Characterizing Response to Elemental Unit of Acoustic Imaging Noise: An fMRI Study

Gregory G. Tamer Jr.,
The Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907 USA

Wen-Ming Luh, and
The Functional MRI Facility, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892 USA

Thomas M. Talavage [Member, IEEE]
The Weldon School of Biomedical Engineering, and the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA

Gregory G. Tamer: gtamer@purdue.edu; Wen-Ming Luh: luhw@mail.nih.gov; Thomas M. Talavage: tmt@purdue.edu

Abstract

Acoustic imaging noise produced during functional magnetic resonance imaging (fMRI) studies can hinder auditory fMRI research analysis by altering the properties of the acquired time-series data. Acoustic imaging noise can be especially confounding when estimating the time course of the hemodynamic response (HDR) in auditory event-related fMRI (fMRI) experiments. This study is motivated by the desire to establish a baseline function that can serve not only as a comparison to other quantities of acoustic imaging noise for determining how detrimental is one's experimental noise, but also as a foundation for a model that compensates for the response to acoustic imaging noise. Therefore, the amplitude and spatial extent of the HDR to the elemental unit of acoustic imaging noise (i.e., a single ping) associated with echoplanar acquisition were characterized and modeled. Results from this fMRI study at 1.5 T indicate that the group-averaged HDR in left and right auditory cortex to acoustic imaging noise (duration of 46 ms) has an estimated peak magnitude of 0.29% (right) to 0.48% (left) signal change from baseline, peaks between 3 and 5 s after stimulus presentation, and returns to baseline and remains within the noise range approximately 8 s after stimulus presentation.

Index Terms

Acoustic noise; auditory system; biomedical image processing; magnetic resonance imaging; modeling

I. Introduction

The acoustic imaging noise associated with the process of acquiring blood-oxygenation-level-dependent (BOLD, [1]) contrast functional magnetic resonance imaging (fMRI, [2]–[5]) data has previously been demonstrated to be a complicating factor when conducting experiments involving auditory stimulation [6], [7]. Complications arise both from acoustic masking, which makes it difficult to hear the presented stimulus [8], and from hemodynamic responses (HDRs) induced by the presence of the additional acoustic energy in the subject environment, which

may obscure or preclude observation of a stimulus-induced HDR [9]–[14]. Complications arising from interaction of the HDR induced by the acoustic imaging noise with the HDR induced by a desired auditory stimulus are expected to be greatest when mapping auditory cortex response properties (e.g., [12], [15], [16]).

Sources of acoustic imaging noise associated with fMRI include flexion of the gradient coils [17], eddy currents [18], RF transmit and slice-selection pulses [18], and ambient noise—the air-handling system, the ventilation fan, and the liquid helium condenser facilitating supercooling of the permanent electromagnet. These sources, except for the ambient noise, combine to produce approximately 70 ms of noise per gradient echo slice acquisition at 1.5 T, with the 40–50 ms flexion of the gradient coils and imager bore proving to be the dominant source of acoustic energy. As current passes through the coils to create the rapidly switched gradient fields within the main (static) magnetic field, Lorentz forces act on the coils to produce loud, distinct "pings" with each image acquisition. The acoustic intensity of the pings has historically ranged from 100 to 130 dB sound pressure level (SPL, [19], [20]), but recent generations of imagers have vacuum-potted coils that can limit sound levels to a typical range of 90–100 db SPL [21]. Note that the audible "clicks" produced by the RF transmit and application of other repetitive gradients (but not at acoustic frequencies of 20–20 000 Hz) produce a small fraction of the total acoustic energy and are additionally likely to be masked in rapid acquisition sequences due to the prolonged duration (up to 200 ms) of forward masking phenomena [22], leaving the ping as the dominant noise source to be characterized.

Strategies, both interventional and preventative, have been developed to reduce the acoustic imaging noise incident on the ear of the subject, but they are limited in the extent to which the residual noise energy may be considered inconsequential. Passive attenuation measures, such as earplugs and noise-rated headphones, are routinely used, decreasing the intensity of the pings by 30–40 dB [23] to 50–85 dB SPL (comparable to a hair dryer). Active attenuation [24], [25] has not yet proven to be the answer, as the high-frequency content of the spectrum of the acoustic imaging noise has proven difficult to cancel, even when using commercial noise-cancellation headphones [26].

Preventative measures can be employed to minimize the effect of the acoustic imaging noise. Use of a clustered volume acquisition (CVA, [11]), also known as "sparse sampling" [27], can be used to provide quiet periods in which the auditory stimuli can be presented in isolation. Also, for extremely long sampling periods (e.g., repetition times, TRs, of 20 s or greater)—the stroboscopic event-related method [28] and also known as sparse temporal acquisition (STA) —can be used to reduce the integrated amount of acoustic imaging noise and provide a better signal-to-noise ratio [29].

However, two complications exist with the use of a CVA and extremely long sampling periods. First, the reduced sampling rate limits the statistical power of the experiment. It is demonstrated in [11] that despite a higher percent signal change with a TR of 8 s than with a TR of 3 s, the statistical power per unit time is greater for a TR of 3 s than for a TR of 8 s. In this case, the reduced sampling rate with a TR of 8 s offsets the benefit of the increased percent signal change.

This limitation of statistical power can be overcome with interleaved silent steady-state (ISSS) imaging [30] or by using a clustered temporal acquisition (CTA, [9], [31], [32]). In both these techniques, a cluster of volume acquisitions instead of a single acquistion are acquired after each stimulus presentation (note that ISSS imaging avoids and CTA accounts for T1-related signal decay that occurs with a clustering of volume acquisitions). These designs permit more sampling than a simple STA in the same amount of time, thus increasing the statistical power.

However, the second complication associated with extremely long sampling periods occurs when measuring the latter portion of the HDR. This typically requires the presentation of a

stimulus soon after the previous volume acquisition (e.g., measuring the response 15 s after stimulus onset with a TR of 20 s requires the presentation of the stimulus only 5 s after the previous volume acquisition). The response to the previous volume acquisition can elevate the baseline signal level, most likely resulting in an overestimate of the latter portion of the HDR relative to the early portion, for which the response to the preceding volume acquisition will have decayed prior to stimulus presentation.

Therefore, while ISSS imaging and CTA techniques help improve the statistical power, there are still concerns about 1) the interaction between the response to the previous volume acquisition and the response to the stimulus and 2) the first volume acquisition in the cluster of volume acquisitions evoking a response that affects the measurement of the latter volume acquisitions in the cluster (especially if the cluster duration is longer than a few seconds). The designs presented in [30]–[32] either ignore these interactions or limit their sampling of the HDR to just the expected locations of the peak of the HDR. These designs can be effectively employed, but to overcome the interaction of the noise and the stimulus, the acquisitions still need to be spaced far apart, and the statistical power of the design is decreased.

Thus, in spite of the successful application of techniques to limit the effects of acoustic imaging noise, no combination of current strategies can overcome the interaction of the acoustic imaging noise and the desired acoustic stimulus without severely limiting the statistical power of the associated experiment relative to the statistical power that can be achieved with standard event-related and rapid-presentation-event-related designs.

