Direct Cortical Recordings from Humans Engaged in a Classical Auditory Streaming Task

Andrew R. Dykstra,[1] Eric Halgren,[2] Thomas Thesen,[3] Chad Carlson,[3] Amy Trongnetrpunya,[3] Ruben Kuzniecky,[3] Werner Doyle,[3] Emad Eskandar,[4] and Sydney S. Cash[5]

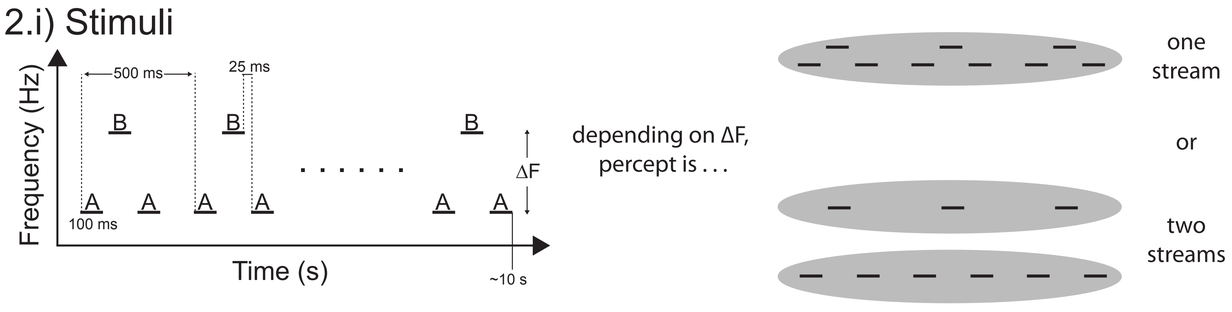
[1]Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology, [2]Departments of Radiology and Neurosciences, University of California, San Diego
[3]Comprehensive Epilepsy Center, Department of Neurology, New York University Medical Center, [4,5]Departments of Neurosurgery and Neurology, Massachusetts General Hospital and Harvard Medical School

1. INTRODUCTION

Despite having only two receptors, the auditory system is able to parse the complex mixture of sound arriving at the ear drums into an accurate representation of the underlying acoustic environment. Crucial in daily life, this process - termed *auditory scene analysis* (Bregman, 1990) - allows us to avoid hazards and attend to sound sources of interest while ignoring irrelevant noise. Despite its behavioral relevance and recent efforts employing diverse methods, the neural mechanisms supporting auditory scene analysis are not well understood. Using a classical "streaming" paradigm (ABA-ABA-...) and direct cortical recordings from the pial surface of patients with epilepsy, the present study combined advantages (high resolution in both space and time, broad coverage, concurrent behavior, direct enquiry) and circumvented limitations of previous work. Specifically, we sought to test previously-proposed models of auditory streaming and to extend the characterization of neural correlates of the streaming effect into brain areas outside the classically-defined auditory cortex.

