Hierarchical processing of temporal information during naturalistic music production and perception

Riesa Cassano[1], Jamal Williams[2], Marius Cătălin Iordan[2], Uri Hasson[2], Elise A. Piazza[1]
[1]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627
[2]Princeton Neuroscience Institute and Department of Psychology, Princeton University, Princeton, NJ 08544
Contact: riesa_cassano@urmc.rochester.edu

Introduction

- Recent work has used scrambled versions of naturalistic stimuli to reveal hierarchies of temporal processing in the brain (Hasson et al., 2008; Lerner et al., 2011).

- Hierarchical processing has been characterized during naturalistic music perception (Farbood et al., 2015).

- In this study, we attempt to extend these findings to naturalistic music production. We also explore new ways of investigating hierarchical processing within an intact piece of music.

Methods

- Expert pianists played a non-ferromagnetic three-octave keyboard in the MRI scanner.

- Subjects sight-read versions of a 4-minute Tchaikovsky medley that were intact as well as scrambled at the phrase level (approximately eight bars), two-bar level, and one-bar level. They played three repetitions of each scrambled condition.

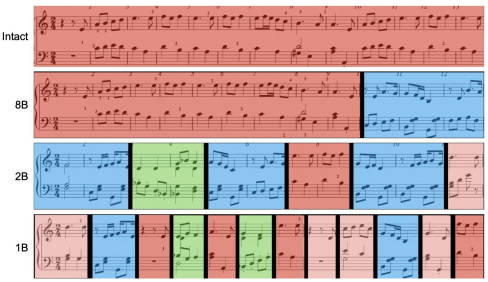

Figure 1 | Stimuli were scrambled versions of naturalistic music. Colors correspond to which section of the original medley the bar comes from.

- Subjects whose data is analyzed here received real-time auditory feedback of themselves playing.

- Intersubject temporal correlation (ISTC) measures the reliability of responses across subjects in a region by averaging over voxels and correlating time series. Intersubject pattern correlation (ISPC) measures the reliability of patterns of neural activity across subjects in a region by averaging voxel activity over time and correlating average patterns.

Results

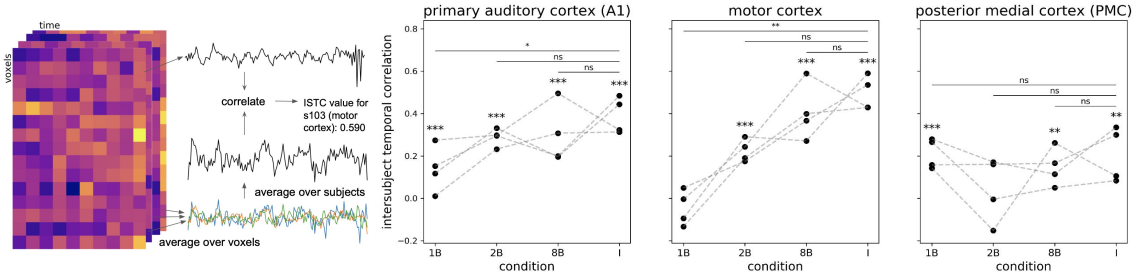

Figure 2 | Intersubject temporal correlation (ISTC) in first repetition of each scrambled condition. Dashed gray lines connect ISTC values of individual subjects. Significance of average ISTC value above chance was determined by phase-randomizing permutation test (1000 samples) and significance of difference between conditions was determined by label-shuffling permutation test (1000 samples). *p<0.05, **p<0.01, ***p<0.001

- In A1 and motor cortex, we expect no change in reliability across scramble conditions. However, there appears to be a systematic increase in ISTC from 1B to Intact.

- In PMC, we expect reliable response *only* in the most intact conditions, since this region has been shown to integrate incoming information over relatively long timescales. However, there is not a difference in reliability across scramble conditions.

