Brain Topogr (2012) 25:182–193
DOI 10.1007/s10548-011-0206-x

ORIGINAL PAPER

Reducing the Interval Between Volume Acquisitions Improves "Sparse" Scanning Protocols in Event-related Auditory fMRI

Franziskus Liem · Kai Lutz · Roger Luechinger · Lutz Jäncke · Martin Meyer

Received: 12 July 2011 / Accepted: 7 October 2011 / Published online: 21 October 2011

Abstract Sparse and clustered-sparse temporal sampling fMRI protocols have been devised to reduce the influence of auditory scanner noise in the context of auditory fMRI studies. Here, we report an improvement of the previously established clustered-sparse acquisition scheme. The standard procedure currently used by many researchers in the field is a scanning protocol that includes relatively long silent pauses between image acquisitions (and therefore, a relatively long repetition time or cluster-onset asynchrony); it is during these pauses that stimuli are presented. This approach makes it unlikely that stimulus-induced BOLD response is obscured by scanner-noise-induced BOLD response. It also allows the BOLD response to drop near baseline; thus, avoiding saturation of BOLD signal and theoretically increasing effect size. A possible drawback of this approach is the limited number of stimulus presentations and image acquisitions that are possible in a given period of time, which could result in an inaccurate estimation of effect size (higher standard error). Since this line of reasoning has not yet been empirically tested, we decided to vary the cluster-onset asynchrony (7.5, 10, 12.5, and 15 s) in the context of a clustered-sparse protocol. In this study sixteen healthy participants listened to spoken sentences. We performed whole-brain fMRI group statistics and region of interest analysis with anatomically defined regions of interest (auditory core and association areas). We discovered that the protocol, which included a short cluster-onset asynchrony (7.5 s), yielded more advantageous results than the other protocols, which involved longer cluster-onset asynchrony. The short cluster-onset asynchrony protocol exhibited a larger number of activated voxels and larger mean effect sizes with lower standard errors. Our findings suggest that, contrary to prior experience, a short cluster-onset asynchrony is advantageous because more stimuli can be delivered within any given period of time. Alternatively, a given number of stimuli can be presented in less time, and this broadens the spectrum of possible fMRI applications.

Keywords Auditory cognition · fMRI protocols · Auditory neuroimaging · Auditory cortex

Electronic supplementary material The online version of this article (doi:10.1007/s10548-011-0206-x) contains supplementary material, which is available to authorized users.

F. Liem (✉) · K. Lutz · L. Jäncke · M. Meyer
Division Neuropsychology, Institute of Psychology, University of Zurich, Binzmühlestrasse 14/25, 8050 Zurich, Switzerland
e-mail: f.liem@psychologie.uzh.ch

K. Lutz
e-mail: k.lutz@psychologie.uzh.ch

L. Jäncke
e-mail: l.jaencke@psychologie.uzh.ch

M. Meyer
e-mail: m.meyer@psychologie.uzh.ch

R. Luechinger
Institute for Biomedical Engineering, University and ETH Zurich, Zurich, Switzerland
e-mail: luechinger@biomed.ee.ethz.ch

L. Jäncke · M. Meyer
Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland

L. Jäncke
International Normal Aging and Plasticity Imaging Center, University of Zurich, Zurich, Switzerland

M. Meyer
Research Unit for Plasticity and Learning of the Healthy Aging Brain, University of Zurich, Zurich, Switzerland

Introduction

Despite the delayed temporal characteristics of the hemodynamic response, functional magnetic resonance imaging (fMRI) is a major tool in present cognitive neuroscience. As the scanner produces auditory noise, a variety of problems emerge, especially in the context of auditory fMRI studies (Amaro et al. 2002; Moelker and Pattynama 2003). Major disadvantages include: acoustical overlap between scanner noise and stimulus, and an enhancement of the apparently "silent" baseline condition in continuous protocols. The overlap between scanner noise and stimulus presentation makes it difficult to perceive auditory stimuli in their full complexity, especially if subtle stimulus manipulation is applied. Perceived spectral characteristics of stimuli are altered by ambient scanner noise that can be as loud as 130 dB. Furthermore, a research participant in the scanner may have to be more attentive, in order to perceptually separate the auditory stimulus from background scanner noise. On a physiological level, auditory scanner noise can lead to saturation of neuron populations in auditory fields because this intense perpetual noise excessively drives auditory cortex activity. Scanner noise induces a BOLD response in auditory-related cortex areas during trials without a proper auditory stimulus. Most interestingly, this appears to happen differentially for the left and the right hemisphere (Herrmann et al. 2000; Tamer et al. 2009; Schmidt et al. 2008) and in a nonlinear manner (Talavage and Edmister 2004). The degree of nonlinearity varies between left and right hemisphere (Hu et al. 2010). The additional auditory input leads to an "inflated" baseline condition, which reduces the possible range of stimulus-induced BOLD response (more detailed accounts of these and further problems have been given for example by Gaab et al. 2007a, b; Hall et al. 1999; Eden et al. 1999).

To overcome these constraints, several groups have published groundbreaking techniques, which are standard in today's auditory fMRI (Eden et al. 1999; Edmister et al. 1999; Talavage et al. 1999; Hall et al. 1999). A variety of names exists to date for similar approaches: Edmister et al. (1999) and Talavage et al. (1999) presented "clustered volume acquisition"; Hall et al. (1999) called their approach "sparse temporal sampling"; Eden et al. (1999) published the "behavior interleaved gradients technique". In the present article we use the term "sparse design" for aforementioned approaches and "clustered-sparse design" for an extension of this design (Schmidt et al. 2008; Zaehle et al. 2007). "Clustered volume acquisition" refers to the temporal clustering of several slices within one volume and is not to be confused with the clustered-sparse protocol, which refers to the clustering of volumes within one trial.

The "sparse" temporal acquisition scheme reduces the scanner's acoustical noise influence by acquiring only one functional image per trial. When repetition time (TR) is long (around 10–14 s), the preceding trial's scanner-evoked BOLD signal returns close to baseline. In this scheme, auditory stimuli are delivered during the silent pause between two image acquisitions. This allows for unobstructed stimulus perception. In addition to these benefits, sparse acquisition schemes show a higher SNR because T1 magnetization can fully recover prior to each image acquisition. This is impossible in continuous protocols. Notwithstanding these undisputable advantages this approach has some drawbacks. The total duration of an experiment is increased and image acquisition needs to be timed around the peak of the BOLD response. Timing is not a pressing issue in block designs where the stimulus-evoked BOLD response almost reaches a steady state at the plateau (e.g. when brief tones or syllables are repeatedly presented for approximately ten seconds followed by an acquisition of one functional volume). However, whenever single-stimulus presentation in the context of an event-related design is desired, timing is a key issue (cf. Fig. 1). For example, timing is a key issue, when investigating the perception of slow modulations in auditory stimuli with durations in the range of a few seconds, such as, prosody in sentences, or melody in brief excerpts of music.