The loss of statistical power due to extended TR values is of greatest concern due to the attractiveness of employing event-related fMRI (ER-fMRI) paradigms that measure both the amplitude and shape (delay and spread) of the HDR to one or more presented stimuli [33], [34]. When individual stimuli are 1) presented at sufficiently long intervals (e.g., 15–20 s) to prevent overlap of corresponding HDRs and 2) there is a minimal amount of interaction with acoustic imaging noise (e.g., a visual experiment), the response to each individual stimulus is accurately estimated. Models of the HDR (e.g., gamma-variate function [35]) can then be fit to the response in a single voxel or region of interest and can be used to detect cortical activity throughout the brain by comparing the modeled HDR to the functional data and identifying areas exhibiting a high correlation with the reference waveform. As with any fit, the potential accuracy increases with the number of sample points. In ER-fMRI, increasing the number of sample points implies more frequent sampling of the HDR through use of a shorter TR. Unfortunately, shortening the TR can be detrimental for auditory studies due to the accompanying increase in acoustic imaging noise.

It is unlikely that auditory stimulus-induced HDRs can always be well-described by traditional HDR models that assume the vascular system is 1) at rest at the time of presentation; 2) linear; and 3) time-invariant. Coincident presentation of both noise and desired stimuli would be expected to alter the HDR observed via fMRI relative to presentation of the stimulus alone. Alterations could arise from both acoustic and vascular sources. The composite acoustic signal (i.e., stimulus plus noise) could produce acoustic masking at the level of the cochlea, altering the stimulus representation reaching cortex, potentially generating HDRs in regions containing neurons that otherwise would not respond to the desired stimulus. Within regions that would be expected to respond to the stimulus, the additional noise signal can generate unintended HDRs likely to combine with the HDR induced by the (perceived portion of the) intended stimulus in a nonlinear fashion, resulting in an observed response that is not equivalent to the summation of the HDRs induced by the two acoustic signals in isolation (e.g., [13]).

The interaction of HDRs from noncoincident presentation of the acoustic imaging noise and the desired stimulus can also produce nonlinear summation of observed HDRs. In such a case,

the desired stimulus is unlikely to be masked, permitting the auditory pathway to deliver an "ideal" representation to the auditory cortex. However, the long duration (15 s or longer) of an HDR is likely to mean that the vascular system has been perturbed from rest by preceding acoustic imaging noise at the time an HDR is induced by the desired stimulus. If the vascular system is not "at rest" at the time an HDR is induced, physical limitations may limit changes in cerebral blood flow (CBF), cerebral blood volume (CBV), and the metabolic rate of consumption of oxygen (CMRO2) (e.g., [36]), resulting in potential nonstationarity of the HDR, relative to that which is expected (e.g., [37], [38]).

It is this concern regarding nonstationarity of the response to a desired stimulus, dependent upon the state of the vascular system at the time of stimulus presentation, that motivates the detailed study of responses to the acoustic imaging noise (to assess how "at rest" the system might be under given circumstances) and how the noise-induced HDR interacts with HDRs induced by a given auditory stimulus. This work focuses on the initial implementation of a framework for the characterization of the HDR associated with acoustic imaging noise, with its primary objective being to characterize and model the amplitude and spatial extent of the HDR to a single unit of acoustic imaging noise (i.e., the "ping") induced by the gradient readout process.

This study at 1.5 T was conducted to estimate the HDR time course and to characterize the HDR amplitude and spatial extent arising from the elemental unit of acoustic imaging noise— a single ping—generated by normal operation of the gradient coils and presented at least 15 s after previous volume acquisitions. The obtained estimate arguably represents the "pure" response in auditory cortex to a genuine single ping—free of residual responses both from previous pings (i.e., stimulus presentations) and volume acquisitions. This estimate shows good agreement with the noise burst response measured by [39] and argues for nonlinearity of HDR duration when compared to the findings of [40].

II. Materials and Methods

A. Acoustic Stimulus

A single-ping stimulus (approximately 46 ms in duration), presented in the quiet period between volume acquisitions (CVA sequence), was generated by performance of a standard gradient echo, blipped, echo-planar imaging (EPI) readout sequence without application of an RF pulse—i.e., a standard slice acquisition was effected with a flip angle of zero degree, so the recovery of longitudinal magnetization (i.e., T_1 relaxation) was not perturbed. The spectrum of the stimulus is presented in Fig. 1.

The timing of each ping presentation was dependent on two factors. First, it was desired that presentation of the stimulus occurs after HDRs arising from all previous acoustic stimulation had returned to baseline. The return to baseline from the volume acquisition was estimated at 15 s, which served as the minimum postvolume acquisition delay prior to presentation of the stimulus. Second, the time at which the ping was presented was varied with respect to the subsequent volume acquisition to permit estimation of the HDR to the ping as a function of time (see Fig. 2). It was conservatively assumed that the HDR to a single ping acoustic stimulus would return to and remain at baseline within 15 s of stimulus onset. Therefore, a minimum TR of 30 s plus the duration of the volume acquisition was required to completely isolate the presented ping from preceding acoustic stimulation while also permitting effective sampling of the full duration of the HDR associated with the ping.

B. Subjects

Eleven normal-hearing subjects (seven males, four females), ages 21–33, participated. All subjects were in good health and gave informed written consent. There was no assessment of gender- or age-specific effects.

C. Task

Subjects were instructed to attend to a visual-only stimulus (a movie without sound delivery) to sustain arousal and to minimize focus on the acoustic imaging noise (the single-ping stimulus and the actual volume acquisitions). Arousal (i.e., attention) was not measured or verified. Measurement of attention while subjects focus on the acoustic imaging noise would likely overestimate the interaction since the focus in a typical auditory fMRI experiment is on the stimulus rather than the acoustic imaging noise.

D. Image Acquisition

All subjects underwent imaging on a 1.5 T GE Signa LX imager, equipped with "standard" body gradients (120 mT/m/ms maximum slew rate) and located in the Department of Radiology at the Indiana University School of Medicine (IUSoM; Indianapolis, IN).

All subjects were fixed in place by straps, padding, and a frame to minimize movements. Subjects were connected to a pneumatic acoustic delivery system by tubes passed through the earmuffs of the surface coils and through earplugs to replicate "normal" conditions in which a stimulus, were one to be presented, would input directly to the ear canals.

All functional data were acquired using auditory surface coils, as described by [41], and all sessions used the same functional imaging parameters: TR = 31.5 s, echo time (TE) = 40 ms, flip angle = 90°, field of view = 20 cm × 20 cm, in-plane resolution = 3.125 mm × 3.125 mm. Fifteen axial slices (5 mm thickness), centered on Heschl's gyrus, were acquired to encompass at least the superior–inferior extent of the Sylvian fissure to ensure imaging of both primary and secondary auditory cortices [42]. These values were chosen to maximize the signal (which will recover fully given the use of a long TR) and the signal-to-noise ratio within brain voxels [43]. The desired 15-slice volume was acquired using a CVA, resulting in approximately 1.5 s of acoustic imaging noise at the beginning of each TR period. A session consisted of six experimental runs, each lasting 504 s.

In a separate imaging session for each subject, a birdcage head coil was used to obtain a 3-D volumetric image of the entire brain (124 slices; in-plane resolution = 0.9375 mm × 0.9375 mm, slice thickness = 1.2 mm). This anatomical scan was used to facilitate identification of primary auditory cortex and other regions, enabling group averaging and comparisons in these areas. Note that this additional acquisition was necessitated by the use of bilateral surface coils for the functional runs; the bilateral surface coils produce minimal imaging signal at the midline, one of the most critical locations for accurate transformation to standardized coordinates.

E. Experimental Paradigm

Each experimental run comprised 16 volume acquisitions preceded by one null acquisition in which the first stimulus presentation took place. One stimulus presentation occurred in each of the 15 s intervals immediately prior to a true volume acquisition. The 16 stimulus conditions were pseudorandomly selected from a set comprising eight delays: {0,2,3,4,5,6,8,10} s. There was no stimulus presentation for the delay time, $t = 0$ s, and therefore, it serves as the control for the experiment. Each of the eight values was used twice for each of the six runs, resulting in 12 trials per delay (i.e., per sampled time point of the measured HDR).

