Modulation Change Detection in Human Auditory Cortex: Evidence for Asymmetric, Nonlinear Edge Detection

Seung-Goo Kim[1], David Poeppel[2,3,4], Tobias Overath[1,5,6]

[1]Department of Psychology & Neuroscience, [5]Duke Institute for Brain Sciences, [6]Center for Cognitive Neuroscience, Duke University, NC, USA

[2]Department of Psychology, [3]Center for Neural Science, New York University, NY, USA, [4]Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany

Introduction

Motivation

- Temporal modulation is a prominent feature of natural sounds (eg. the syllabic rhythm of 4–7 Hz; the phonetic rhythm of 30–50 Hz [1]). Thus the detection of its changes is a crucial task of the auditory system, but the underlying neural mechanisms remain poorly understood.

Previous works

- ***Asymmetry:*** An emergence of an *auditory object,* as compared to its disappearance, evokes distinctive neural responses in terms of morphology and topology of [3-4]
- ***Nonlinearity:*** Human auditory cortex shows a compelling sensitivity to slow temporal modulations [5-7]

Hypotheses

- Stronger evoked responses to transitions to slow modulations than from them
- Phase tracking of temporal modulations in respective frequency bands (eg. higher phase coherence in low-frequency bands for slow temporal modulation rates)

Methods

Stimuli

- 20 random frequency bins from 101 loglinear bins between 246–4435 Hz
- Correlation between amplitude vectors in adjacent 20-ms frames: r = 0, 0.5, 0.8, 0.95 (linear increase in Z = 0, 0.55, 1.1, 1.83)
- 3 exemplars x 4 correlation levels
- 12 segments x 10 blocks x 4 sessions
- Exported at 44.1 kHz & 16 bit, presented at ~75 dB SPL

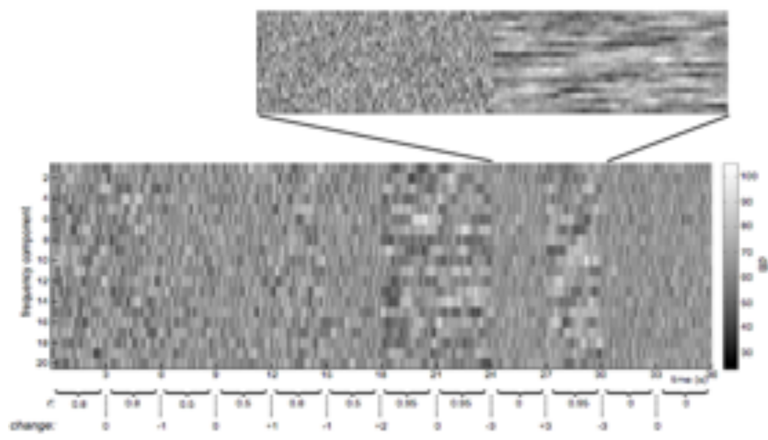

Figure 1. Schematic of the stimulus. The amplitude values of the 20 ms frames. The inset provides a more detailed view of a change from a segment with r = 0 to a segment with r = 0.95.

MEG data

- 160-channel whole-head MEG system (KIT, Kanazawa, Japan)
- Sampled at 1 kHz and online-filtered for 1–200 Hz with a notch at 60 Hz
- Task: to detect a transition of temporal modulation rate (after a pre-MEG session)

Participants

- Sixteen participants (18–33 years, 8 females), 3 excluded for poor MEG signal or chance task performance (total N = 13)

Data analysis

- Software: FieldTrip Matlab Toolbox, MNE-Python package, and custom codes
- Preprocessing: Butterworth-filtering 1-60 Hz, ICA-based artifacts correction
- Evoked response analysis: denoising source separation (DSS) [6] to extract reliably evoked (bias) components, absolute timeseries was further analyzed
- Phase tracking analysis:
  - cross-trial phase coherence (CTPC) [7] with Morlet Wavelet (3–11 cycles) $\mathrm{CTPC}(t,f) = \left(\frac{1}{N}\sum_{n=1}^{N}\cos\theta(n,t,f)\right)^2 + \left(\frac{1}{N}\sum_{n=1}^{N}\sin\theta(n,t,f)\right)^2$
  - inter-trial correlation (ITC) [8] with bandpass FIR-filtering for 5 freq. bands