To ensure that the BOLD response's peak is sampled, the clustered-sparse temporal acquisition (CTA) protocol has been devised (Zaehle et al. 2007; Schmidt et al. 2008). Derived from the sparse protocol, the clustered-sparse scheme allows for the collection of a cluster of (usually three) consecutive functional volumes per trial. As a result, the likelihood of recording the hemodynamic response's peak increases, as does the number of acquired images and therefore the number of observations. This makes the clustered-sparse protocol superior to the sparse approach with respect to statistical power. Especially in single subject analyses, this superiority becomes manifest in more precisely estimated effects, namely, beta-values (Zaehle et al. 2007). The duration of the silent pause in CTA designs is determined by the cluster-onset asynchrony (COA, the time between two consecutive cluster-onsets).

Interestingly, there does exist an approach similar to the clustered-sparse scheme: silent gradient protocols (Mueller et al. 2011; Schwarzbauer et al. 2006; Schmithorst and Holland 2004) also sample data by using a cluster of several volumes. Clusters are separated by silent intervals during stimulus presentation. In contrast to our approach, longitudinal magnetization is held constant via silent slice-selective excitation pulses. As a result, the T1-decay-related signal-to-noise ratio (SNR) improvement (for at least

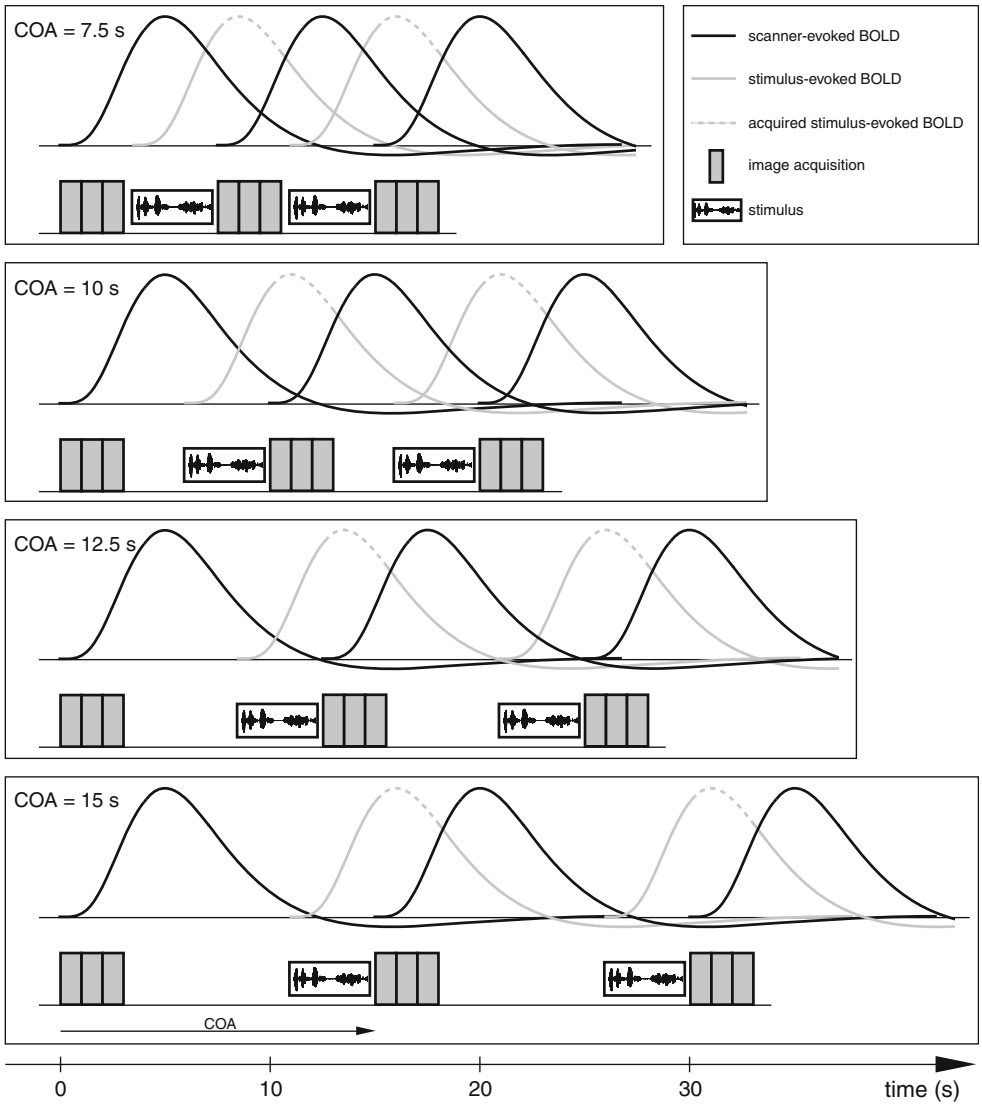

Fig. 1 Assumed BOLD response evoked by auditory stimuli and auditory scanner noise, respectively

the first image per cluster), which the clustered-sparse protocol benefits from, is absent.

Especially in studies focusing on auditory-related areas silent protocols, in comparison to conventional continuous acquisition protocols, have been shown to be beneficial in terms of SNR and effective power (Schmidt et al. 2008; Gaab et al. 2007a, b; Hall et al. 1999). However, a slow timing and the resulting inflated duration of the scanning session can make such studies a tedious and gruelling experience for participants. Therefore, the aim of this study is to empirically improve the timing setup by varying the COA. Compared with long COAs (e.g. 15 s), short COAs (e.g. 7.5 s) lead to a notable increase in the amount of acquired data in a given period of time. This should increase the accuracy of the parameter estimates, as manifested in the parameter estimate's lower standard error of the mean (SEM).

