEUROSCI.1455-20.2021>

Keywords:

multi-speaker environment, attention, phonetic feature encoding, EEG

# Background

- Cocktail Party Problem: the ability to focus on a single talker amidst multiple sounds
- Various theories model attention as a selective filter that rejects unattended message beyond a certain information processing stage (Cherry 1953; Broadbent 1958; Moray 1959; Deutsch and Deutsch 1963; Treisman 1964; Johnston and Wilson 1980)
- Question: Does this filter operate at early or late stages of speech processing?

# Background

- Question: How does attention affect processing in different brain areas and at different latencies?
- Neuroimaging studies
  - Single talker: core auditory regions code low-level features and combined in higher areas to yield more abstract neural codes (Binder et al. 2000; Davis and Johnsrude 2003; Okada et al. 2010)
  - Multi-speaker selective attention: primary auditory cortical responses represent all talkers irrespective of attentional state but only attended speaker is represented at higher areas ie. superior temporal gyrus (Mesgarani and Chang 2012; Zion Golumbic et al. 2013; J. O'Sullivan et al. 2019)
- EEG/MEG studies
  - All talkers co-represented in early components, with distinct responses to the attended speaker appearing only in later components (Ding and Simon 2012b; Power et al. 2012; Puvvada and Simon 2017)

# Background

- Question: How does attention influence encoding of different speech representations thought to be encoded by the hierarchical network of the human cortex?
- Encoding/decoding methods:
  - Neural indices of lexical and semantic processing can only be found for attended speech, in contrast to acoustically-driven measures (ie. those based on the amplitude envelope) that are less affected by attention (Brodbeck et al., 2018)
- All of these findings supports notion that higher-order regions and higher-level representations are more greatly modulated by attention

# Research questions or hypotheses

- Examine how attention might differentially affect neurophysiological indices of hierarchical acoustic and linguistic speech representations at pre-lexical levels
- 2 main questions:
  - **1. Does the brain process speech at the level of phonemes and, if so, does it do so for both attended and unattended speech?**
    - Categorical responses to phonetic features OR simple acoustic features
    - To determine this, study will explore brain responses to speakers with very different acoustics (male vs. female)
  - **2. At what hierarchical levels does attention influence processing?**
    - Hypothesis: will see stronger attention effects on isolated measures of phonetic feature processing than on isolated measures of acoustic processing

# Speaker Information

- 14 subjects (9 female + 5 male)
- Between ages 19 – 30
- Right-handed
- Primary language: English
- No history of hearing impairment or neurological disorder

# Materials & Procedures

- Undertook 40 one-minute trials in 2 separate blocks
- Stimuli presented through headphones, with one side playing female talker and one side playing male talker
- Subjects instructed to attend to one of two talkers
  - Counterbalanced paradigm shift used to ensure subjects attend to both male and female talkers and at both locations
- Attended story segments presented contiguously and unattended segments presented in random order
- After each 60s trial, subjects answered 4 multiple choice comprehension questions on both stories (attended and unattended)

# EEG Data

- Conducted in soundproof room
- Biosemi ActiveTwo system used to record EEG data from 128 electrode positions on scalp and 2 over mastoid processes
- Digitized at 512Hz



# Stimuli

- Stimuli taken from 2 works of fiction, one narrated by a female talker and one narrated by a male talker
  - Sherlock Holmes Novels
    - Male Speaker: The Hound of the Baskervilles
    - Female Speaker: A Study in Scarlett
- Stimuli filtered using Head-Related Transfer functions, giving rise to perception of talkers at 90 degrees to left and right of subject

# Stimuli (specific phonetic contrasts)

- Speech Representations Measured:

- Acoustic features:
  - Spectrogram ( $s$ )
  - Spectrogram Derivative ( $sD$ )
- Phonetic features:
  - Phonetic Features ( $f$ )
  - Phonetic Feature Onsets ( $fo$ )