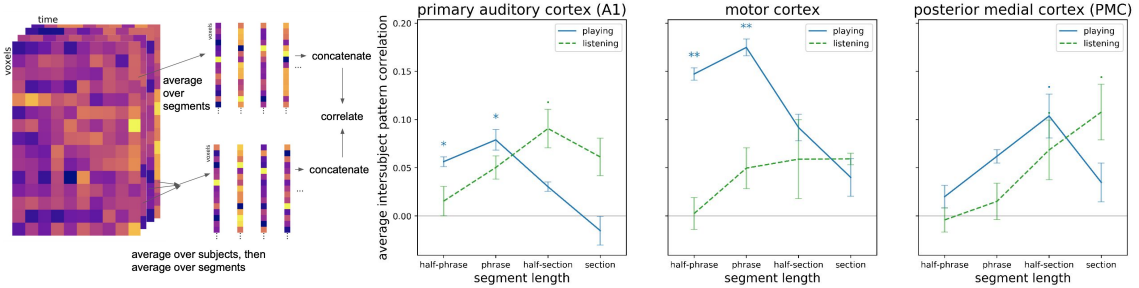

Figure 3 | Intersubject pattern correlation (ISPC) in first repetition of Intact playing condition and average repetitions of listening condition. Patterns of activity were averaged over stimulus-defined segments at each level, then concatenated and correlated with the average of concatenated patterns of other subjects. Error bars show standard error of the mean (N = 4). Significance of pattern correlation values were determined by shuffling brain data in time, imposing original boundaries, and recomputing pattern correlation values (1000 samples). ·p<0.10, *p<0.05, **p<0.01, ***p<0.001

- Due to possible differences in attentional load across scramble conditions, ISPC analysis within the Intact condition may be more sensitive to hierarchical effects.

- During playing, A1 and motor cortex show reliable patterns of activity over shorter segments, suggesting that subjects are chunking what they are playing into shorter segments (like phrases).

- During listening, A1 and PMC show more reliable response patterns over longer segments, suggesting that subjects might devote more attention to longer-term chunking while listening. The listening condition was presented at the end of the session, so fatigue effects may impact the data in this condition.

Discussion

Ongoing analysis and future directions

- Results presented here are from only four subjects. We are planning to scan additional performers.

- To better understand how the subjects might be segmenting the stimuli as they are playing, we are working on an analysis to detect event structure in neural data (Williams et al., 2021).

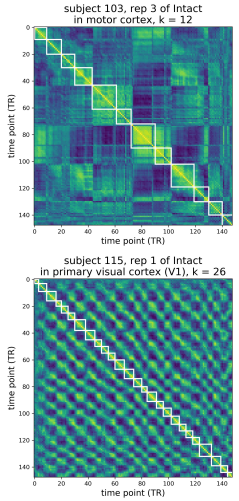

Figure 4 | Preliminary event segmentation findings. Hidden Markov model (HMM) was fit to data averaged over all subjects. Best number of events was selected from a set of k values according to best model fit, measured by greatest difference in correlation within event vs across event. Event boundaries are plotted on correlation matrices for single subjects.

- HMM finds longer events in motor cortex and shorter events (related to eye movements) in V1.

- We are planning to collect data on listeners' perception of event boundaries. We expect that HMM-predicted boundaries in higher order areas match to the listener-perceived boundaries.

References

Farbood, M. M., Heeger, D. J., Marcus, G., Hasson, U., & Lerner, Y. (2015). The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience, 9*. https://doi.org/10.3389/fnins.2015.00157

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A Hierarchy of Temporal Receptive Windows in Human Cortex. *The Journal of Neuroscience, 28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *The Journal of Neuroscience, 31*(8), 2906–2915. https://doi.org/10.1523/JNEUROSCI.3684-10.2011

Williams, J. A., Margulis, E. H., Nastase, S. A., Chen, J., Hasson, U., Norman, K. A., & Baldassano, C. (2021). *High-order areas and auditory cortex both represent the high-level event structure of music* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.01.26.428291