The SEM is of major interest because it directly influences t values and, therefore, the statistical significance. Generally spoken, if two designs show identical beta-values but different standard errors the t value is higher in the design with the smaller standard error (Mechelli et al. 2003). We applied an approach introduced by Mechelli et al. (2003) to segregate the SEM ($(\sigma^2 c^T(X^TX)^{-1}c)^{1/2}$) into error variance ($\sigma^2$) and design variance ($c^T(X^TX)^{-1}c$). The design variance is solely depending on the design matrix and the contrast and represents a measurement of the variance of the explanatory variables and their correlations. The error variance is variance in the data that cannot be explained by the model. While short COAs allow for the presentation of a large number of trials, long COAs permit a clear separation of stimulus-evoked and scanner-noise-evoked BOLD response, as well as a return of BOLD response to baseline level. In addition, saturation effects

are not expected to occur and, thus, do not influence the signal. This should result in a wider dynamic range of the BOLD signal (and a larger potential influence of auditory stimuli on the BOLD signal) and, therefore, in higher effect sizes (beta-values). Whereas the design variance is expected to be favourably influenced (reduced) by the higher number of images in a short COA setting, this effect should vanish if the amount of analyzed data is kept equal across the conditions by analyzing only part of the samples in short COA settings. In the present study, this was done by analyzing only the first 30 trials of each condition. In contrast, effect sizes should be fairly unaffected by such an analysis. Error variance is also not expected to be influenced by the number of trials. However, neurophysiological processes, for instance saturation effects or changes in the shape or temporal characteristics of the hemodynamic response, might render the model less appropriate, and therefore increase error variance. Whether decreases in design variance translate to decreases in SEM depends on error variance.

The current investigation focuses on event-related clustered-sparse designs, which shall exam hemodynamic response to processing of stimuli spanning over several seconds, namely spoken sentences and music stimuli, because these designs can elucidate neuronal mechanisms supporting slow acoustic modulation.

For this purpose, four differential COA settings were implemented (cf. Fig. 1): 7.5, 10, 12.5 and 15 s [the latter was tested in prior clustered-sparse fMRI studies (Zaehle et al. 2007; Schmidt et al. 2008)]. To evaluate to what extent and how differential COA settings may influence the ability to identify different brain responses to stimuli that varied in loudness, we presented participants with sentences of two different intensities. Prior studies have convincingly shown an increase in the number of significantly activated voxels or percent signal change or both in auditory-related cortex areas as a result of increasing stimulus intensity (e.g. Jäncke et al. 1998; Hart et al. 2003; Mulert et al. 2005; Brechmann et al. 2002). In the present study, the variations in intensity were implemented merely as a vehicle to show differences in the settings' sensitivities.

It is assumed that the preceding trial's scanner-evoked BOLD response has less influence on the current image at long COAs because the signal has time to return to, or near baseline (cf. Fig. 1; Hall et al. 2000). Therefore, we expect a systematic increase in effect size (beta-values) at longer COA settings. Notably, a short COA setting provides more data within a given period of time, which reduces design variance. If all COAs yield equal error variance, SEM will decrease as a function of decreasing design variance, which leads to more precisely estimated effect sizes. Our aim is to find a COA setting that balances adverse impact on effect sizes and their estimation accuracy.

Materials and Methods

Participants

Sixteen subjects (eight female) took part in this experiment. Participants were between 20 and 27 years old ($M = 23$, $SD = 2$). They were screened for hearing impairments, tinnitus, dyslexia, neurological and neuropsychological history. Participants were also asked if they had any metal implants or devices in their bodies. Subjects had normal or corrected vision. All subjects were German or Swiss German native speakers and right-handed according to the Annett questionnaire (Annett 1992). They gave written informed consent and were paid for their participation. This study was approved by Canton Zurich's Ethics Committee (application E-40/2009).

Stimuli, Experimental Conditions and Task

In this study we presented spoken German sentences to participants whilst they were placed in an MR scanner. The sentences were spoken by a trained female speaker and were recorded in a soundproof chamber at the University of Zurich Phonetics Lab. Each of the 146 sentences lasted on average about 3.0 s ($SD = 0.4$ s). Each sentence was presented only once during the course of the experiment (example below).

Sentences were delivered across four randomized runs (=four different conditions), each of which varied in COA (7.5, 10, 12.5 and 15 s). The COA is the single within-subjects factor of interest in this experiment. One run lasted 7.5 min and was composed of the maximum number of trials possible, depending on the COA (cf. Table 1).

Two other within-subjects factors of no direct interest were varied in this experiment: sentence intensity and sentence accent. In order to measure the COA setting's influence on the ability to detect differences between two classes of stimuli, sentences were presented pseudorandomly with either of two sound pressure levels (SPLs). By using the Praat software (v5.1.09; http://www.fon.hum.uva.nl/praat/) we set the sentences' mean intensity to 60 and 80 dB SPL, respectively. The stimuli were recorded with accents that were placed either on the first, or the second part of the sentence (for example: "Laura empfiehlt Martin, den Computer zu kaufen."/"Laura advises Martin, to buy the computer." The underline indicates the possible position of sentence accent). To control for participants' attention, they were asked to indicate by button press whether the sentence they just heard had a sentence accent on the first or the second noun phrase. Manipulation of intensity and emphasis was conducted in an orthogonal, randomized manner (within one run: 50% of sentences were presented at 60 dB, 50% at 80 dB; of each intensity

Table 1 Number of trials for all experimental conditions

COA (s)	Number of trials per condition			
	60 dB	80 dB	Empty	Total
7.5	26	26	8	60
10	19	19	7	45
12.5	15	15	6	36
15	13	13	4	30

group 50% with emphasis on the first part, 50% with emphasis on the second part). As a baseline measurement, empty trials were also included at a lower rate (cf. Table 1).

Data Acquisition

An event-related clustered-sparse fMRI design was employed in this study (Schmidt et al. 2008; Zaehle et al. 2007). Via MR-compatible headphones with an incorporated piezoelectric auditory stimulation system, one auditory stimulus per trial was binaurally presented in an interval devoid of auditory scanner noise. Throughout the experiment, sentence onset was four seconds prior to acquisition onset (cf. Fig. 2). A fixation cross preceded the presentation of each sentence. The fixation cross was projected onto a screen and could be seen through a mirror mounted on the head coil. Subsequently, three functional volumes were recorded, each with an acquisition time of 1000 ms. During this interval, subjects indicated the noun phrase on which the accent was present via a button press with either their right index finger, or with their right middle finger (index finger for accents on the first noun phrase, middle finger for second noun phrase). During empty trials participants were asked to randomly press a button once, which enabled us to control for motor activity. The headphones' volume was calibrated with an SPL-meter prior to each session.

Data was collected on a Philips 3T Achieva whole-body MR unit (Philips HealthCare, Best, The Netherlands) that is equipped with an eight-channel Philips head coil. Functional time series were collected from 16 transverse slices covering the entire perisylvian cortex with a spatial resolution of 2.75 × 2.75 × 2.75 mm^3, using a single-shot, gradient-echo planar sequence (EPI acquisition matrix 80 × 80 voxels, field of view (FOV) = 220 mm, echo time (TE) = 35 ms, flip angle (FA) = 68°, SENSE factor = 2). Volume acquisition time of each EPI scan was 1000 ms. Cluster-onset asynchrony was systematically varied across the four runs (7.5, 10, 12.5 and 15 s, see Stimuli, experimental conditions and task section) but was kept constant within each run. Additionally, one whole-brain EPI (60 slices) was recorded prior to the experiment, so as to improve the spatial normalisation process according to an established procedure.