The use of this set of delays was based on the analysis of preliminary data from three sessions conducted on a 1.5 T GE Signa CVi imager at the Medical College of Wisconsin (Milwaukee, WI) [44], in which it was observed that the estimated HDR to the single-ping stimulus had returned to baseline by 10 s after stimulus presentation, and remained at baseline (i.e., within the noise range) out to the 15 s measurement.

F. Experimental Control

To verify that systematic and experimental factors (e.g., the variable delay between the single-ping stimulus and subsequent volume acquisition) did not produce a false response (e.g., eddy currents in gradient coils changing measured signal levels), one functional run was performed on a phantom. A birdcage head coil was used, and the imaging parameters were the same as used for the functional sessions.

G. Data Analysis

A standardized processing procedure was applied to the event-related functional data acquired for a given subject. After initial image reconstruction, a 3-D registration (AFNI, [45]) aligned all experimental runs acquired in a given session to the first run of that session (i.e., the run acquired closest in time to the high-resolution dataset). After this intrasession registration, data from one session were discarded due to excessive head motion (up to 2 mm) within each run, sometimes occurring between successive time points (believable due to the long TR employed). Therefore, data from ten functional sessions were analyzed and evaluated.

The registration step permitted intrasession averaging of trials across experimental runs and permitted conversion to a standardized coordinate system—a procedure for which minimal registration error is required between the imaged volume and the high-resolution images. To facilitate this conversion, the last preprocessing step involved registration of the high-resolution images to the 3-D volumetric dataset acquired with the birdcage head coil. A combination of FreeSurfer (CorTechs and the Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA) and AFNI was used for registration. The resulting rotation, translation, and scaling matrix was then applied to the experimental runs using in-house software to align them with the 3-D volumetric dataset.

The novel pulse sequence that was implemented for this experiment acquired the first volume acquisition of each run with a slightly longer TE (42 ms) than specified (40 ms) to generate a field map that would be used to obtain higher quality reconstructions. As a result, the signal measured with the first volume acquisition of each run is lower than it should be due to the longer echo time in which more T2* decay has occurred. Theoretically, the first volume acquisition will be approximately lower in signal by a factor of the exponential of the negative ratio of the difference in TE divided by T2*. To account for this lower signal, an experimental adjustment factor was calculated for each slice of each run and then applied to the respective first volume acquisition. This experiment adjustment factor was calculated for each brain voxel by first taking the ratio of 1) the average signal intensity across all volume acquisitions except the first volume acquisition and 2) the signal intensity of the first volume acquisition. The calculated ratios of the brain voxels within one slice of one run were then averaged together to form a first-volume-acquisition adjustment factor for that particular slice and run. Each brain voxel's signal intensity was then simply multiplied by this adjustment factor. The calculated experimental adjustment factors ranged from 1.03 to 1.06, which is consistent with the theoretical expectation.

H. Statistical Maps

Group-averaged statistical maps were generated and overlaid onto Talairach-transformed anatomical datasets to permit examination of estimated cortical activity on a location-specific

basis. For each subject, the registered, motion-corrected volume files were averaged together using AFNI. Each voxel's time-series in the resulting averaged dataset was then compared with several reference waveforms by means of cross-correlation statistical tests.

The reference waveforms were generated using the gamma-variate model [35]

$$s(t) = \left(\frac{t-\delta}{\tau}\right)^2 e^{-\frac{(t-\delta)}{\tau}} u(t-\delta).$$

(1)

Parameter δ is the delay until a measurable response is observed, and τ is a parameter that controls the spread (width) of the idealized response, Note that $u(\cdot)$ in the equation denotes the unit-step function.

Reference waveforms were created using the following parameter sets: 1) $\delta = 1$ s, $\tau = 1$ s; 2) $\delta = 1.5$ s, $\tau = 1.25$ s; and 3) $\delta = 2.5$ s, $\tau = 1.25$ s. These parameter sets were selected based on HDR estimates observed using preliminary data [44] and also wanting to check for responses with delayed starts at various times. A double-gamma function was not employed to account for any poststimulus undershoot since 1) a significant undershoot was not observed in the preliminary data [44].

The statistical datasets were converted to z-scores, which were then adjusted based on the number of volume files averaged. The statistical datasets for a session were subsequently converted to Talairach coordinates. Finally, the Talairach-converted, statistical datasets for all the subjects were averaged together, and the z-scores corrected based on the number of subjects. This Talairach-converted, group-averaged statistical dataset was overlayed onto a Talairach-transformed anatomical dataset, and the correlation test z-score statistic was thresholded for display and interpretation. The threshold for a statistical map was set by first determining the single-session z-score that corresponded to a p-value of 0.05 and then multiplying this z-score by the square root of the number of sessions that had been averaged together.

I. Region of Interest (ROI) Analysis

ROIs were selected by identifying regions of significant clustered activation present in the generated statistical maps. Each ROI selected was 7 mm × 7 mm × 7 mm in size. The ROIs and Talairach (x, y, z) centers selected were: 1) left lateral TTG (−58, −16, 11 mm); 2) left medial TTG (−47, 22, 11 mm); 3) left STG (−57, −32, 11 mm); 4) right insula (39, −38, 22 mm); 5) right medial TTG (39, −29, 12 mm); 6) right lateral TTG (50, −17, 10 mm); 7) right STG (50, −27, 15 mm); and 8) right STG again (50, −27, 6 mm).

Across voxels within an ROI, an average, estimated HDR to the presented stimulus was computed. Each time point's value was calculated as a percent signal change relative to the no-stimulus condition—i.e., the magnitude value at any given time point is compared to the magnitude value at an effective post-stimulus time of $t = 0$ s. An estimated HDR for a session was generated by averaging all trials acquired for each delay. Averaged data from each session were then used to compute a group average of the estimated HDR to the stimulus.

To justify the averaging of results across runs, voxels, and subjects, it is assumed that the noise will be independent of poststimulus time measurement, and that the random placement of trials within each run should result in a comparable set of nonstationary errors for each measurement delay, thus resulting in a grand average with uniform bias over time. It is also assumed that the cortical response to the stimulus is reproducible across runs and subjects.

It is important to note that response onset delays are assumed to be approximately equal for all voxels within auditory cortex. While other regions may have different onset delays of response to acoustic imaging noise, potentially not correlating well with the reference HDRs used in this test, the primary concern is demonstrating that auditory cortex significantly responds to acoustic imaging noise and that this particular response is localized to auditory cortex.

J. Control With Phantom

For the control phantom run, two square, disjoint ROIs (16 voxels each, 1024 cubic mm) were randomly selected within the phantom, and the data were processed using the same processing stream as used for the (human) functional data. Statistical maps were not generated.

K. Hemodynamic Response Modeling

For the purpose of modeling the HDR to the single-ping stimulus and to compare HDRs in the different selected ROIs, a nonlinear regression was conducted on the group-averaged estimated HDRs for each ROI. The chosen model was a double-gamma-variate model (see [2]) since a small poststimulus undershoot was observed in the estimated group-averaged HDRs

$$s(t) = A_1 \left(\frac{t-\delta_1}{\tau_1}\right)^2 e^{-\frac{(t-\delta_1)}{\tau_1}} u(t-\delta_1)$$
$$- A_2 \left(\frac{t-\delta_2}{\tau_2}\right)^2 e^{-\frac{(t-\delta_2)}{\tau_2}} u(t-\delta_2).$$

(2)

This first part of the equation permits modeling of the general shape and strength of the main response and the second part of the equation permits the modeling of the poststimulus undershoot. A nonlinear regression was performed using MATLAB's (The MathWorks, Inc., Natick, MA) *lsqcurvefit* function to estimate A_1, δ_1, τ_1, A_2, δ_2, and τ_2 for each ROI.