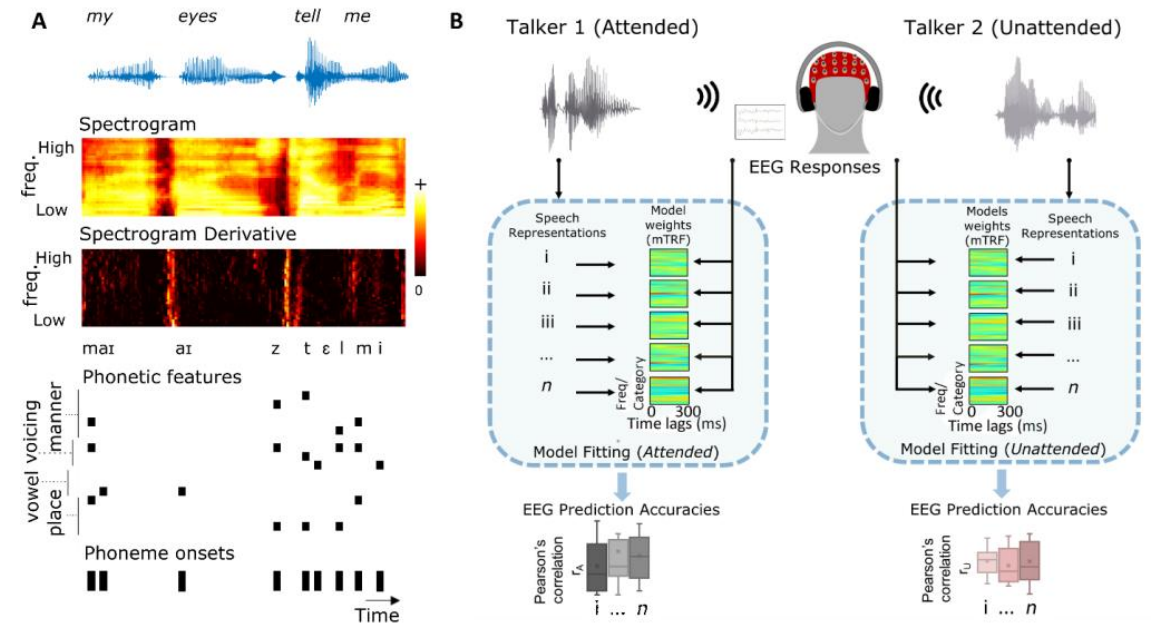


Fig. 1

# Two-stage Data Analysis

## **1) Does the brain process speech at the level of phonemes and, if so, does it do so for both attended and unattended speech?**

- Compared performance of different models in predicting EEG responses to speech
  - Determined whether or not inclusion of categorial phonetic feature representation of speech ( $f$ ) could improve prediction of EEG beyond other features
  - Determined whether improved prediction would hold for both attended and unattended speech

## **2) At what hierarchical levels does attention influence processing?**

- Identified ability of each stimulus feature to predict unique variance in the EEG responses and assessing whether that unique predictive contribution was affected by attention

# Findings

- 1. For attended speech, including 19-dimensional phonetic features representation improved the prediction of the EEG responses beyond that obtained when only using acoustic features
  - Was not true for unattended speech
- 2. Unique predictive power of the phonetic feature representation was enhanced for attended versus unattended speech
  - Was not true for any other feature