Furthermore, a standard 3D T1-weighted scan with 1 × 1 × 1 mm^3 spatial resolution (160 sagittal slices, FOV = 240 mm, TE = 3.7 ms, TR = 8.1 ms, FA = 8°) was collected, in order to obtain individual anatomical regions of interests (ROIs).

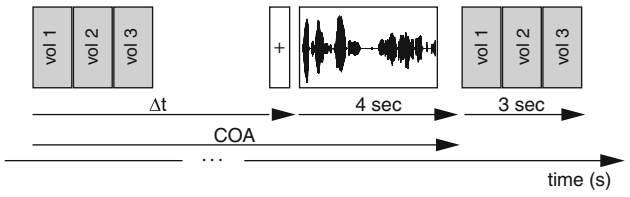

Fig. 2 Sequence of one trial. *Gray squares*: image acquisition. The interval between acquisition onset and the following stimulus onset varies with COA: Δt = COA – 4 s

Data Analysis

Behavioural data analysis and ROI statistics were performed using PASW Statistics 18.0 (SPSS Inc.).

Whole-brain Analysis[1]

Since a whole-brain group analysis is standard in fMRI experiments, we included this analysis; however, the origin of a significant effect is more systematically to evaluate in a post-hoc ROI analyses of effect sizes and its SEM. Analysis of fMRI data was carried out using SPM8 (Wellcome Department of Cognitive Neurology, London). To account for movement artefacts, the functional images were realigned to the first volume. Each run was entered as a separate session. Since the functional brain volumes where comprised of only 16 slices, the realigned images were co-registered with the whole brain EPI. This permitted an overall improved normalisation. The whole-brain EPI was normalised onto the SPM8 EPI template. Resulting spatial normalisation parameters were applied to all functional volumes. This transformed them into MNI space. Finally, the images were smoothed with an FWHM kernel of 5 × 5 × 5 mm^3 (Buchsbaum et al. 2005).

After pre-processing the data, a General Linear Model (GLM; subject level) was separately set up for each run (=COA setting). Sentence-events were entered as two separate conditions (60 and 80 dB). Due to the low number of sampling points, a boxcar function (first order, window

[1] Note that the functional brain scans covered only about 50 percent of the brain in the inferior-posterior direction, namely the entire perisylvian cortex. For the sake of simplicity, we still refer to this processing step as whole-brain analysis.

length = 3 s) was modelled for each trial. In accordance with the approach established by Zaehle et al. (2007), two regressors of no interest were included to account for the T1-decay along the three consecutive volumes. Three contrasts were calculated: an auditory default contrast (all auditory events vs. empty trials) and two direct comparisons: 80 vs. 60 dB and 60 vs. 80 dB. For each COA, individual contrast images were subjected to a random-effects second level analysis (one-sample *t* test against zero for all three abovementioned first level contrasts). Family wise error (FWE) correction was applied to the resulting statistical parametric maps. For each COA, suprathreshold voxels at the 80 vs. 60 dB contrast in the temporal lobe were counted and averaged across the two hemispheres for better statistical power.

ROI Analysis

To elaborate on effect sizes (mean beta-values) and distinct anatomical regions comprising auditory core and adjacent auditory-related cortex, a post-hoc ROI analysis was performed. Two different approaches were taken to define the ROIs. We used both (a) automatically processed, anatomically defined individual ROIs of Heschl's gyrus (HG) and planum temporale (PT) and (b) the well established cytoarchitectonically defined region TE1.0, which is included in the SPM Anatomy toolbox (v1.7; Eickhoff et al. 2005). The TE1.0 region corresponds to the normal location of the core auditory cortex on the medial portion of Heschl's gyrus (Morosan et al. 2001; Rademacher et al. 2001). We applied an individual approach, as well as a normalized ROI approach, in order to help generalize our results to different methodologies.

To obtain individual ROIs, the T1 scan was co-registered onto the whole-brain EPI. Subsequently, the whole-brain EPI's normalisation parameters were applied to the T1 scan. This procedure moved it into standard stereotactic space. The normalised anatomical brain scan was then automatically processed with the FreeSurfer software package (v4.5.0; http://surfer.nmr.mgh.harvard.edu; Dale et al. 1999; Fischl et al. 1999). After completing the default FreeSurfer processing stream, ROI masks, as provided by FreeSurfer's aparc.a2009 s parcellation, were exported individually for each subject's HG and PT. HG includes only the most anterior transverse temporal gyrus. Possible additional transverse temporal gyri are attributed to PT, which comprises both horizontal and vertical aspects (planum parietale; Destrieux et al. 2010). The ROIs only included areas that were fully covered by the functional volumes. Structural overlap maps of the individual ROIs for the entire sample are provided in S1.

Mean beta-values were extracted from first level's contrast images via an in-house tool. This was done for each ROI, COA, and hemisphere for the contrast 80 dB vs. 60 dB. Therefore, reported beta-values represent an increase in effect size from 60 to 80 dB. Since we had no interest to explore functional lateralisation, the beta-values for each ROI and COA were averaged across the two hemispheres, so as to improve statistical power. Per ROI, the resulting values were entered into a separate one-way repeated measures ANOVA with COA (7.5, 10, 12.5 and 15 s) as the within-subjects factor for each ROI. Subsequently, linear and quadratic trend analysis was performed on the significant effects.

Standard Error, Design Variance and Error Variance

To obtain information about the first level GLM's error variance ($\sigma 2$), mean values within TE1.0 were collected from the ResMS.img of each subject's first level model (Zaehle et al. 2007). The design variance was calculated from the design matrix and the contrast vector ($c^T(X^TX)^{-1}c$). The standard error was calculated from these measures ($(\sigma^2 c^T(X^TX)^{-1}c)^{1/2}$; see Mechelli et al. 2003, Eq. 5). As with the beta-values, data was averaged across the hemispheres and entered into one-way repeated measures ANOVAs with COA as the within-subjects factor.

To control for the influence of different number of trials per run, all whole-brain and subsequent ROI analyses were performed twice: first entering all collected functional volumes into the model, second entering only the first 30 trials into the model. The first 30 trials of each run were arranged to comprise an equal number of sentences and empty trials (13 60-dB-sentences, 13 80-dB-sentences and 4 empty trials). As a result of this processing step, eight SPM t-tests (80 vs. 60 dB: 2 × 4 COA conditions), six ANOVAs at the ROI effect size analyses (2 × HG, PT and TE1.0), two ANOVAs at the analysis of SEM of parameter estimates, and two ANOVAs at the analysis of error variance will be reported. Since the run with COA = 15 s only contained 30 trials, the all-trials-analysis and the 30-trials-analysis at COA = 15 s are by definition identical.

