PNAS
www.pnas.org

Supplementary Information for

Responses in area hMT+ reflect tuning for both auditory frequency and motion after blindness early in life

Elizabeth Huber, Fang Jiang, Ione Fine

Elizabeth Huber
Email: ehuber@uw.edu

This PDF file includes:

Supplementary Information Text

Methods

Auditory stimulus presentation

Auditory stimuli (see below) were generated in MATLAB using custom software and the Psychophysics Toolbox (www.psychtoolbox.org) and were delivered via insert earphones (MR compatible Sensimetrics S14). All stimuli were presented at a sampling rate of 44.1 kHz.

Frequency calibration was carried out in a standard two-step process. We began by adjusting intensity as a function of frequency to ensure flat frequency transmission (equal amplitudes at the ear canal) from 100 Hz to 8 kHz, thereby compensating for the acoustic frequency transmission profile of our particular earphones.

After this first stage of sound system calibration, stimulus sound intensities were adjusted according to an International Standard equal-loudness curve created for insert earphones (ISO 226) to approximate equal perceived loudness across frequency. Actual sound intensities (65–83 dB SPL) were matched to the perceived loudness of a 1 kHz tone (reference frequency) at 70 dB SPL.

Acoustic noise from the scanner was attenuated by expanding-foam ear-tips, as well as protective earmuffs placed over the ear following earphone insertion. Subjects confirmed that all tones were indeed clearly audible, and of roughly equal loudness across the frequency range.

Stimuli and task

Stationary stimulus

The *stationary* stimulus contained blocks of sinusoidal pure tones (with ramped onset and offsets), which varied in frequency (88-8000 Hz, sampled in half octave steps), originally developed by (1).

Each 2s stimulus block contained eight pure tone bursts of the same frequency. Individual bursts were randomly assigned a duration of 50 or 200 ms (10 ms onset/offset ramping), with an interstimulus interval of 50 ms. Tone durations were alternated in pseudo-randomized order, switching durations at least 4 times during each 2s block, resulting in a “Morse code” like pattern of long and short tones which served to increase the perceptual salience of the stimuli over the regular pattern of background scanner noise.

Each scan consisted of a sequence of 3-4 ascending or descending (in half-octave steps) blocks, with, with the starting frequency and sequence (ascending vs. descending) selected pseudo randomly at the beginning of each sequence. The pseudo randomization ensured that the distribution of presented frequencies between 88-8000Hz was uniform, and that there were equal numbers of ascending and descending sequences.

Motion stimulus

Because the perception of auditory motion is weak for pure tone stimuli, the *moving* stimulus consisted of band pass noise bursts centered at one of 7 frequencies (125-3952 Hz, stimulus bandwidth of approximately 0.8 octaves).

The band pass noise bursts were generated by filtering noise in the Fourier domain (using MATLAB) and adding a ramped (10ms) onset and offset. We then simulated the ITD cues associated with constant-velocity 30m/s motion along a frontoparallel direction (a straight-line oriented perpendicular to the listener's facing direction) at a distance of 10 m, centered at the midline. Simulated ITDs ranged from 0 to 0.42 ms.

Each block lasted 2s and contained a pair of 1s band pass moving noise bursts. The pair of bursts were always identical and were treated as a single event. Each block was presented in a pseudo-randomized (across frequency and direction of motion) order. Randomization constraints were set to ensure that each frequency and direction of motion was presented the same number of times. Noise bursts were centered on the following frequencies: 100, 177, 316, 562, 1000, 1778, and 3162 Hz.

Traditional ascending/descending sequences lead to biases in pRF estimation due to hemodynamic blurring, neural adaptation, and subject expectations, among other factors (7,5). A full randomization of individual frequencies is optimal but results in relatively poor signal to noise. This motivated the use of short ascending/descending blocks where the beginning of each sequence was pseudorandomized. Randomization constraints were set to ensure even coverage of the entire frequency range, and to ensure that each starting point differed from the end point of the previous sequence. This provided reliable estimates with reasonable signal to noise, without the bias in pRF fits toward low and high frequencies that are observed using traditional ascending/descending designs (5). All of the relevant code is now publicly available on GitHub.

Visual localizer

In sight recovery subjects we defined hMT+ with a traditional moving vs. stationary dot-field localizer. Individual dots were relatively large (subtending 1 degree in diameter) and presented at full contrast to ensure that the stimulus was visible for both of the sight-recovery subjects, despite their acuity limitations. Voxels selective for visual motion were selected based on a threshold of qFDR < 0.05.