L. IRB Approval

The use of human subjects for the fMRI sessions in this study has been approved by the Institutional Review Board of Purdue University and by the Institutional Review Board of the Indiana University School of Medicine.

III. Results

A. Assessment of Artifactual Activation

Fig. 3 presents the average estimated HDR for the control run obtained from two randomly placed disjoint ROIs within the phantom. To assess correlation between the signal changes and the stimulus paradigm for presumably nonactivated voxels, a cross-correlation test was conducted that compared the presented time courses, concatenated together, to a reference HDR time course created using [1] with parameters $\delta = 1.0$ s, $\tau = 1.0$ s, and $A = 1$. The calculated R-squared value is 0.24 (*p-value* = 0.18). Therefore, the systematic noise factor in the ROI signals appears to be approximately +/−0.1%.

B. Statistical Maps

Fig. 4(a)–(c) illustrates areas of activation in all three planes for the group-averaged data observed when the gamma-variate function (with $\delta = 1$ s and $\tau = 1$ s) was used as a reference waveform for cross-correlation analysis. Significant activation was not observed with the other two reference waveforms. The z-scores depicted in Fig. 4 range from 6.2 (green, $p < 0.05$) to 12.5 (red). Note that the HDR to the acoustic imaging noise is localized to auditory regions of

the cortex. The coordinates presented are Talairach coordinates, and Fig. 4(b) presents left-hemisphere slices.

Fig. 4(d)–(f) illustrates the selected ROIs. Fig. 4(d) illustrates ROIs within left lateral transverse temporal gyrus, left medial transverse temporal gyrus, and left superior temporal gyrus. Fig. 4 (e) illustrates ROIs within right insula and right medial transervse temporal gyrus. Fig. 4(f) illustrates ROIs within right lateral transverse temporal gyrus and right superior temporal gyrus (two ROIs).

C. Estimation of Hemodynamic Response

Fig. 4(g)–(h) presents the group average (with standard error bars) of the estimated HDR for eight ROIs in left (g) and right (h) auditory cortex (and surrounding areas) for the ten analyzed sessions. Auditory cortex HDRs are observed that are comparable to HDRs presented in previous auditory fMRI studies [28], [40]. The HDRs peak between 2 and 4 s and appear to return to baseline by 8 s after stimulus presentation.

D. Hemodynamic Response Modeling

Table I presents the estimated parameter values from the group-averaged nonlinear regressions performed on the data from the eight ROIs. It also includes the average of the parameters for the three left-hemisphere ROIs and the average for the five right-hemisphere ROIs.

IV. Discussion

In the assumed absence of all residual responses from previous stimuli and volume acquisitions, the HDR in left and right auditory cortex to the elemental unit (a single ping) of acoustic imaging noise has successfully been estimated for blipped-EPI fMRI at 1.5 T. The response to a single ping induces a response that has an estimated peak of 0.48% above baseline in left auditory cortex and 0.29% in right auditory cortex. The estimated response has a short delay time ($\delta = 1.3$ s for left auditory cortex and $\delta = 1.2$ s for right auditory cortex), peaks between 2 and 4 s, and returns to baseline and remains within the noise range from approximately 8 s after stimulus presentation.

A paradigm design and pulse sequence design were created in tandem to estimate the "pure" response to acoustic imaging noise. These designs are not limited to investigating estimated responses to acoustic imaging noise and can also be used to estimate "pure" responses to other auditory stimuli. The advantage, of course, is that the estimated response to the stimulus will be in isolation and will not contain any residual response from previous stimuli and volume acquisitions. The disadvantage of this experiment design, though, is that for any one session, a low number of trials are acquired per time point of the HDR, most likely requiring averaging of responses across subjects to obtain the sufficient statistical power needed for confidence in the results.

The objective of this experiment was to obtain a group-level estimate of the HDR to a single-ping stimulus rather than focus on the individual differences and variability across the particular subjects of this experiment. Due to the intentionally limited number of trials per session (subject) for each time point of the HDR and due to the assumed presence of significant systemic and physiologic noise, it was anticipated that a large variability of results across subjects would occur and that only group-averaged results would have sufficient statistical power and yield statistically significant results. This conservative approach was supported by experimental results not presented—c.f., there was significant variability of the estimated responses across subjects and the lack of significant responses for several individual subjects. The result is then a general working model of an average group-level response that can then

be used in a correction algorithm to account for the acoustic imaging noise (see the text under Section IV-C). Any accounting of the acoustic imaging noise will inevitably not be perfect, and trying to implement a model on an individual-subject basis will most likely yield minimal improvement through a significant increase in workload by the experimenter (and subjects).

A. Other Contributions to the Acoustic Imaging Noise

In addition to the "ping" produced by the gradient readout process, the full slice acquisition includes several "click" stimuli arising from short-duration current ramps (e.g., RF transmit pulse, crusher gradients). While these clicks are perceptible, it has been demonstrated that clicks of less than 20–30 ms in duration and presented at rates outside of 5–8 Hz are generally poor stimuli for eliciting cortical activation [39]. Therefore, the underlying neuronal response to these clicks is expected to be small relative to the response driven by the gradient readout process. Further, during acquisition of multiple slices in a relatively short period of time (as is typical for a CVA sequence), the clicks are most likely masked by the close temporal proximity (<20 ms) to the (loud) "ping" of the readout that will produce forward masking of brief stimuli —a phenomenon that can persist as long as 200 ms. Therefore, the focus of this paper on the "ping" component of the acoustic imaging noise is justified.

B. Comparison to Previous Work

The HDR estimate obtained for the 1-ping stimulus allows us to evaluate responses to other volume acquisition sizes and to assess the linearity of increased volume acquisition durations. Critically, to serve as a basis in this fashion, it is important to note that the obtained estimate for the response to a 1-ping stimulus is consistent with the response model presented by [39] for a noise burst stimulus with a duration of 25 ms. Both of these responses peak around 3 to 4 s postonset and return to baseline after approximately 8 s.

Linearity of increased volume acquisition durations can be assessed by a comparison to the work of [40] and their estimate of the HDR to a 16-ping stimulus. While a quantitative analysis is not appropriate given the significant differences in experimental design, image processing, and field strengths, a qualitative analysis supports the notion that the amplitude of response to acoustic image noise increases nonlinearly with volume acquisition duration. This suggests that the HDR is likely low-pass filtering the neural response to a train of pings, resulting in a cumulative HDR that is not the linear combination of responses to independent, single presentations. This qualitative analysis also illustrates the need to characterize different durations (or levels) of acoustic imaging noise since the scaling of the response to a single ping of acoustic imaging noise will not lead to accurate modeling of responses to multiple pings.

C. Implications and Future Work

The estimated HDR obtained for the 1-ping stimulus by this study serves as the first element of a basis set that will be useful for improving the quality of results obtained when using auditory stimuli in fMRI. For example, the response amplitude and duration difference observed between the 1-ping (46 ms) (from this study) and the 16-ping (1150 ms) (from [40]) stimuli may simply be due to the previously reported nonlinear behavior of responses to short- and long-duration stimuli [13], [46], [47]. Previous preliminary work [44] does indeed suggest a nonlinear behavior of responses, as estimated HDRs for 1-ping and 5-ping stimuli were found to be similar, inconsistent with a system in which superposition holds. However, further investigation with a sufficient number of subjects for each of several multiping stimuli is needed to confirm the suggested nonlinear behavior. The later return to baseline for the longer stimulus could also be an indication that late responses to the 16-ping stimulus (presented closer in time to the preceding volume acquisition) were not obtained with the auditory hemodynamic system "at rest." Answering this question will require greater understanding of how the multiple pings of a volume acquisition are processed—as a single stimulus or as a train of bursts?