Results

Behavioural Data

Overall, the evaluation of sentence accent made by research participants was accurate [percent correct (SD): COA = 7.5: 97.4% (2.5), COA = 10: 98.5% (1.9), COA = 12.5: 97.9% (2.9), COA = 15: 98.8% (2.3)]. It did

not differ between the different COA settings (ANOVA for repeated measures: F (3, 45) = 1.15, ns).

Whole-brain Analysis

For whole-brain second level group analysis, three t contrasts were calculated for each COA setting separately: an auditory default contrast (all auditory events vs. empty trials) and the two comparisons 80 vs. 60 dB and 60 vs. 80 dB.

To ensure the general integrity of our analysis, the auditory default contrast was calculated. This revealed bihemispheric clusters for all COA settings in the superior temporal lobe (cf. S2).

Since the comparison of the 80 versus 60 dB contrast over different COA settings indicates differences in the settings' sensitivity in detecting intensity variations in stimuli, this contrast was of major interest. This analysis yielded bihemispheric neuronal activation in the superior temporal plane for each COA setting (FWE, $P < .05$). Figure 3 shows the contrast 80 versus 60 dB for all COA settings. The highest amount of significant voxels can be observed at COA = 7.5 s (cf. Table 2; Fig. 4).

To control for the amount of acquired data, additional models were calculated in a second step of analysis. The number of functional brain volumes per COA setting was reduced to the first 30 trials and, as a result, was equal for each COA setting. Once again, the COA = 7.5 s setting revealed the largest clusters. Overall, a lower number of significant voxels can be observed when comparing this analysis to the analysis that comprised all scans.

The reversed t-contrast (60 vs. 80 dB) did not result in suprathreshold voxels at any of the COA settings.

ROI Analysis

Analyses for the three ROIs were also conducted. Mean beta-values for the 80 versus 60 dB contrast were collected from the individual automatic delineations of HG and PT (created by the FreeSurfer software), as well as the TE1.0 region. Each ROI's mean differential beta-value (80 vs. 60 dB) was subjected to a repeated-measures ANOVA, with COA setting as the within-subjects factor. In each ROI a significant main effect of COA could be found, even when the number of recorded trials was equally balanced (cf. Table 3; Fig. 5; absolute values are depicted in S3). All linear trends reached statistical significance; in contrast, none of the quadratic trends achieved statistical significance.

In general, the COA = 7.5 s setting resulted in significantly enhanced differential beta-values compared to that of the longer COAs. Therefore, the 7.5 s COA setting results in the largest sensitivity in detecting stimuli differences. This holds true for all three ROIs.

Standard Error, Design Variance and Error Variance

To elaborate on the models' SEM, design variance and error variance, the design variance was calculated for the contrast 80 versus 60 dB and values for mean error variance were extracted from TE1.0 (auditory core region). This was done only for this ROI, in order to rule out effects of interindividual variability. Then, the standard error was calculated from these measures (cf. Table 4). Subsequently, standard error and error variance were subjected to ANOVAs. Values for design variance increased with decreasing number of trials; therefore, increased with increasing COA. The mean error variance increased with increasing COA, though not significantly (when all trials were analysed: F (1, 19) = 1.4, ns., Greenhouse–Geisser correction; when an equal number of trials were analyzed: F (1, 20) = 2.0, ns., Greenhouse–Geisser correction). SEM of parameter estimates did significantly increase with increasing COA at the all-trials-analysis (F (1, 20) = 29.2, $P < .001$, Greenhouse–Geisser correction; post-hoc linear trend: F (1, 15) = 36.2, $P < .001$). When equalising the number of trials, the COA settings do not differ significantly with regards to SEM of parameter estimates (F (1, 22) = 2.3, ns., Greenhouse–Geisser correction; See Table 4).

Discussion

"Silent" scanner protocols are an important tool in auditory fMRI research. Only they allow the presentation of auditory stimuli without disturbance from auditory scanner noise. Standard procedure for sparse and clustered-sparse acquisition schemes is to introduce relatively long silent pauses for stimulus presentation between image acquisitions. It has been previously demonstrated that this approach renders the stimulus-evoked BOLD response largely unaffected by scanner-noise-evoked BOLD response. Therefore, this approach produces data with larger effects (Schmidt et al. 2008). On the downside, relatively long pauses between trials limit the amount of recorded images. This weakens the accuracy of estimated parameters. While the aforementioned reasoning was derived from knowledge about the characteristics of the BOLD response, the aim of this article is to empirically determine the optimal cluster-onset asynchrony setting for clustered-sparse acquisition schemes in the context of an auditory event-related fMRI design; therefore, determining the optimal duration of the silent pause between image acquisitions. We presented participants with sentences of