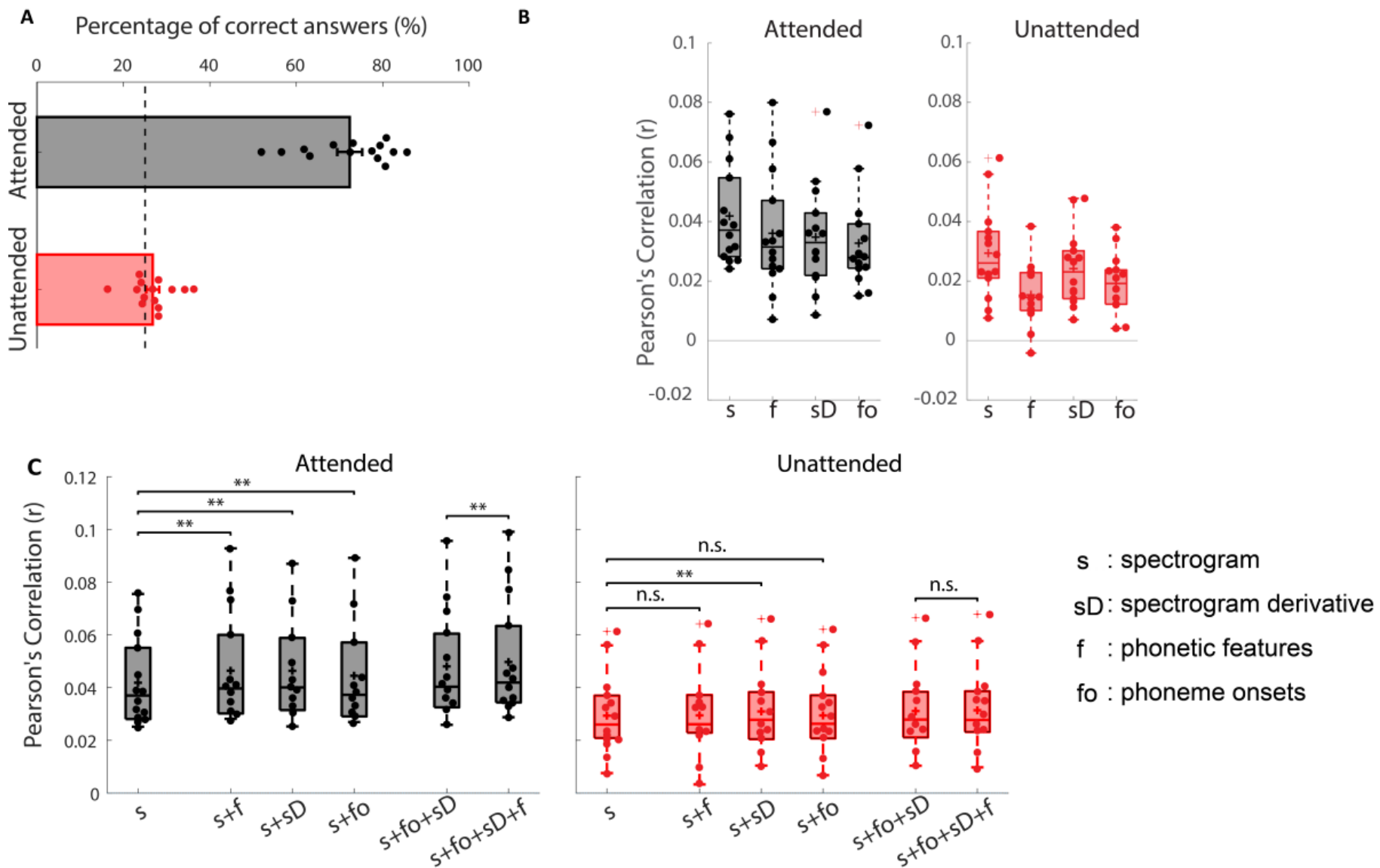


Fig. 2

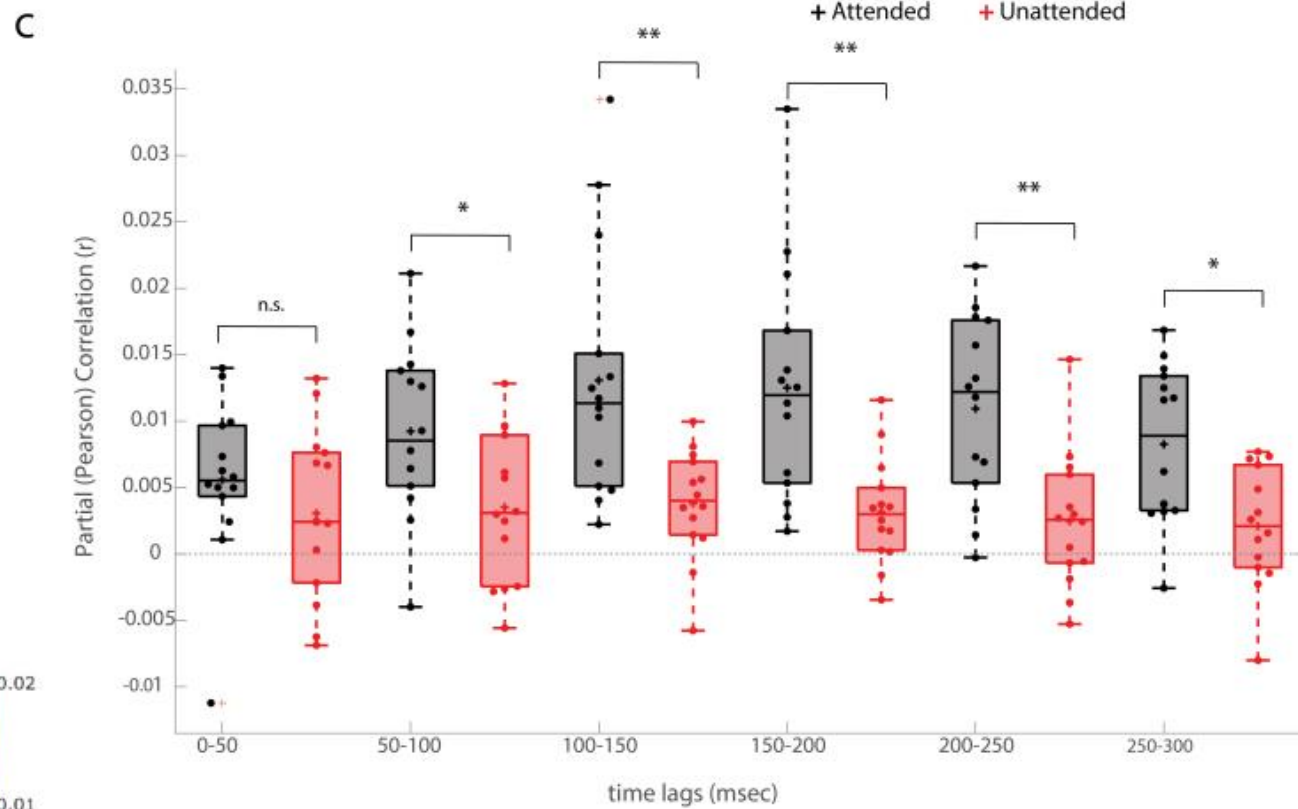