Fundamentally, we must have an answer to the question as to the number of slices, and the temporal spacing of slice acquisitions at which the auditory hemodynamic system behaves as if being stimulated by multiple stimuli versus a single (albeit discontinuous) stimulus. Such knowledge will then serve as an entry point into improving our understanding of the interaction of the response to the acoustic imaging noise with that arising from a desired stimulus [13], and potentially allow us, under appropriate conditions, to tease apart these two responses.

It is critical to note that while acoustic imaging noise from blipped EPI induces a significant response in auditory cortex, results from this study do not invalidate previous auditory fMRI studies. However, it is probable that these results will be of most significant benefit for studies using similar image acquisition procedures. Due to the robust nature of the observed response to acoustic imaging noise, and the presence of some form of this noise across imager platforms, previous findings that do not account for it are most likely misestimating the HDR in auditory cortex. This misestimation can occur by either overlooking areas in which responses are being masked by interaction with the response to acoustic imaging noise, or by including areas in which responses arise primarily from acoustic imaging noise. It is important to determine how the results of this study might change as a consequence of imager field strength, manufacturer, or installation. The amplitude and frequency differences between field strengths precludes incorporating the current findings into analysis when using a 3.0 T or higher field strength. While the general shape of the HDR is not expected to change, the amplitude and localization of observed activity could shift if sufficient variation is present in both the acoustic frequency content of the "ping" and imager sensitivity. Therefore, it is probable that these results will be of the most significant benefit for studies using blipped-EPI on 1.5 T GE Signa imagers.

One possible mechanism of compensating for the noise-induced responses in the cortex is the construction of an additive or compressive model that can be incorporated into the analysis of auditory fMRI data. This is further motivated by the fact that the simple gamma-variate model used in this research cannot account for the potentially complicating factors such as the predip and the poststimulus undershoot, thus future efforts should explore the use of other HDR models. Such a model might be constructed using nonlinear regressions on estimated HDRs obtained over a variety of acoustic imaging noise conditions, including different manufacturers and field strengths.

V. Conclusion

This ER-fMRI study demonstrates that even a single, isolated presentation of acoustic imaging noise with duration of 46 ms can induce a significant HDR in right and left auditory cortex. Incorporating this with results found in previous acoustic imaging noise studies [9]–[14], [40], it is evident that current ER-fMRI models and analysis techniques for auditory studies need to be refined to predict and compensate for signal changes that arise due to the presence of the acoustic imaging noise throughout the auditory experiments. A baseline has now been established that can serve as a comparison to other quantities of noise for determining how significantly different (and worse off) from baseline may be the environment of a given experiment. These results can now serve as the foundation for a model that will compensate for the responses to acoustic imaging noise.

Acknowledgments

References

1. Ogawa S, Menon R, Tank D, Kim S, Merkle H, Ellerman J, Ugurbil K. Functional brain mapping by blood oxygenation level dependent contrast magnetic resonance imaging. Biophys J 1993;64:803–812. [PubMed: 8386018]

2. Belliveau J, Kennedy D, McKinstry R, Buchbinder B, Weisskoff R, Cohen M, Vevea J, Brady T, Rosen B. Functional mapping of the human visual cortex by magnetic resonance imaging. Science 1991;254:716–719. [PubMed: 1948051]

3. Kwong K, Belliveau J, Chesler D, Goldberg I, Weisskoff R, Poncelet B, Kennedy D, Hoppel B, Cohen M, Turner R, Cheng H, Brady T, Rosen B. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 1992;89:5675–5679. [PubMed: 1608978]

4. Bandettini P, Wong E, Hinks R, Tikofsky R, Hyde J. Time course EPI of human brain function during task activation. Magn Reson Med 1992;25:390–397. [PubMed: 1614324]

5. Ogawa S, Tank D, Menon R, Ellermann J, Kim S, Merkle H, Ugurbil K. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping using MRI. Proc Natl Acad Sci USA 1992;89:5951–5955. [PubMed: 1631079]

6. Amaro E Jr, Williams S, Shergill S, Fu C, MacSweeney M, Picchioni M, Brammer M, McGuire P. Acoustic noise and functional magnetic resonance imaging: Current strategies and future prospects. J Magn Reson Imag 2002;16:497–510.

7. Moelker A, Pattynama P. Acoustic noise concerns in functional magnetic resonance imaging. Hum Brain Mapp 2003;20:123–141. [PubMed: 14601139]

8. Shah N, Jancke L, Grosse-Ruyken M, Muller-Gartner H. Influence of acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination. J Magn Reson Imag 1999;9(no 1):19–25.

9. Bandettini P, Jesmaowicaz A, Kylen JV, Birn R, Hyde J. Functional MRI of imager acoustic noise induced brain activation. Magn Reson Med 1998;39:410–416. [PubMed: 9498597]

10. Ulmer J, Biswal B, Yetkin F, Zerrin F, Mark L, Mathews V, Prost R, Estkowski L, McAuliffe T, Haughton M, Daniels D. Cortical activation response to acoustic echo planar scanner noise. J Comput Assist Tomogr 1998;22(no 1):111–119. [PubMed: 9448773]

11. Edmister W, Talavage T, Ledden P, Weisskoff R. Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 1999;7:89–97. [PubMed: 9950066]

12. Talavage T, Edmister W, Ledden P, Weisskoff R. Quantitative assessment of auditory cortex responses induced by imager acoustic noise. Hum Brain Mapp 1999;7:79–88. [PubMed: 9950065]

13. Talavage T, Edmister W. Nonlinearity of fMRI responses in human auditory cortex. Hum Brain Mapp 2004;22:216–228. [PubMed: 15195288]

14. Langers D, van Dijk P, Backes W. Interactions between hemodynamic responses to scanner acoustic noise and auditory stimuli in functional magnetic resonance imaging. Magn Reson Med 2005;53:49–60. [PubMed: 15690502]

15. Talavage T, Sereno M, Melcher J, Ledden P, Rosen B, Dale A. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. J Neurophysiol 2004;91:1282–1296. [PubMed: 14614108]

16. Formisano E, Kim D, Salle FD, de Moortele PV, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 2003;40:859–869. [PubMed: 14622588]

17. Mansfield P, Glover P, Beaumont J. Sound generation in gradient coil structures for MRI. Magn Reson Med 1998;39:539–550. [PubMed: 9543415]

18. Edelstein W, Hedeen R, Mallozzi R, El-Hamamsy S, Ackermann R, Havens T. Making MRI quieter. Magn Reson Imag 2002;20:155–163.

19. Ravicz M, Melcher J. Isolating the auditory system from acoustic noise during functional magnetic resonance imaging: Examination of noise conduction through the ear canal, head, and body. J Acoust Soc Amer 2001;109(no 1):216–231. [PubMed: 11206150]

20. Price D, De Wilde J, Papadaki A, Curran J, Kitney R. Investigation of acoustic noise on 15 MRI scanners from 0.2 T to 3 T. J Magn Reson Imag 2001;13:288–293.

21. Tseng, G.; Talavage, T.; Hinks, R. Repeatability and variability of noise generated during MRI. Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.; San Francisco, CA. 2004. p. 1046-1099.

22. Viemeister, N.; Plack, C. Time analysis. In: Yost, W.; Popper, A.; Fay, R., editors. Human Psychophysics. New York: Spring-Verlag; 1993. p. 116-154.