2. METHODS

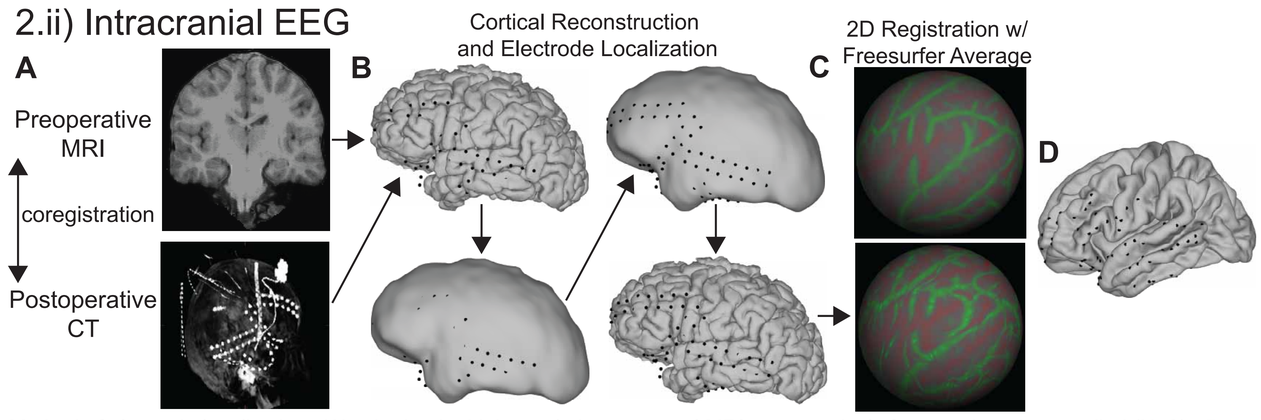
2.i) Stimuli



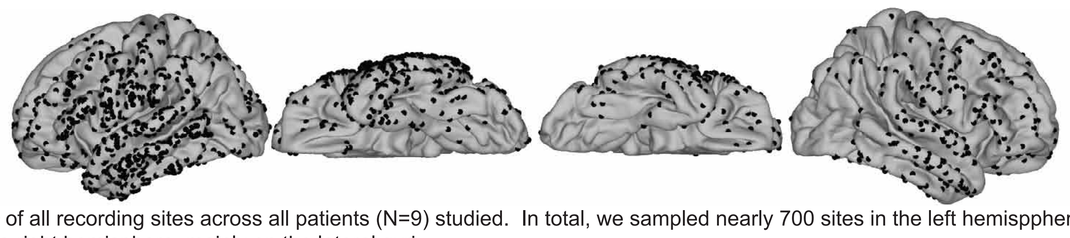
Stimuli were sequences of pure tones alternating in frequency. Individual tones were 100 ms long with 10-ms rise and fall times. The tone-onset asynchrony was 125 ms; the triplet-onset asynchrony was 500 ms. The frequency separation (ΔF) between tones varied between 0 and 12 semitones (0 and 100%, respectively). At the end of each sequence lasting between 6.5 and 10 seconds, patients indicated by button press whether they heard one or two "streams." Each experimental session lasted approximately 0.5 hours, and between one and five sessions were completed in each patient.

2. METHODS (cont'd)

2.ii) Intracranial EEG



Method of electrode localization. We obtained a high-resolution structural MRI preoperatively from each patient studied (N=9). 5 patients underwent postoperative CT scans; 4 underwent postoperative MRI.
(A) Preoperative MRI was coregistered with either postoperative CT (shown) or MRI (not shown) in order to obtain 3D electrode coordinates in the patient's native anatomical space. (B) After computing a 3D cortical reconstruction using Freesurfer, electrodes were "snapped" to the nearest cortical surface vertex (first on the dural surface, then on the pial surface). (C) Then, a 2D spherical surface of each patient was computed and coregistered with the Freesurfer average brain. (D) Electrode coordinates were transformed into this space by nearest-neighbor matching in the 2D spherical space, and finally overlaid onto the Freesurfer average pial surface.

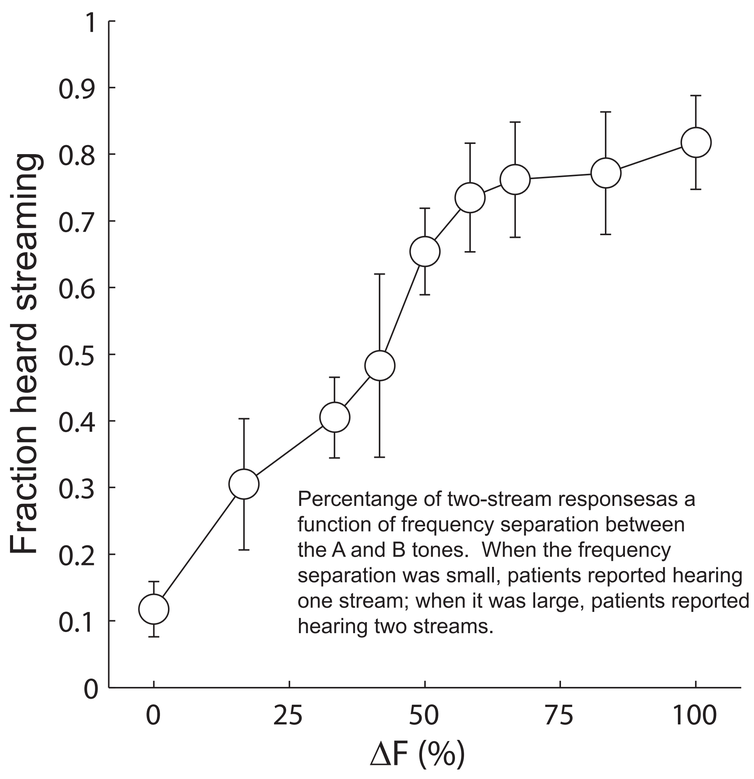


Summary of all recording sites across all patients (N=9) studied. In total, we sampled nearly 700 sites in the left hemisppshere and 250 sites in the right hemisphere, mainly on the lateral gyri.