References

(1) Pickett JM (1999) The Acoustics of Speech Communication: Fundamentals, Speech Perception Theory, and Technology. Boston, MA: Allyn & Bacon.
(2) Chait M, Poeppel D, de Cheveigné A, Simon JZ (2007) Processing asymmetry of transitions between order and disorder in human auditory cortex. J Neurosci 27:5207-5214.
(3) Chait M, Poeppel D, Simon JZ (2008) Auditory temporal edge detection in human auditory cortex. Brain Res 1213:78-90.
(4) Overath T, Kumar S, Stewart L, von Kriegstein K, Cusack R, Rees A, Griffiths TD (2010) Cortical mechanisms for the segregation and representation of acoustic textures. J Neurosci 30:2070-2076.
(5) Overath T, Zhang Y, Sanes DH, Poeppel D (2012) Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: fMRI evidence. J Neurophysiol 107:2042-2056.
(6) de Cheveigne A, Simon JZ (2008) Denoising based on spatial filtering. J Neurosci Methods 171:331-339.
(7) Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron 54:1001-1010.
(8) Ding N, Simon JZ (2013) Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. The Journal of Neuroscience 33:5728-5735

Evoked respones

DSS1 timeseries

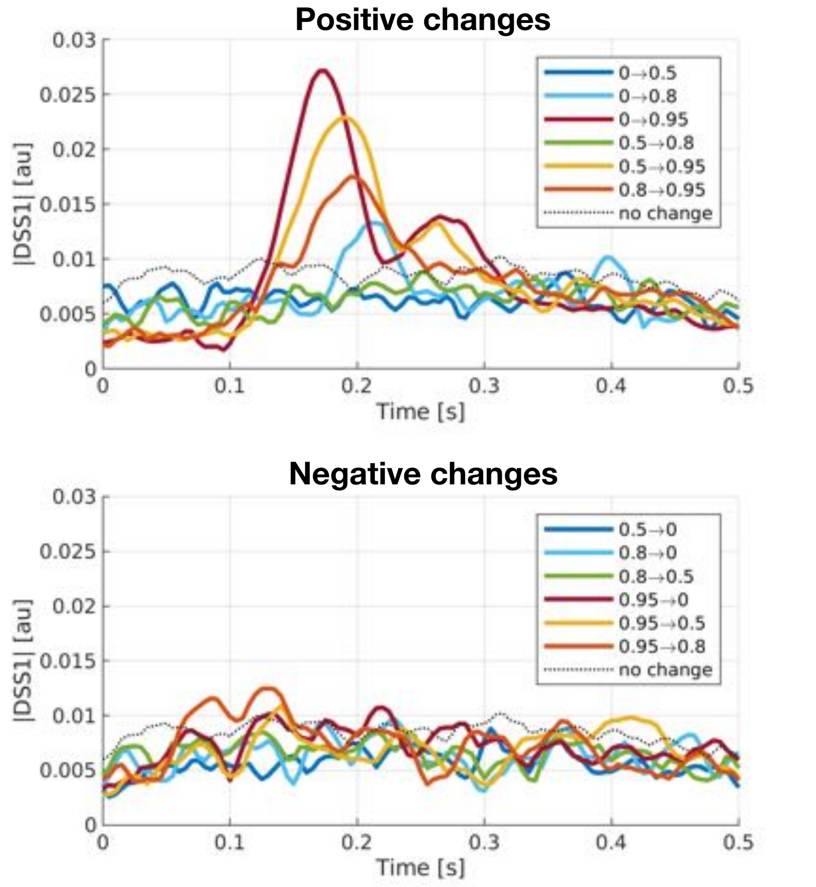

Figure 2. DSS1 timeseries averaged for conditions with positive changes (upper) and negative changes (lower) are plotted separately. For the "no change" condition, DSS1s were extracted from conditions without changes (e.g., 0→0, 0.5→0.5, …) separately, then averaged.