23. Ravicz M, Melcher J, Kiang N. Acoustic noise during functional magnetic resonance imaging. J Acoust Soc Amer 2000;108(no 4):1683–1696. [PubMed: 11051496]

24. Chapman B, Haywood B, Mansfield P. Optimized gradient pulse for use with EPI employing active noise control. Magn Reson Med 2003;50:931–935. [PubMed: 14587003]

25. Chambers J, Akeroyd M, Summerfield A, Palmer A. Active control of volume acquisition noise in functional magnetic resonance imaging: Method and psychoacoustical evaluation. J Acoust Soc Amer 2001;110(no 6):3041–3054. [PubMed: 11785805]

26. Geris, R.; Mechefske, C.; Rutt, B. Active noise reduction in a 4T whole body MR image. presented at the Int. Soc. Magn. Reson. Med., 11th Sci. Meeting Exhib.; Toronto, ON, Canada. 2003.

27. Hall D, Haggard M, Akeroyd M, Plamer A, Summerfield A, Elliott M, Gurney E, Bowtell R. Sparse temporal sampling in auditory fMRI. Hum Brain Mapp 1999;7:213–223. [PubMed: 10194620]

28. Belin P, Zatorre R, Hoge R, Evans A, Pike B. Event-related fMRI of the auditory cortex. NeuroImage 1999;10:417–429. [PubMed: 10493900]

29. Gaab N, Gabrieli J, Glover G. Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. Hum Brain Mapp 2007;28(no 8):703–720. [PubMed: 17080440]

30. Schwarzbauer C, Davis M, Rodd J, Johnsrude I. Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI. NeuroImage 2006;29:774–782. [PubMed: 16226896]

31. Zaehle T, Schmidt C, Meyer M, Baumann S, Baltes C, Boesiger P, Jancke L. Comparison of 'silent' clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. NeuroImage 2007;37:1195–1204. [PubMed: 17644001]

32. Schmidt C, Zaehle T, Meyer M, Geiser E, Boesiger P, Jancke L. Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. Hum Brain Mapp 2008;29:46–56. [PubMed: 17318832]

33. Buckner R. Event-related fMRI and the hemodynamic response. Hum Brain Mapp 1998;6:373–377. [PubMed: 9788075]

34. Josephs O, Turner R, Friston K. Event-related fMRI. Hum Brain Mapp 1997;5:243–248.

35. Dale A, Buckner R. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 1997;5:329–340.

36. Buxton R, Wong E, Frank L. Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. Magn Reson Med 1998;39:855–864. [PubMed: 9621908]

37. Malonek D, Dirnagl U, Lindauer U, Yamada K, Kanno I, Grinvald A. Vascular imprints of neuronal activity: Relationships between the dynamics of cortical blood flow, oxygenation, and volume changes following sensory stimulation. Proc Natl Acad Sci USA 1997;94:14826–14831. [PubMed: 9405698]

38. Mandeville J, Marota J, Kosofsky B, Keltner J, Weissleder R, Rosen B, Weisskoff R. Dynamic functional imaging of relative cerebral blood volume during rat forepaw stimulation. Magn Reson Med 1998;39:615–624. [PubMed: 9543424]

39. Harms M, Melcher J. Sound repetition rate in the human auditory pathway: Representations in the waveshape and amplitude of fMRI activation. J Neurophysiol 2002;88(no 3):1433–1450. [PubMed: 12205164]

40. Hall D, Summerfield A, Goncalves M, Foster J, Palmer A, Bowtell R. Time-course of the auditory bold response to scanner noise. Magn Reson Med 2000;43:601–606. [PubMed: 10748437]

41. Talavage T, Ledden P, Benson R, Rosen B, Melcher J. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. Hear Res 2000;150:225–244. [PubMed: 11077206]

42. Galaburda A, Sanides F. Cytoarchitectonic organization of the human auditory cortex. J Comp Neurol 1980;190(no 3):597–610. [PubMed: 6771305]

43. Petridou, N.; Ye, F.; McLaughlin, A.; Bandettini, P. Relationship between S/N and fMRI sensitivity. Proc. Int. Soc. Magn. Reson. Med., 9th Sci. Meeting Exhib.; Glasgow, U.K.. 2001. p. 1181

44. Tamer, G., Jr; Talavage, T.; Luh, WM.; Ulmer, J. Characterizing the amplitude and spatial extent of the cortical response in auditory cortex to acoustic scanner noise generated during echo-planar image acquisition in functional magnetic resonance imaging. Proc. 26th Annu. Int. Conf. IEEE EMBS; San Francisco, CA. Sep. 2004; p. 1899-1902.

45. Cox R. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996;29:162–173. [PubMed: 8812068]

46. Birn R, Saad Z, Bandettini P. Spatial heterogeneity of the nonlinear dynamics in the FMRI BOLD response. NeuroImage 2001;14:817–826. [PubMed: 11554800]

47. Friston K, Fletcher P, Josephs O, Holmes A, Rugg M, Turner R. Event-related fMRI: Characterizing differential responses. NeuroImage 1998;7:30–40. [PubMed: 9500830]

Biographies

Gregory G. Tamer, Jr. received the B.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1999 and 2005, respectively.

From 2005 to 2007, he was a Postdoctoral Fellow in the Psychology Department, University of Illinois at Urbana-Champaign. In 2007, he returned to Purdue University to serve as the Manager of Operations at the new Purdue Magnetic Resonance Imaging (MRI) Facility.

Wen-Ming Luh received the Ph.D. degree in biophysics from the Medical College of Wisconsin, Milwaukee, in 1999. He received postdoctoral training in radiology at the University of California at San Diego. Prior to 2001, he was an Applications Development Engineer of the Global MR Engineering of GE Medical Systems. He is currently a Staff Scientist of the Functional MRI Facility of the Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD.

Thomas M. Talavage (S'90–M'98) received the B.S.C.E.E. and M.S.E.E. degrees from Purdue University, West Lafayette, IN, in 1992 and 1993, respectively, and the Ph.D. degree in speech and hearing sciences from the Harvard-MIT Division of Health Sciences and Technology, Cambridge, in 1998.

