Auditory Object Formation in Temporally Complex Acoustic Scenes

*Berfin Bastug[1,2,3], Yue Sun[1,5], Erich Schröger[2,3], David Poeppel[2,4]

1. Ernst Strüngmann Institute for Neuroscience, Frankfurt am Main, 2. Max Planck School of Cognition, 3. Wilhelm-Wundt-Institute of Psychology, Leipzig University, Leipzig, 4. New York University, New York, 5. Cooperative Brain Imaging Center (CoBIC), Goethe University Frankfurt

The auditory system decomposes boundary-less sensory input into meaningful units through Auditory Scene Analysis (Bregman, 1990). Repetition helps listeners segregate overlapping sounds, and identify distinct auditory objects (McDermott et al., 2011). Studies suggest that repeated units in noisy contexts can eventually be perceived as stable auditory objects (Barczak et al., 2018; McDermott et al., 2011), but the behavioral signature of this dynamic process remains largely unexplored.
We investigated this using "tone clouds" —randomly generated clusters of 50-ms tones lacking explicit boundary cues. Repetition strength was manipulated by adjusting the ratio of repeated to regenerated tones, creating a continuum from random to repeated sequences. This formed an auditory analog to motion coherence tasks. To perceive repetition, listeners had to group repeated tones into auditory objects, allowing us to probe the minimal sensory evidence required.
There were two tasks: repetition detection and sensorimotor synchronization (SMS). In detection, participants judged if sequences repeated. We varied unit duration to examine how temporal structure affects this process. In SMS, participants tapped in sync with the repeating pattern, providing a real-time behavioral measure of perceptual organization.
We show sigmoidal performance across repetition levels in both experiments. Auditory object formation depends on repetition strength and longer durations need more evidence. But once repetition is detectable, ˜4 cycles are needed to make a judgment, regardless of unit duration. This suggests the evidence is integrated over cycles. In the SMS, sigmoidal curves converge across unit durations, eliminating the interaction effect. Trial progression analysis reveals two stages during object formation: when repetition is detectable, performance gradually builds up before reaching a saturation point, suggesting a categorical perceptual shift in strong repetition conditions, in which the additional evidence no longer enhances performance.

Keywords: auditory perception, repetition detection, auditory objects, sensorimotor synchronization