Projected DSS1 topography

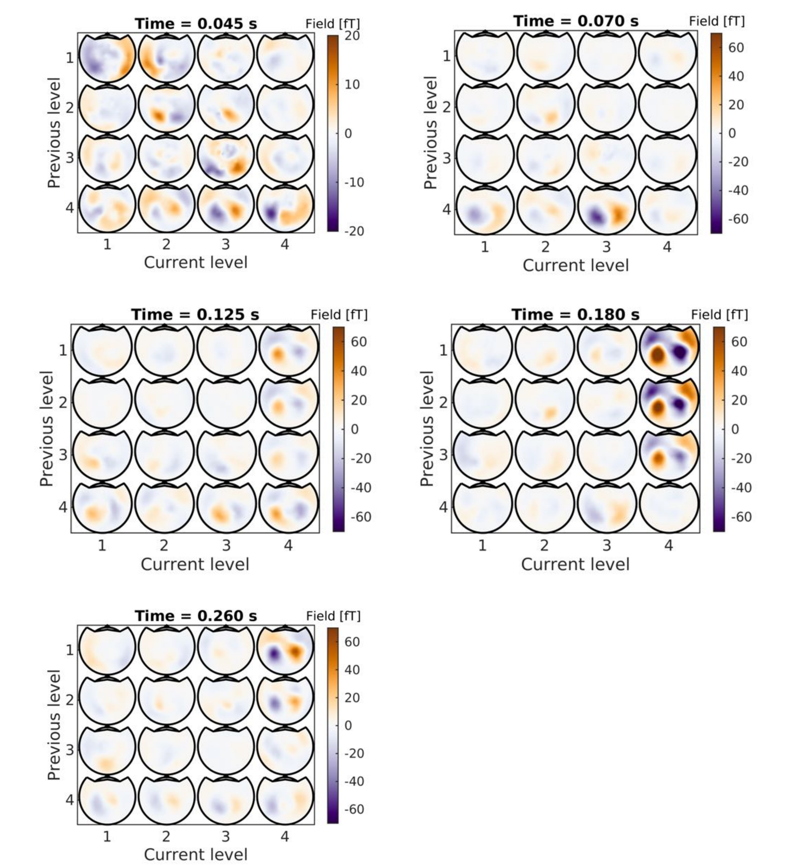

Figure 3. DSS1 components at peak latencies of rm-ANOVA (Figure 4 & 5) projected onto the sensor space and averaged across subjects. Previous and Current Levels of a transition correspond to rows and columns of a 4-by-4 matrix, respectively.

Two-way repeated-measures ANOVA models:

Previous Level x Current Level

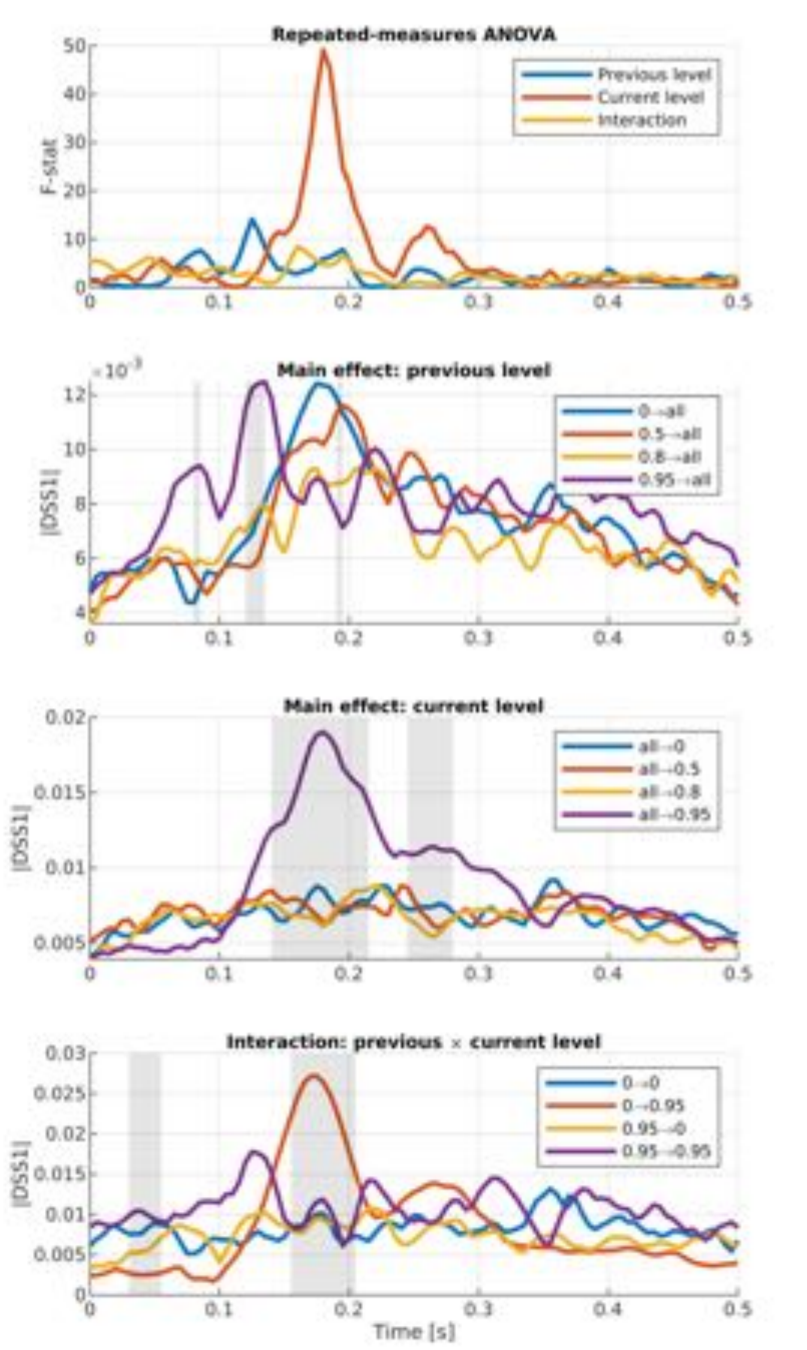

Figure 4. F-statistics for main effects and interaction of Previous and Current Levels (top), and grand-average DSS1 timeseries demonstrating the effects with significant latencies (cluster-based corrected P < 0.05) marked in gray shades (upper middle to bottom).