Thereafter, he joined the Faculty of Purdue University, where he holds a joint appointment as Associate Professor in both the School of Electrical and Computer Engineering and the Weldon School of Biomedical Engineering. He is also currently a Co-Director of the Purdue Magnetic Resonance Imaging (MRI) Facility.
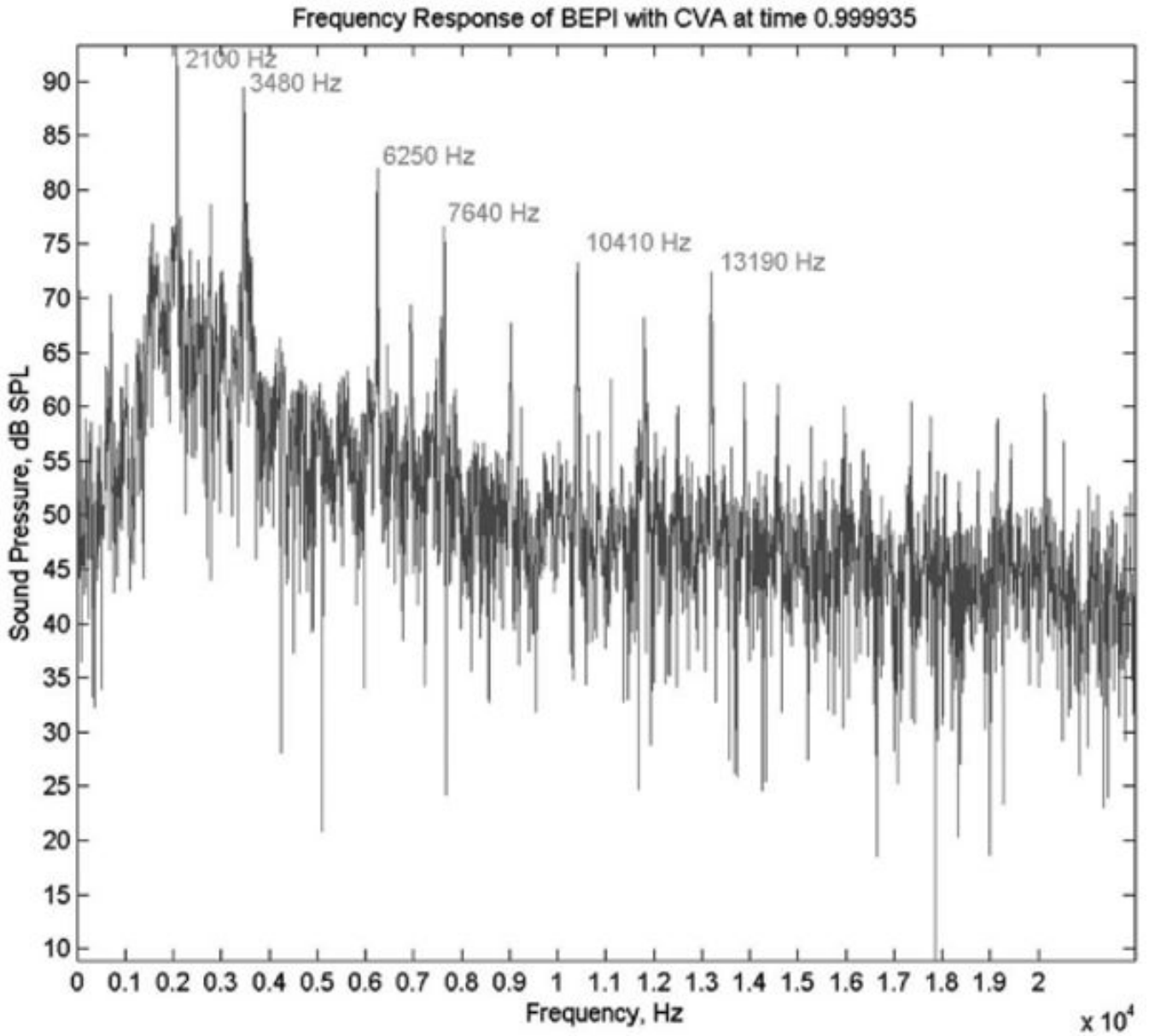
Fig. 1.
Frequency spectrum produced by the gradient readout process for the single-ping stimulus at the 1.5 T General Electric Signa LX horizon imager located at the Indiana University School of Medicine (Indianapolis, IN). The maximal acoustic intensity is 107 dB SPL, with the most prominent peaks at 2100 Hz (95 dB SPL) and 3480 Hz (90 dB SPL). The selected readouts were performed as part of a CVA sequence using a maximum slew rate of 120 mT/m/ms to acquire data from a 24 cm field of view with a 64×64 acquisition matrix. The spectrum was obtained from the approximate position of the subjects ears, when supine in the bore and recorded using Audio-Technica AT899 microphones and a National Instruments DAQCard AI-16E-4, with sampling of the signal at 44.1 kHz.
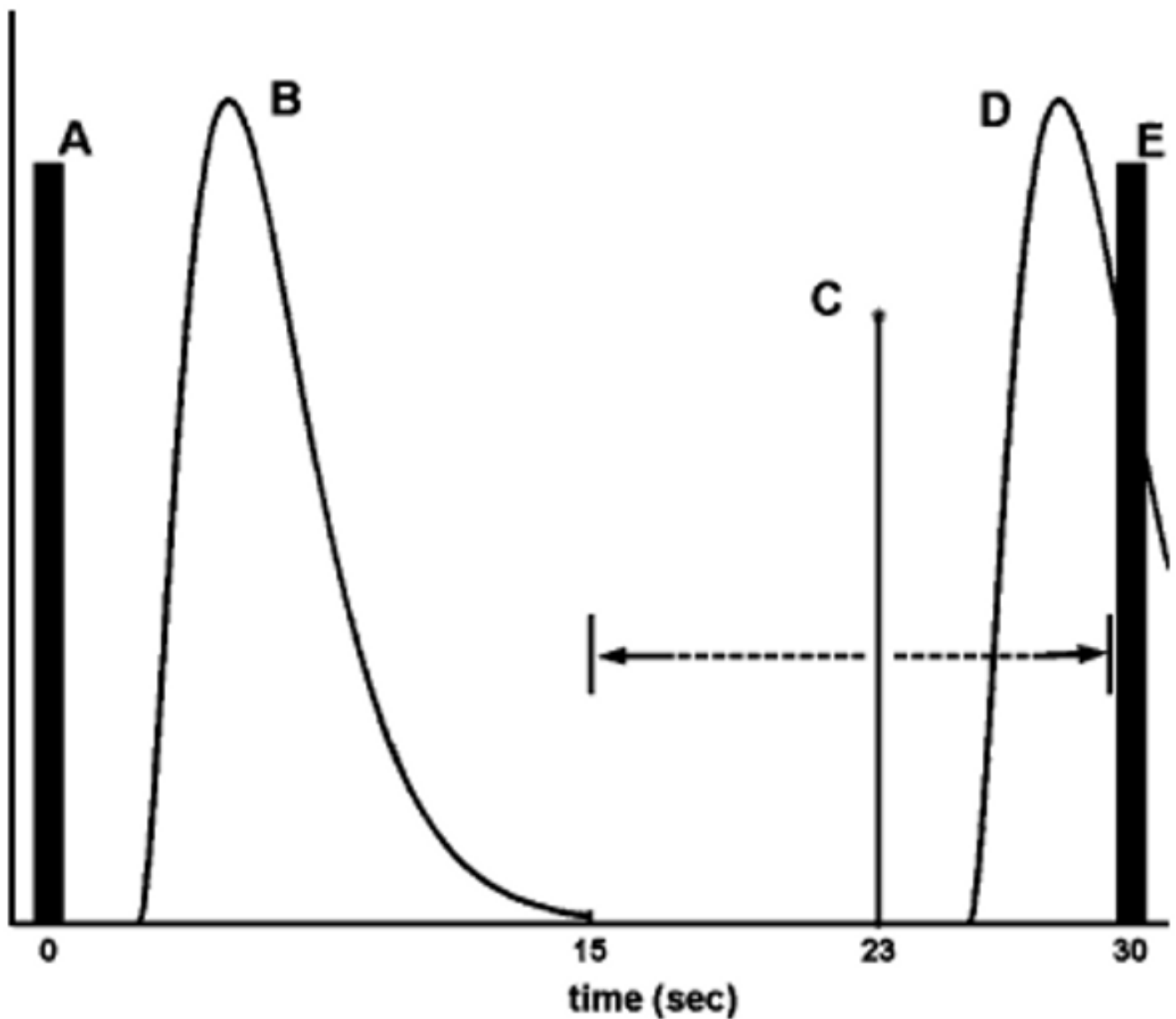
Fig. 2.
Illustration of the paradigm in which the ideal hemodynamic response (HDR) (object B) to the previous volume acquisition (A) is displayed. The single-ping stimulus (C) is presented after the HDR to the previous volume acquisition (B) has returned to baseline, roughly 15 s after the previous volume acquisition (A). The time between the presentation of the single-ping stimulus (C) and the subsequent volume acquisition (E) is jittered from trial to trial to permit the measurement of the time course of the response