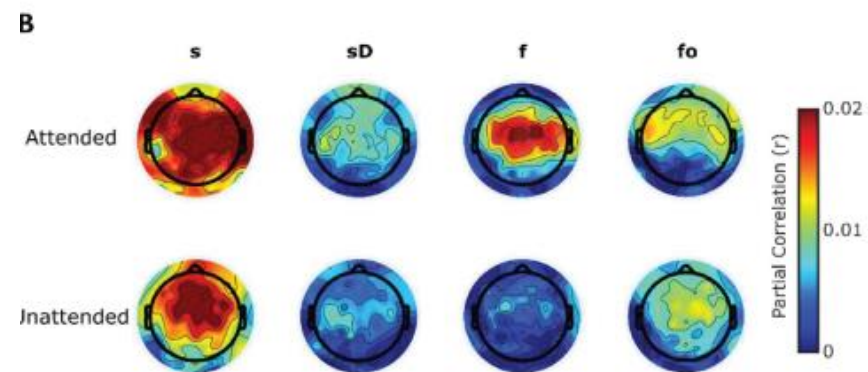
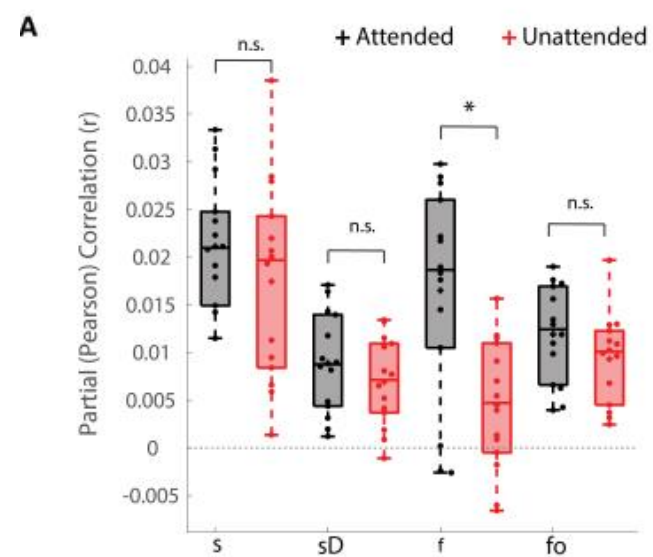