Direction x Absolute Step

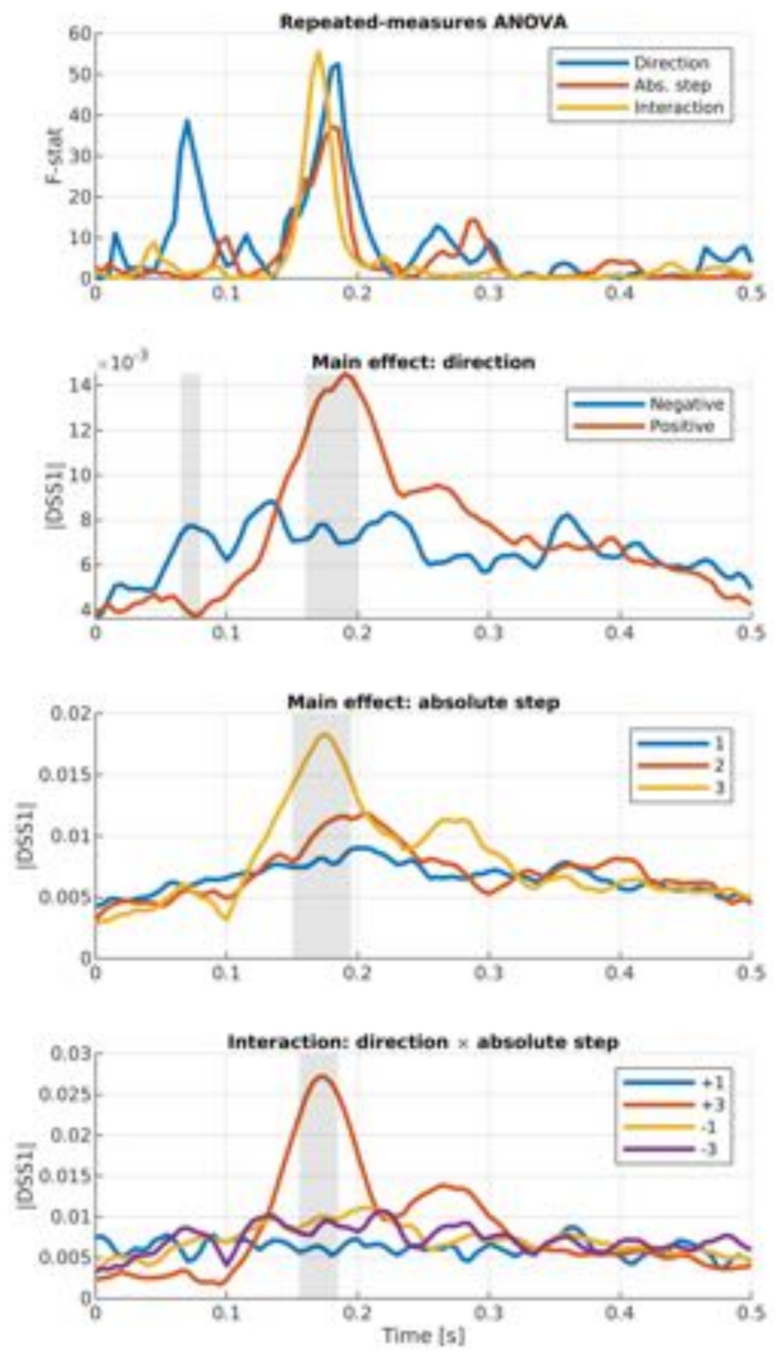

Figure 5. F-statistics for main effects and interaction of Direction and Absolute Step (top), and grand-average DSS1 timeseries demonstrating the effects with significant latencies (cluster-based corrected P < 0.05) marked in gray shades (upper middle to bottom).

Behavioral results

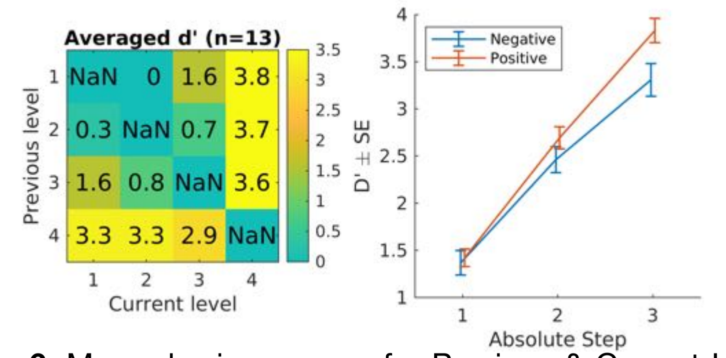

Figure 6. Mean d-prime scores for Previous & Current Levels (left) and Direction & Absolute Step (right)

Phase tracking

Cross-trial phase coherence

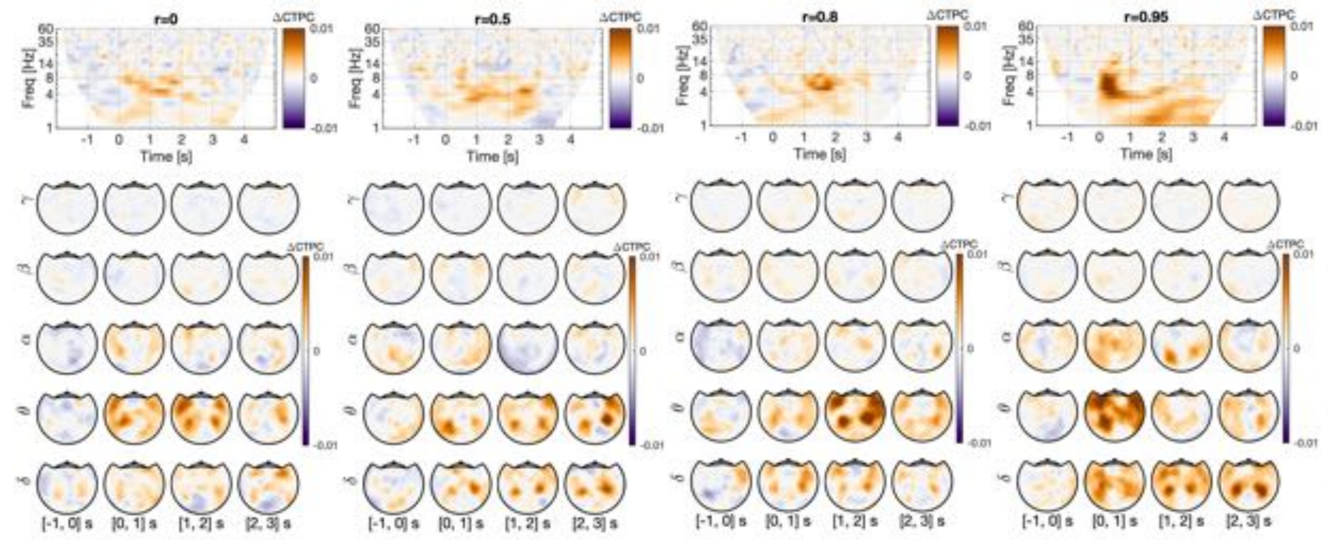

Figure 7. Average cross-trial phase coherence (CTPC) difference (across-exemplar CTPC subtracted from within-exemplar CTPC). CTPC differences averaged across all channels and topography averaged for 1-sec time bins and frequency bands (δ, 1-3 Hz; θ, 4–7 Hz; α, 8–12 Hz; β, 13–29 Hz; γ, 30–60 Hz) for each correlation level.