at 1-s resolution. In the illustration, the single-ping stimulus (C) is presented 23 s after the previous volume acquisition (A) and 7 s before the subsequent volume acquisition (E). This particular subsequent volume acquisition (E) theoretically measures the response (D) at poststimulus time $t = 7$ s. The stimulus (C) is jittered in the range (illustrated with dashed lines) of 15–30 s after the previous volume acquisition (A). Note that no ping is actually presented when a delay of 30 s (poststimulus time $t = 0$ s) is used.
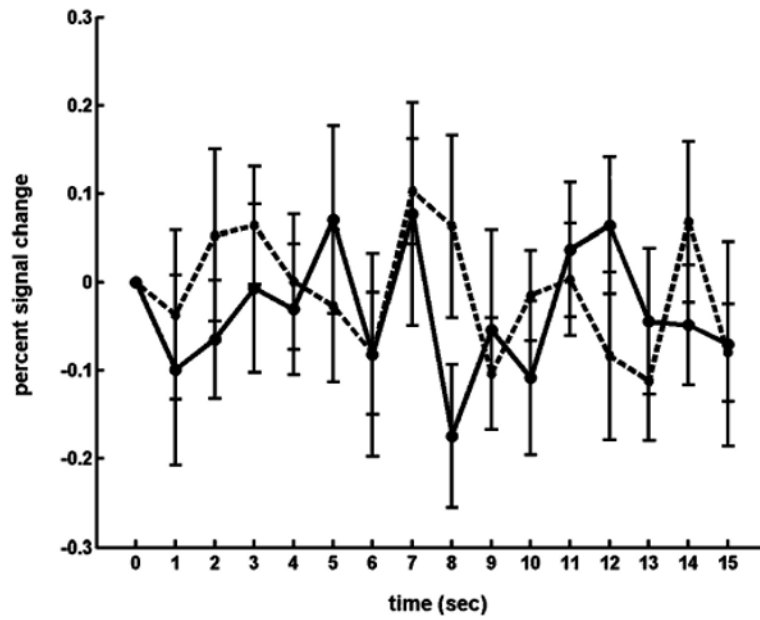
Fig. 3.
Two estimated responses (solid and dashed lines) to the single-ping stimulus for the control run using the phantom. The mean percent signal change value at each time point is presented, along with standard error bars.
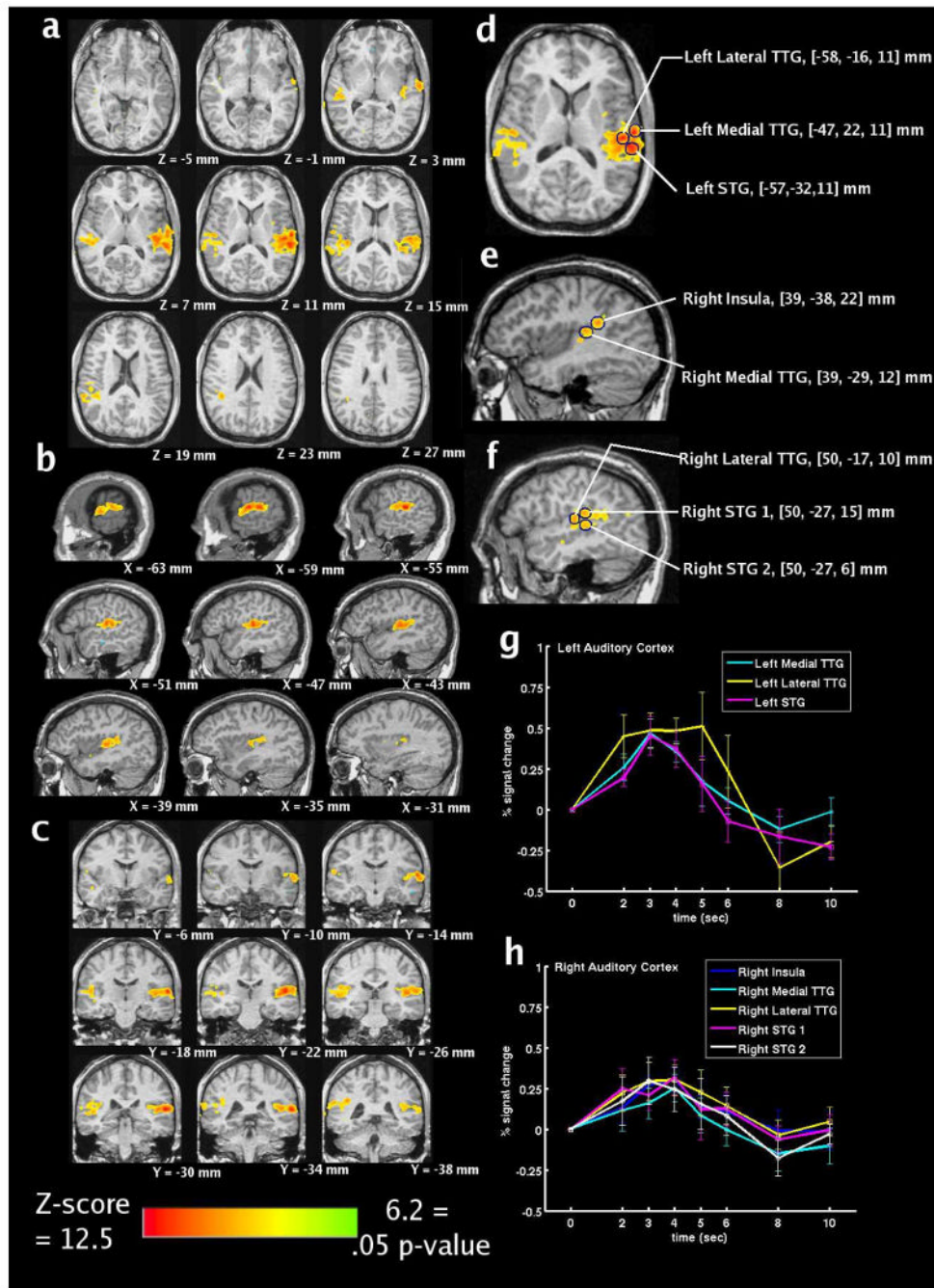
Fig. 4.
(a)–(c) Statistical maps, (d)–(f) selected regions of interest, and (g), (h) group-averaged, estimated HDRs within the selected ROIs to the single "ping" for the ten sessions.

TABLE I

Estimated Parameter Values (A_1, δ_1, τ_1, A_2, δ_2, τ_2) of Double-Gamma-Variate HDR Model from Nonlinear Regression for Each ROI, as well as Averages Across ROIs Within Each Hemisphere

ROI	A_1	δ_1	τ_1	A_2	δ_2	τ_2
Left Medial TTG	1.52	1.22	1.11	0.79	1.25	1.57
Left Lateral TTG	4.45	1.77	0.60	4.82	2.69	0.28
Left STG	1.07	1.37	0.95	0.45	0.53	3.61
Left Average	2.35	1.46	0.89	2.02	1.49	1.82
Right Insula	1.30	1.22	1.24	0.79	1.15	1.48
Right Medial TTG	1.18	1.28	1.47	0.94	1.28	2.04
Right Lateral TTG	1.44	1.09	0.85	0.96	1.41	0.61
Right STG 1	1.22	0.74	0.78	1.86	1.85	0.24
Right STG 2	1.28	1.11	1.40	0.86	1.16	1.92
Right Average	1.28	1.09	1.15	1.08	1.37	1.26