Comparing the Timescale of Cortical Speech and Music Analysis with Sound Quilts

Dana Boebinger*[2,3], Sam Norman-Haignere*[1,2], Josh McDermott[2,3]

*Co-First Authors

[1] HHMI Postdoctoral Fellow of the LSRF

[2] Brain and Cognitive Sciences, MIT

[3] Speech and Hearing Bioscience and Technology, Harvard University

Recent work suggests that human non-primary auditory cortex contains distinct neural populations that respond selectively to speech and music, respectively. However, little is known about what these neural populations compute. We have previously used “quilting” – a technique for scrambling sound waveforms at different durations – to investigate the sensitivity of speech-selective populations to sound structure at different timescales. Here, we apply the same technique to music, and compare the results to those for speech. We used voxel decomposition to estimate the fMRI response of speech- and music-selective populations to the sound quilts. Replicating prior findings, the response of speech-selective populations increased with increasing segment duration up to approximately 500ms, after which the response plateaued. The response of music-selective populations was qualitatively similar to that of speech-selective populations. Importantly, neither neural population showed a significant response to the quilting manipulation of the non-preferred stimulus type. We then applied the same approach to four other components that reflect selectivity for certain acoustic features of the sounds, and none of them showed a substantial effect of quilting. Sensitivity to temporal structure in speech and music appears to be a signature of higher-order processing in nonprimary regions of auditory cortex.