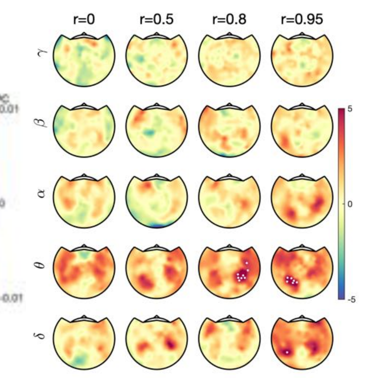

Figure 8. One-sample T-test. Averaged over [0, 3] s period.

Inter-trial correlation

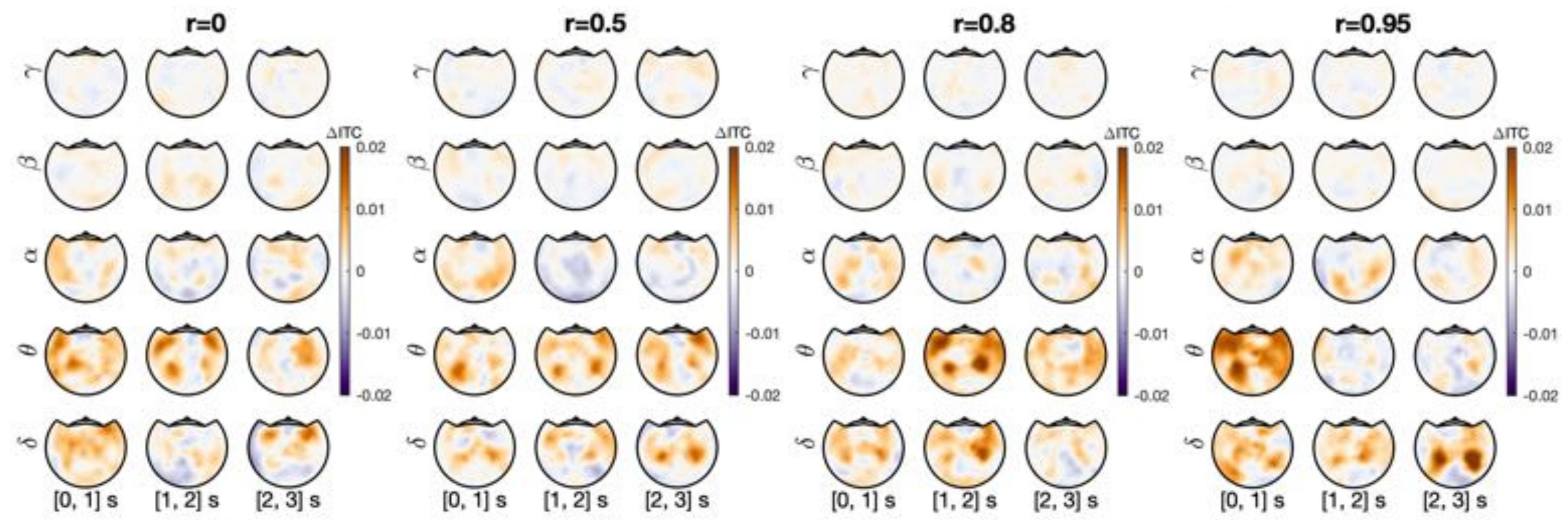

Figure 8. Average inter-trial correlation (ITC) difference (across-exemplar ITC subtracted from within-exemplar ITC). ITC differences computed from 1-sec time bins and frequency bands for each correlation level.

Conclusions

- Behavioral and evoked neural responses showed asymmetrical responses (i.e., stronger responses to transitions to the slowest temporal modulation as compared to transitions from it).
- The magnitude of responses was not linearly dependent on the absolute step size of changes (i.e., nonlinear sensitivity to slow temporal modulation)
- Contrary to our hypothesis, phase tracking (CPTC, ITC) in respective frequency bands was not significant.