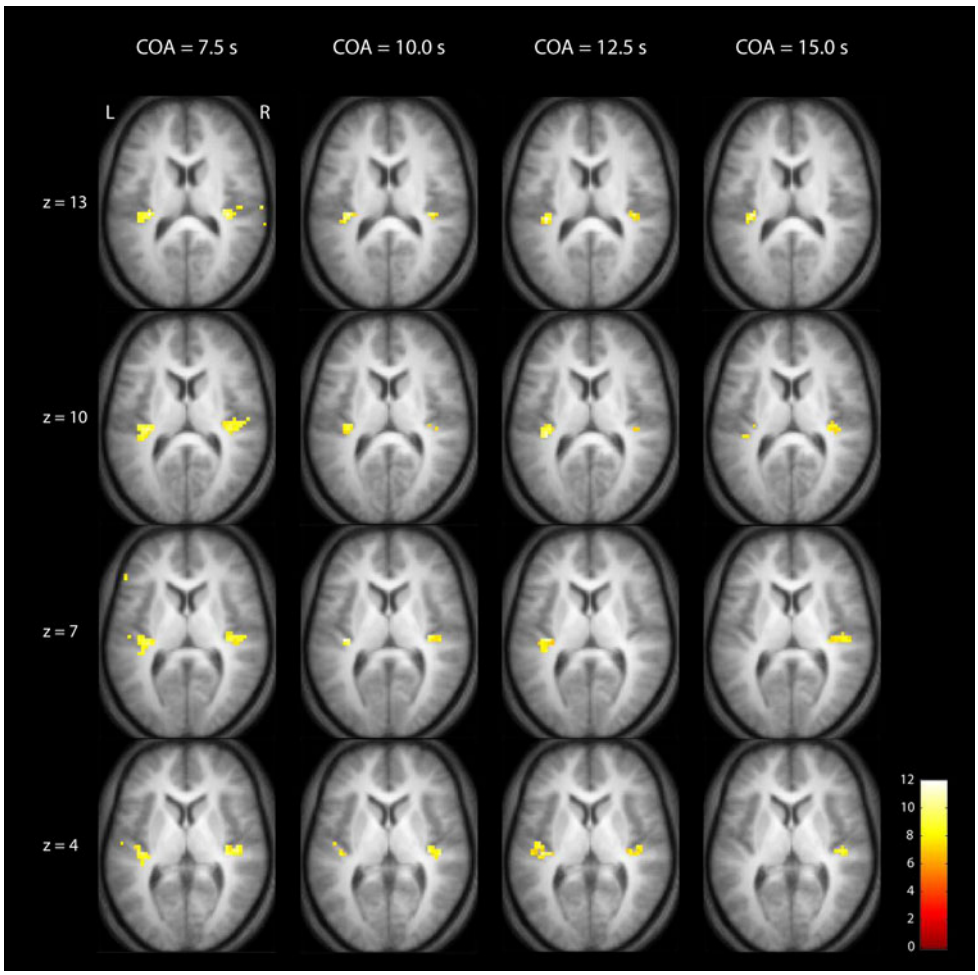

Fig. 3 Horizontal slices of the t-contrast 80 versus 60 dB at all four COA settings (all trials) for all subjects ($N = 16$) projected onto all subjects' mean T1 image (FWE, $P < .05$, $T > 6.8$). MNI space. Neurological convention

Table 2 Second level *t* test 80 dB versus 60 dB (FWE, $P < .05$)

COA (s)	All trials		30 trials	
	T(15)	# Voxels	T(15)	# Voxels
7.5	10.8	93.5	10.8	77.5
10	10.6	34	8.6	11
12.5	11.9	45	11.6	39.5
15	10.9	30.5	10.9	30.5

t value of the peak voxel and number of significant voxels in the superior temporal planes. Mean values pooled for both the left and right hemisphere

different intensities in a variety of COA settings, which were implemented in a clustered-sparse acquisition scheme. It has been demonstrated that the clustered-sparse protocol is most advantageous when working with auditory stimuli that consist of a few seconds, presented in an event-related fashion (Zaehle et al. 2007).

In accordance with present knowledge about the characteristics of the BOLD response, we hypothesised that short COA settings would lead to more accurately estimated effects (lower SEM of parameter estimates); this is because short COA settings allow to acquire a higher number of data points. On the other hand, long COA settings, which are presently the standard procedure, should show higher effect sizes (beta-values). Our data confirms the former, but consistently shows that the latter is not the case.

A ROI analysis was performed within auditory core and association areas. Error variance increases with increasing COA, though not significantly. This indicates that at COA = 7.5 s the model is at least as appropriate as at COA = 15 s. Contrary to our expectations, saturation effects might not be a problem at a short COA. Additionally, differences in the subjects' attentional state between different COA settings might influence the characteristics of the BOLD response, which alters error variance. This might vary over different brain regions (Jäncke et al. 1999; Mechelli et al. 2003; Woods et al. 2009). As both, error variance and the design variance, increase with increasing COA, a short COA setting yielded, as expected, lower SEM of parameter estimates. Nonetheless, this difference vanishes when equalising the amount of acquired data across the COA settings, as predicted by theory and by our

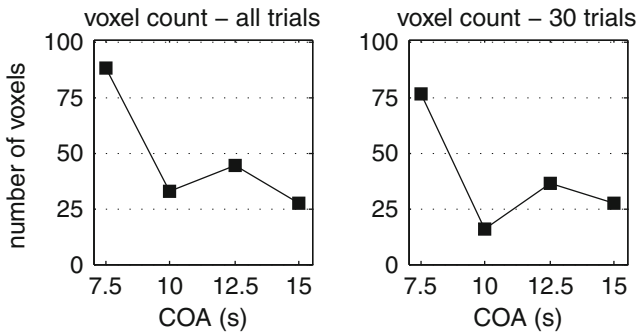

Fig. 4 Number of significant voxels in the temporal lobe produced by the second level t-contrast 80 versus 60 dB (FWE, $P < .05$; mean values pooled for both the left and right hemisphere). *Left column*: all trials. *Right column*: equal number of trials (30) for each COA setting

hypothesis. Contrary to our expectations, a short COA setting, relative to a long COA setting, led to higher effect sizes (beta-values).

Since the effect sizes at short COAs are larger and more accurately estimated (lower SEM of parameter estimates) than at long COAs, it came as no surprise that a short COA setting yielded a higher number of significant voxels at the whole-brain group statistical analysis. It should be noted that a disproportional large drop in cluster and effect size can be observed at COA = 10 s (cf. Figs. 4, 5). However, the fact that the linear trend analyses gained significance while the quadratic did not suggests a linear decline in beta-values and cluster size. Nevertheless, to draw sound conclusions about the linearity of the decline the number of COA increments needed to be larger. To directly asses the advantages of a short COA in a whole-brain analysis, we calculated the comparison COA = 7.5 s versus COA = 15 s (for the 80 vs. 60 dB contrast), which yielded bihemispherical suprathreshold voxels in the temporal lobes (cf. Fig. 6; Note that since the rather conservative voxelwise FWE correction did not yield significant results, we adopted a slightly more liberal approach. We applied a clusterwise FWE $P < .001$ with clusters selected on a voxelwise $P < .001$). Taken together, these findings demonstrate that a short COA setting (7.5 s) results in higher, more accurately estimated effects. Although not explicitly investigated here, we expect that these results should also be applicable to sparse imaging protocols.

While, error variance and SEM do not differ significantly between the COAs if the number of volume acquisitions is equal, the effect sizes do. This indicates a general advantage of a short COA, irrespective of the amount of acquired data. We conclude that in the context of clustered-sparse designs the hemodynamic response at short COAs (7.5 s in comparison to 15 s) is more dynamically susceptible to external auditory stimuli, at least as far as the auditory-related cortex is concerned. The timing of a short COA setting (7.5 s) results in each sentence being presented shortly after the peak of a previous BOLD response (cf. Fig. 1). One might reason that this occurs during a phase in which a BOLD response can be more easily elicited, than when each sentence was presented later during a phase of possible BOLD signal undershoot.

Investigations of the optimal timing (TR) in auditory fMRI designs have been previously conducted by Shah et al. (2000) and Edmister et al. (1999) in the format of fMRI on/off block designs. Nevertheless, the different stimulus-induced BOLD characteristics in block designs, as compared to event-related clustered-sparse designs, make a comparison between these studies and our approach difficult. Furthermore, as these studies only implemented TRs up to 9 s, no conclusion can be drawn about longer silent pauses.