MRI acquisition

Functional magnetic resonance images were acquired with a 3T Phillips Achieva scanner (Philips, Eindhoven, Netherlands) at the University of Washington Diagnostic Imaging Sciences Center (DISC) using a 32-channel head coil. Foam padding minimized head motion. Overhead lighting was dimmed throughout MRI acquisition. During auditory sequences, subjects were instructed to keep their eyes closed and this was monitored through a closed-circuit camera system. Each session lasted approximately 1 hour.

In the *stationary* session, four functional scans were acquired using a standard EPI sequence (36 slices, TR/TE = 2000/25 ms, no slice gap). After discarding the first 5 timeframes, each scan consisted of 255 volumes at an effective voxel size of 3 mm isotropic. In the *moving* session, 4 functional scans were acquired using a standard EPI sequence (30 slices, TR/TE = 2000/25 ms, no slice gap). After discarding the first 5 timeframes, each scan consisted of 144 volumes at an effective voxel size of 2.75 x 2.75 x 3.00 mm (3mm in-plane).

The slight modification to the acquisition was motivated by a scanner upgrade. Importantly, the slight change in voxel size for the moving sequences maintained similar signal-to-noise, so differences in results between the 'static' and 'moving' sequences cannot be attributed to either the upgrade or the change in acquisition protocol. Additionally, all subjects were run on the same sets of scans, so group differences cannot be attributed to differences between the two acquisition sequences.

Anatomical voxel selection

For each subject, anatomical regions of interest (ROIs) were selected from the partially inflated left and right hemisphere cortical surface meshes using drawing tools within BrainVoyager QX (version 2.3.1 Brain Innovation B.V., Maastricht, The Netherlands).

In auditory cortex, ROI borders were drawn generously to include all voxels within a contiguous region between the lateral border on the crown of the superior temporal gyrus, the medial border within the fundus of the lateral sulcus, the posterior border of the supramarginal gyrus, and the anterior border of the most anterior portion of the temporal lobe.

hMT+ was defined using the Jülich probabilistic atlas thresholded at 25%. To ensure proper alignment to individual subject anatomical scans, an ROI based on the Jülich atlas was first converted into a surface patch at the gray-white matter border of the standard MNI anatomical image. Each subject's individual anatomy was then aligned to the standard MNI surface using cortex-based alignment procedures in BrainVoyager QX. The resulting inverse transformation file was used to project the Jülich atlas definition onto each individual subject's cortical surface. This means that hMT+ location, by definition, could not differ across the subject groups with reference to cortical folding, which is established as the best way to identify hMT+ location given its variability with respect to Talairach/MNI coordinates (3).

All surface-defined ROIs were mapped back into the brain volume and expanded to include voxels from -1 to 3 mm around the gray-white matter boundary. Preprocessed time-course data for each 3D anatomical voxel within the volume ROIs were then exported to MATLAB for further analysis.

Pre-processing

Standard pre-processing of fMRI data was carried out using BrainVoyager QX software, including 3D motion correction and high-pass filtering (cut-off: 3 cycles per scan). Functional data were aligned to the T1-weighted anatomical image acquired in the same session (MPRAGE, 1 × 1 × 1 mm^3). The BrainVoyager QX automatic segmentation routine was used to reconstruct the cortical surface at the white–gray matter border (with hand-editing to minimize segmentation errors) and the resulting smooth 3D surface was partially inflated.

Population Receptive Field Analysis

For each voxel within an anatomically defined ROI, we assumed a one-dimensional Gaussian sensitivity profile on a log auditory frequency axis. Using custom software written in MATLAB, we found, for each voxel, the center (f_0, best frequency) and standard deviation (σ) of the Gaussian that, when multiplied by the stimulus over time after convolution with a canonical hemodynamic response function (4), produced a predicted time course that best correlated with the measured time course within each voxel.

More specifically, the pRF analysis begins with a definition of the stimulus time course, *s(f, t)*, provided to the model as a matrix of binary values marking the presence or absence of auditory stimulation over frequency and time. To generate a hemodynamically blurred stimulus time course, *r(f, t)*, the stimulus time course was convolved with the estimated hemodynamic response function modeled as a gamma function with parameters tau and delay optimized for individual subjects as described elsewhere (5, briefly, we iterated between fitting a subject specific tau and delay using a subset of voxels, and f_0 and σ for each individual voxel).