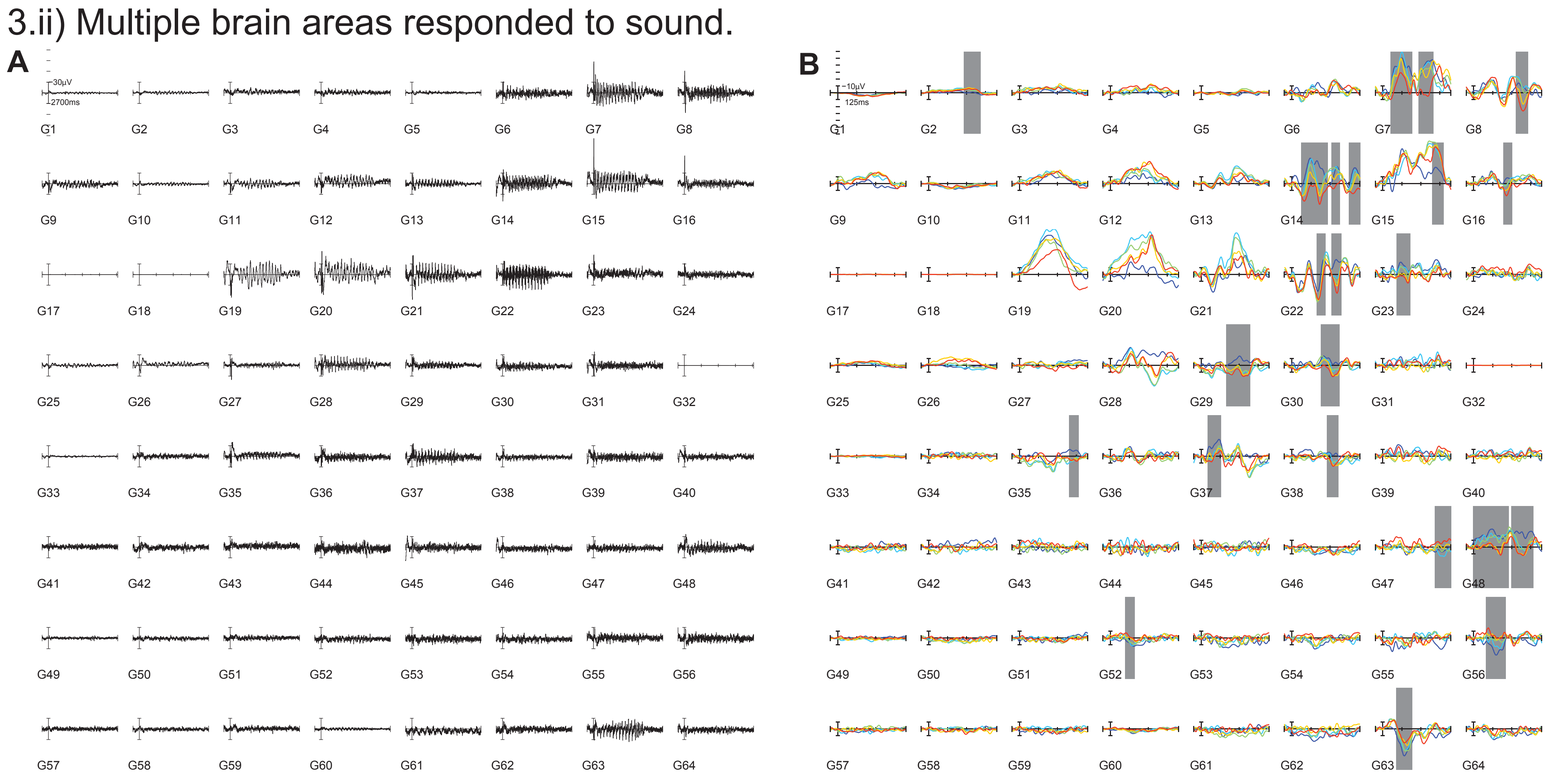
Recording and pre-processing details:
Platinum electrodes (4-mm diameter, 2.3-mm exposed, 1-cm center-to-center spacing). Referenced either to an intracranial electrode facing the inner skull table or a skull-mounted screw.
Sampling frequencies between 250 and 2000 Hz. On-line high-pass filtering at 0.1 Hz. Off-line notch filtering at 60 Hz and harmonics. Low-pass filtering at 30 Hz for display.
Artifact rejection was performed with independent component analysis, automated absolute and gradient voltage thresholds, and visual inspection.