Compared with continuous fMRI acquisition schemes, sparse and clustered-sparse designs sample less data in a given period of time. In spite of a less noisy, unobstructed signal this results in lower statistical power. Thus, improving existing protocols is of utmost importance. By all means, silent protocols are exceptionally beneficial in auditory experiments as they permit the investigation of stimulus perception, which remains unaffected by auditory scanner noise (Schmidt et al. 2008). While it is always of vital interest to reduce sources of error in any measurement tool and to obtain precise data, it becomes even more important when relating two measurements to one another, e.g. behavioural with brain imaging data.

Sentences with a duration of 3 s were used in the present study. Whether or not the present findings also apply to other stimulus durations (e.g. auditory stimuli in the range of hundreds of milliseconds or long speech or music stimuli) might be subject to future studies. Furthermore, the

Table 3 ROI analysis

Main effect COA and linear trend analysis of six separate one-way repeated measures ANOVA

ROI	All trials ANOVA	Linear trend	30 trials ANOVA	Linear trend
TE1.0	F (3, 45) = 8.1	F (1, 15) = 21.2	F (3, 45) = 8.5	F (1, 15) = 17.0
	$P < .001$	$P < .001$	$P < .001$	$P < .001$
HG	F (3, 45) = 8.0	F (1, 15) = 16.4	F (3, 45) = 7.9	F (1, 15) = 11.2
	$P < .001$	$P < .001$	$P < .001$	$P < .01$
PT	F (3, 45) = 3.7	F (1, 15) = 9.0	F (3, 45) = 3.1	F (1, 15) = 6.7
	$P < .05$	$P < .01$	$P < .05$	$P < .05$

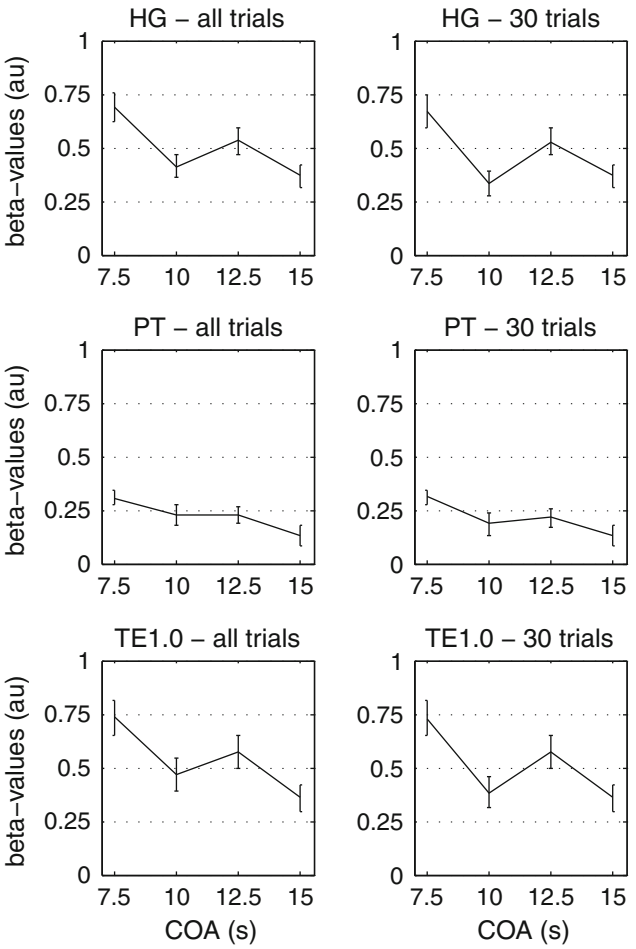

Fig. 5 Mean differential beta-values (80 versus 60 dB) for each of the four COAs in the anatomically defined ROIs (±1 SEM). *Left column*: all trials. *Right column*: equal number of trials (30) for each COA setting. *HG* Heschl's gyrus, *PT* planum temporale, *TE1.0* auditory core region. HG and PT ROIs were constructed automatically and individually for each participant ($N = 16$). All effects are significant (cf. Table 3)

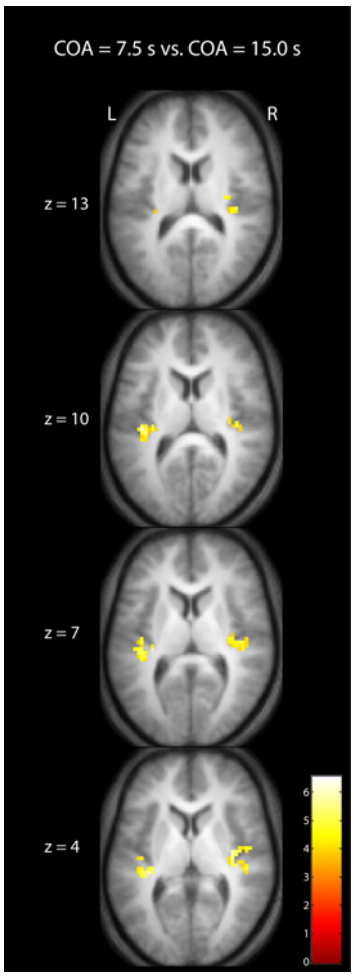

Fig. 6 Horizontal slices of the t contrast COA = 7.5 s versus COA = 15 s (for the comparison 80 vs. 60 dB; all trials) for all subjects ($N = 16$) projected onto all subjects' mean T1 image (clusterwise FWE $P < .001$, clusters selected on a voxelwise $P < .001$). MNI space. Neurological convention

Table 4 Standard error, design variance and error variance for TE1.0

COA	Standard error $(\sigma^2 c^T(X^TX)^{-1}c)^{1/2}$	Design variance $c^T(X^TX)^{-1}c$	Error variance σ^2
All trials			
7.5	0.201	0.026	1.638
10	0.245	0.035	1.768
12.5	0.283	0.044	1.920
15	0.310	0.051	2.000
30 trials			
7.5	0.277	0.051	1.542
10	0.294	0.051	1.726
12.5	0.303	0.051	1.911
15	0.309	0.051	1.999

potential influence of attention-related effects might be a topic for further research. For instance, altering the number of stimuli or conditions or changing the task difficulty may alter the participants' attentional state and, as a result, the error variance.

Taken together, these results consistently prove an overall advantage for a short cluster-onset asynchrony setting (7.5 s, i.e. … 3 s of image acquisition followed by 4.5 s of scanner-silence and stimulus presentation followed by 3 s of image acquisition…) implemented in clustered-sparse fMRI designs. In this study, a short COA resulted in higher differential beta-values and number of activated voxels, irrespective of the number of recorded trials. This opens up the future possibility of increasing the number of trials presented to research participants in a given period of time, therefore increasing the statistical power. Alternatively, an equal number of trials can be delivered in a

shorter period. This is especially convenient when working with patients or children because short experiment durations may help make this experience more pleasant and in some cases even possible.