We estimated the population response using a one-dimensional Gaussian function, *g(f)*, defined over log frequency. The frequency center (f_0) corresponds to the best frequency, while the standard deviation (σ) was used to estimate bandwidth by transforming the values into octaves and then calculating the full width half maximum (FWHM) of the Gaussian function. To create the predicted time series, we calculated the linear sum of the overlap between the input stimulus (after hemodynamic blurring) with the Gaussian receptive field.

Model fits for each voxel were obtained in two stages. We began with a grid search to find values that maximized the correlation between the predicted and actual fMRI time-courses. The initial grid search parameters for frequency centers (f_0) spanned the range of the stimulus (88 to 8000 Hz for the stationary stimulus, 100-3162 Hz for the moving stimulus), and initial standard deviation (σ) values ranged from 0.5 to 4. The best fitting parameters from this grid search were then used as initial parameters for a nonlinear search algorithm (Matlab's fminsearch function) which uses unconstrained nonlinear minimization to find the pRF model parameters f_0 and σ that maximize the correlation between the pRF predicted time-series and the BOLD data.

After fitting, voxels that met the following criteria were retained for further analyses: (1) the correlation between the observed fMRI time-course and the time-course predicted by the best-fitting pRF (our goodness-of-fit index) was higher than 0.20, (2) the center (f_0) of the best fitting pRF fell within the range of tested values (88-8000 Hz for the stationary stimulus, or 100-3162 Hz for the moving stimulus), and (3) the standard deviation (σ) of the best fitting pRF fell between the minimum inter-tone interval and the full presented frequency range, in log frequency.

Statistical Analysis – pRFs

When assessing whether a voxel was tuned for auditory frequency, we used a leave-one-out cross-validation procedure, in which we trained the pRF model using all but one scan for each subject, and then calculated the correlation between the predicted and obtained time-courses for each left-out scan. Voxels with a mean cross-validated correlation of $r > 0.2$ were categorized as 'frequency tuned'.

The expected false-positive rate for a cross-validated correlation of $r > 0.2$ was estimated using a stimulus-label permutation null model created using a randomized stimulus representation. To preserve the temporal structure within individual stimulus blocks, we randomized the stimulus block order, while retaining the ascending or descending series structure within each block. We then fit pRFs (as described above) using this randomized stimulus representation. Null distributions generated using 1000 iterations per subject were used to calculate the false-discovery rate for a threshold of $r > 0.2$.

Within hMT+, collapsed across all subjects, this stimulus-label permutation generated a mean false discovery rate of 5.2% with an upper 5% confidence bound of 5.35%; i.e. only 5% of permutations resulted in a false-positive rate greater than 5.35%). We report significant frequency tuning within a given ROI only where the number of frequency tuned voxels (cross-validated model fit with $r > 0.20$) was greater than this upper confidence limit of 5.35%.

Statistical Analysis – Group Differences

Group differences were assessed using a non-parametric approach, since we had small subject numbers and could not be assured that our data met the assumptions of parametric tests.

Statistical significance of group comparisons was estimated using the Wilcoxon rank sum test. All comparisons were also carried out using conventional parametric tests, which yielded compatible results.

Comparisons of frequency distributions across groups were performed using a bootstrapped $\chi 2$ test of independence (6). The significance of the $\chi 2$ value was estimated using a bootstrapping procedure implemented with custom MATAB software, which re-estimated $\chi 2$ after randomly assigning subjects across groups (sight-recovery, blind, sighted-control).