3. RESULTS

3.i) Patients perceive sequences as healthy subjects do.

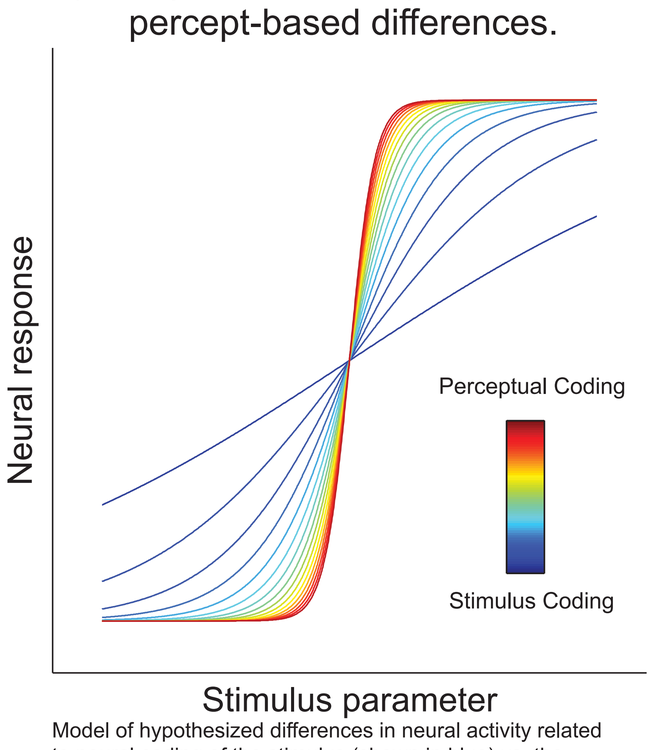


3.ii) Multiple brain areas responded to sound.

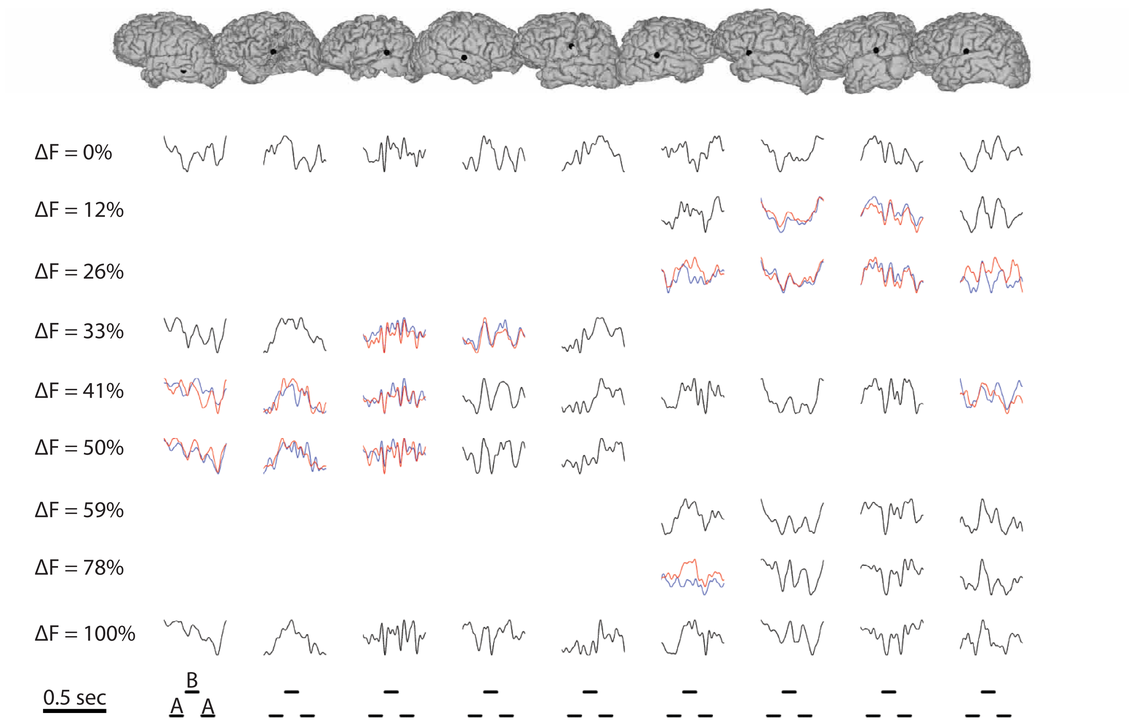


(A) Grand average response from a single patient collapsed across all conditions and time-locked to the onset of the stimulus. Each panel shows the response from an invidual electrode. The entire grid covered frontotemporoparietal regions of the right hemisphere. (B) Triplet-locked responses across stimulus conditions for the same patient. The color traverses from small ΔF (in blue) to large ΔF (in red). Significant cluster-based correlations are indicated by shaded patches.