Acknowledgments This research was supported by the Swiss National Foundation (Grant No. 320030-120661 to MM) and by the "Fonds zur Förderung des akademischen Nachwuchses (FAN) des Zürcher Universitätsvereins (ZUNIV)". We are indebted to Sarah McCourt-Meyer and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

References

Amaro E, Williams SCR, Shergill SS, Fu CHY, MacSweeney M, Picchioni MM, Brammer MJ, McGuire PK (2002) Acoustic noise and functional magnetic resonance imaging: current strategies and future prospects. J Magn Reson Imaging 16(5): 497–510. doi:10.1002/jmri.10186

Annett M (1992) Five tests of hand skill. Cortex 28(4):583–600

Brechmann A, Baumgart F, Scheich H (2002) Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study. J Neurophysiol 87(1):423–433

Buchsbaum BR, Olsen RK, Koch PF, Kohn P, Kippenhan JS, Berman KF (2005) Reading, hearing, and the planum temporale. Neuroimage 24(2):444–454. doi:10.1016/j.neuroimage.2004.08.025

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I Segmentation and surface reconstruction. Neuroimage 9(2):179–194. doi:10.1006/nimg.1998.0395

Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53(1):1–15. doi:S1053-8119(10)00854-2[pii]10.1016/j.neuroimage.2010.06.010

Eden G, Joseph J, Brown H, Brown C, Zeffiro T (1999) Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: the behavior interleaved gradients technique. Magn Reson Med 41(1):13–20

Edmister WB, Talavage TM, Ledden PJ, Weisskoff RM (1999) Improved auditory cortex imaging using clustered volume acquisitions. Hum Brain Mapp 7(2):89–97

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25(4):1325–1335. doi:10.1016/j.neuroimage.2004.12.034

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9(2):195–207. doi:10.1006/nimg.1998.0396

Gaab N, Gabrieli JDE, Glover GH (2007a) 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 28(8):703–720. doi:10.1002/hbm.20298

Gaab N, Gabrieli JDE, Glover GH (2007b) Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. Hum Brain Mapp 28(8):721–732. doi:10.1002/hbm.20299

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7(3):213–223

Hall DA, Summerfield AQ, Gonçalves MS, Foster JR, Palmer AR, Bowtell RW (2000) Time-course of the auditory BOLD response to scanner noise. Magn Reson Med 43(4):601–606

Hart HC, Hall DA, Palmer AR (2003) The sound-level-dependent growth in the extent of fMRI activation in Heschl's gyrus is different for low- and high-frequency tones. Hear Res 179(1–2): 104–112

Herrmann CS, Oertel U, Wang Y, Maess B, Friederici AD (2000) Noise affects auditory and linguistic processing differently: an MEG study. Neuroreport 11(2):227–229

Hu S, Olulade O, Castillo JG, Santos J, Kim S, Jr GGT, Luh W-M, Talavage TM (2010) Modeling hemodynamic responses in auditory cortex at 1.5 T using variable duration imaging acoustic noise. NeuroImage 49 (4):3027–3038. doi:10.1016/j.neuroimage.2009.11.051

Jäncke L, Shah NJ, Posse S, Grosse-Ryuken M, Müller-Gärtner HW (1998) Intensity coding of auditory stimuli: an fMRI study. Neuropsychologia 36(9):875–883

Jäncke L, Mirzazade S, Shah NJ (1999) Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci Lett 266(2):125–128

Mechelli A, Price CJ, Henson RN, Friston KJ (2003) Estimating efficiency a priori: a comparison of blocked and randomized designs. Neuroimage 18(3):798–805

Moelker A, Pattynama PMT (2003) Acoustic noise concerns in functional magnetic resonance imaging. Hum Brain Mapp 20(3): 123–141. doi:10.1002/hbm.10134

Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13(4):684–701. doi:10.1006/nimg.2000.0715

Mueller K, Mildner T, Fritz T, Lepsien J, Schwarzbauer C, Schroeter ML, Moller HE (2011) Investigating brain response to music: a comparison of different fMRI acquisition schemes. Neuroimage 54(1):337–343. doi:10.1016/j.neuroimage.2010.08.029

Mulert C, Jäger L, Propp S, Karch S, Störmann S, Pogarell O, Möller H-J, Juckel G, Hegerl U (2005) Sound level dependence of the primary auditory cortex: simultaneous measurement with 61-channel EEG and fMRI. Neuroimage 28(1): 49–58. doi:10.1016/j.neuroimage.2005.05.041

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage 13(4):669–683. doi:10.1006/nimg.2000.0714

Schmidt CF, Zaehle T, Meyer M, Geiser E, Boesiger P, Jancke L (2008) Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. Hum Brain Mapp 29(1):46–56. doi:10.1002/hbm.20372

Schmithorst VJ, Holland SK (2004) Event-related fMRI technique for auditory processing with hemodynamics unrelated to acoustic gradient noise. Magn Reson Med 51(2):399–402. doi:10.1002/mrm.10706

Schwarzbauer C, Davis MH, Rodd JM, Johnsrude I (2006) Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fMRI. Neuroimage 29(3):774–782. doi:10.1016/j.neuroimage.2005.08.025

Shah NJ, Steinhoff S, Mirzazade S, Zafiris O, Grosse-Ruyken ML, Jäncke L, Zilles K (2000) The effect of sequence repeat time on auditory cortex stimulation during phonetic discrimination. Neuroimage 12(1):100–108. doi:10.1006/nimg.2000.0588

Talavage TM, Edmister WB (2004) Nonlinearity of FMRI responses in human auditory cortex. Hum Brain Mapp 22(3):216–228. doi:10.1002/hbm.20029

Talavage TM, Edmister WB, Ledden PJ, Weisskoff RM (1999) Quantitative assessment of auditory cortex responses induced by imager acoustic noise. Hum Brain Mapp 7(2):79–88

Tamer GG, Luh W-M, Talavage TM (2009) Characterizing response to elemental unit of acoustic imaging noise: an FMRI study. IEEE Trans Biomed Eng 56(7):1919–1928. doi:10.1109/TBME.2009.2016573

Woods DL, Stecker GC, Rinne T, Herron TJ, Cate AD, Yund EW, Liao I, Kang X (2009) Functional maps of human auditory cortex: effects of acoustic features and attention. PLoS ONE 4(4):e5183. doi:10.1371/journal.pone.0005183

Zaehle T, Schmidt CF, Meyer M, Baumann S, Baltes C, Boesiger P, Jancke L (2007) Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. Neuroimage 37(4):1195–1204. doi:10.1016/j.neuroimage.2007.04.073