Fig. 3

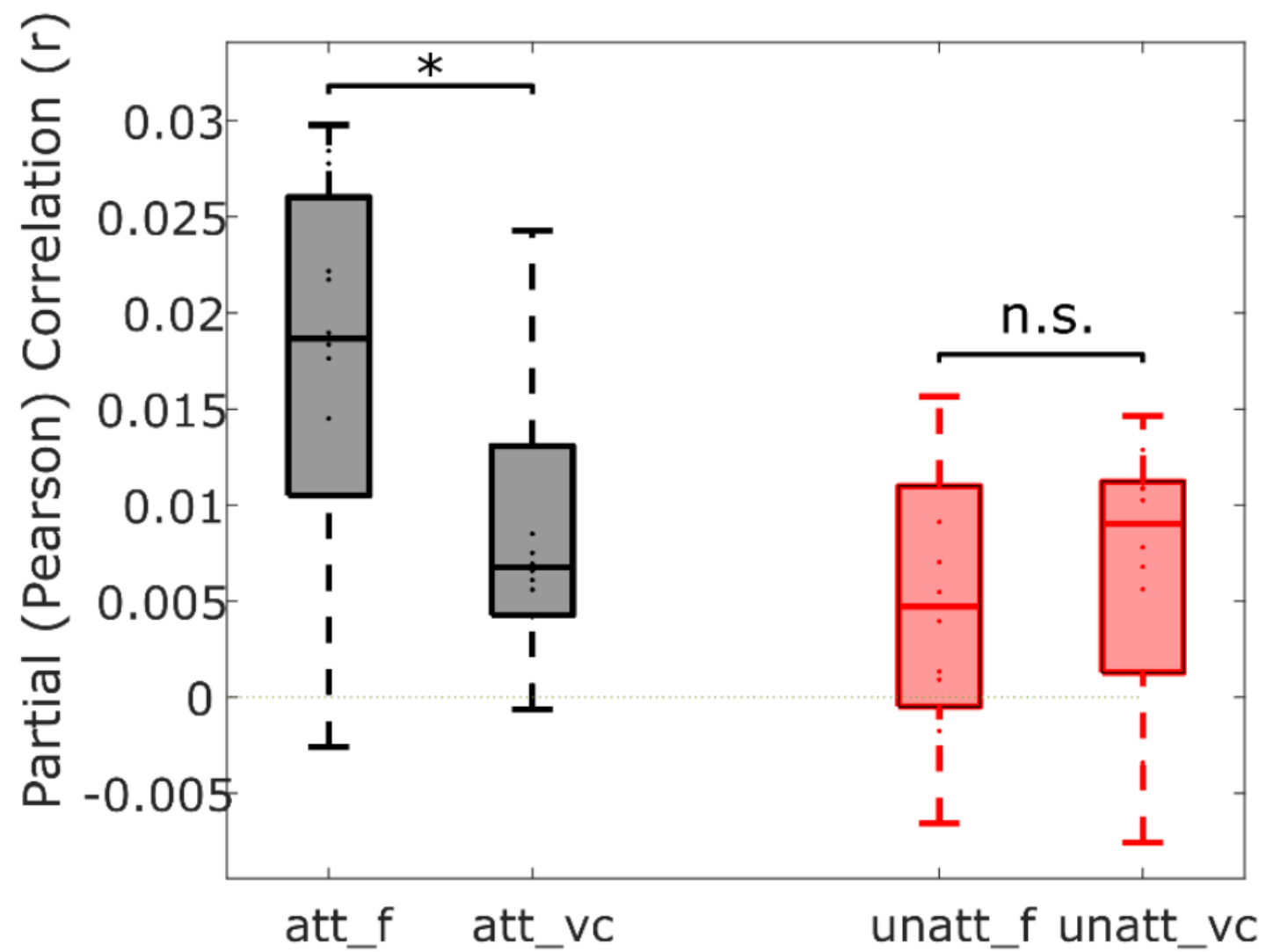


Fig. 4

# Most relevant info

- 2 contributions to literature
  - 1. contribute to debate around pre-lexical speech processing in cortex
  - 2. contribute to debate on how selective attention affects processing of speech
- Study shows that a phonetic feature representation has unique predictive power when modelling responses to attended speech, but not for unattended speech
  - Suggests processing attended speech may involve mapping from an acoustic to a categorical phonemic presentation
    - Controversial idea in current literature
  - Cocktail Party Attention debate: Study shows attention differentially modulates cortical processing of acoustic and phonetic information
    - Confirms stronger attention effects at higher levels of the speech processing hierarchy
- These results combined with previous literature suggest that attention is categorically selective for speaker-invariant representations and, at most, attenuates lower-level acoustic (speaker dependent) measures