3.iii) Triplet-locked responses over or near the superior temporal gyrus correlated with ΔF, but did not show percept-based differences.

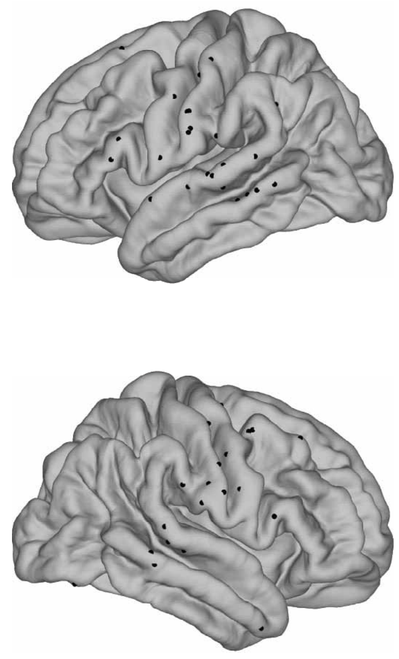


Model of hypothesized differences in neural activity related to neural coding of the stimulus (shown in blue) vs. the percept (shown in red). If activity measured in a particular brain area shows a unimodal distribution, then that area is likely to be involved in processing aspects of the stimulus. If, however, the response shows a bimodal distrubtion, especially for stimuli that produce bistable percepts, than that activity likely relates to the percept directly.



Responses from individual electrode sites over or near the superior temporal gyrus in individual patients. When the percept was bistable, the responses were binned into 1- and 2-stream conditions (blue and red, respectively). Otherwise, traces are black. Black dots show the electrode location - on an individual patient's cortical reconstruction - from where each response was obtained. The timing of individual tones is shown in the bottom panel.

3.iv) Triplet-locked responses in brain areas outside the canonical auditory cortex also correlated with ΔF.



Summary of electrodes showing significant correlations with ΔF across all patients and all electrode sites studied.

4. CONCLUSIONS

4.i) Patients were able to do the task and, like studies conducted in healthy participants, heard one stream when ΔF was small and two when it was large.
4.ii) Though variable across patients, stimulus-specific neural correlates of ΔF were observed in diverse brain areas including, but not limited to, the superior temporal gyrus.
4.iii) Unlike previous studies using non-invasive methods (EEG, MEG), evoked waveforms were highly variable both within and across brain areas.
4.iv) Surprisingly, we observed no percept-based differences in neural activity in any part of the brain in the absence of physical stimulus differences.
4.v) Our results call into question previously-proposed models of auditory streaming and suggest that the neural correlate of bistability with these stimuli are likely to be found on a finer spatiotemporal scale than that observed in the present study. Animal models of auditory streaming may prove essential.

REFERENCE

Bidet-Caulet A, Bertrand O (2009) Neurophysiological mechanisms involved in auditory perceptual organization. Front Neurosci 3(2):182-191
Bregman AS (1990) Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press, Cambridge, MA, USA
Cherry EC (1953) Some experiments on the recognition of speech, with one and two ears. J Acoust Soc Am 25: 975-979
Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25:5382-5388
Micheyl CM, Carlyon RP, Gutschalk A, Melcher JR, Oxenham AJ, Rauschecker JP, Tian B, Wilson EC (2007) The role of auditory cortex in the formation of auditory streams. Hear Res 229(1-2):116-131
Micheyl CM, Tian B, Carlyon RP, Rauschecker JP (2005) Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 48(1):139-148
Snyder JS, Alain C, Picton TW (2006) Effects of attention on neuroelectric correlates of auditory stream segregation. J Cogn Neurosci 18:1-13
Snyder JS, Alain C (200
Wilson EC, Melcher JR, Micheyl C, Gutschalk A, Oxenham AJ (2007) Cortical FMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming. J Neurophysiol 97(3):2230-2238