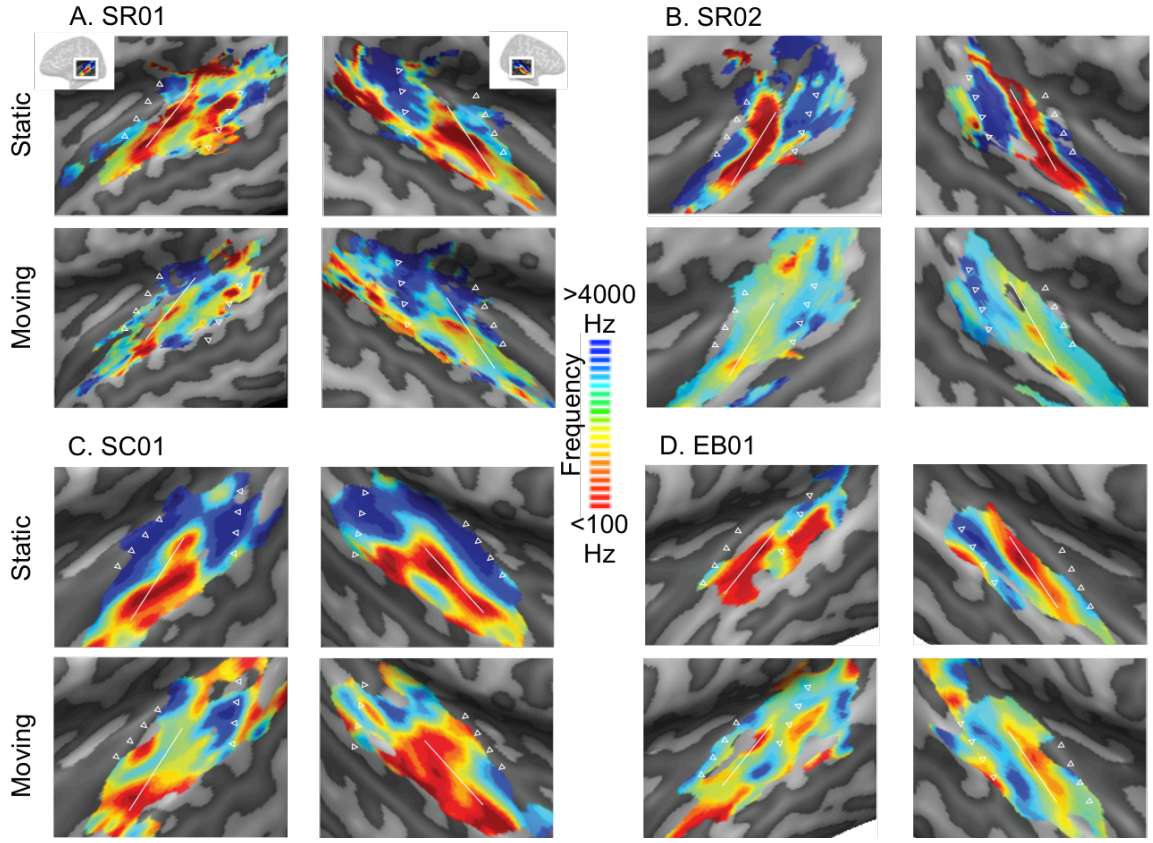

Fig. S1. Supplementary Figure 1. Tonotopic maps in primary auditory cortex. Auditory responses for both the stationary and the moving stimuli are shown for both sight-recovery subjects (Panels A & B), and example sighted (Panel C) and early blind (Panel D) subjects. The estimated anterior/posterior boundary of PAC is defined with a black dashed line for each subject. A solid black line marks the location of Heschl's gyrus. To maximize visual similarity, given that the motion stimulus had a smaller frequency range (100-3162 Hz) than the stationary stimulus (88-8000 Hz), the color map is restricted to the frequency range of the moving stimulus. Voxels with frequency centers greater than 4000 Hz have the same color-coding. PRFs were estimated using the full set of runs, and voxels above a threshold of 0.2 (not cross-validated) were included.

SUBJECT	HEMISPHERE	
	RH	LH
SR01	33.03%	39.23%
SR02	67.26%	59.155%
EB01	46.18%	55.02%
EB02	31.27%	25.34%
EB03	2.15%	7.28%
EB04	28.19%	36.48%
SC01	32.02%	29.95%
SC02	58.15%	64.85%
SC03	37.83%	48.18%
SC04	37.13%	26.25%

Table S1. Percentage of frequency tuned voxels within PAC, estimated using the moving stimulus. The PAC ROI was defined using a combination of anatomy and the frequency map based on the stationary stimulus (see 1). In the case of the 2 sighted controls for whom we did not collect data on the static stimulus, PAC was defined using frequency maps based on the moving stimulus.

References

1. Da Costa S*, et al.* (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci* 31(40):14067-14075.
2. Jiang F, Stecker GC, & Fine I (2014) Auditory motion processing after early blindness. *J Vis* 14(13):4.
3. Dumoulin SO*, et al.* (2000) A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex* 10(5):454-463.

4. Boynton GM, Engel SA, Glover GH, & Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16(13):4207-4221.
5. Thomas JM*, et al.* (2015) Population receptive field estimates of human auditory cortex. *Neuroimage* 105:428-439.
6. Bock AS*, et al.* (2015) Resting-State Retinotopic Organization in the Absence of Retinal Input and Visual Experience. *J Neurosci* 35(36):12366-12382.
7. Binda P, Thomas JM, Boynton GM, Fine I (2013) Minimizing biases in estimating the reorganization of human visual areas with BOLD retinotopic mapping. J Vis 13:13.