NeuroImage 52 (2010) 1410–1419

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing

Jonathan E. Peelle [a,*], Rowena J. Eason [a], Sebastian Schmitter [b], Christian Schwarzbauer [a,1], Matthew H. Davis [a]

[a] *MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK*
[b] *German Cancer Research Center, Heidelberg, Germany*

ARTICLE INFO

Article history:
Received 22 February 2010
Revised 14 April 2010
Accepted 6 May 2010
Available online 16 May 2010

Keywords:
Sparse imaging
Quiet EPI
BOLD sensitivity
Listening effort

ABSTRACT

Echoplanar MRI is associated with significant acoustic noise, which can interfere with the presentation of auditory stimuli, create a more challenging listening environment, and increase discomfort felt by participants. Here we investigate a scanning sequence that significantly reduces the amplitude of acoustic noise associated with echoplanar imaging (EPI). This is accomplished using a constant phase encoding gradient and a sinusoidal readout echo train to produce a narrow-band acoustic frequency spectrum, which is adapted to the scanner's frequency response function by choosing an optimum gradient switching frequency. To evaluate the effect of these nonstandard parameters we conducted a speech experiment comparing four different EPI sequences: Quiet, Sparse, Standard, and Matched Standard (using the same readout duration as Quiet). For each sequence participants listened to sentences and signal-correlated noise (SCN), which provides an unintelligible amplitude-matched control condition. We used BOLD sensitivity maps to quantify sensitivity loss caused by the longer EPI readout duration used in the Quiet and Matched Standard EPI sequences. We found that the Quiet sequence provided more robust activation for SCN in primary auditory areas and comparable activation in frontal and temporal regions for Sentences>SCN, but less sentence-related activity in inferotemporal cortex. The increased listening effort associated with the louder Standard sequence relative to the Quiet sequence resulted in increased activation in the left temporal and inferior parietal cortices. Together, these results suggest that the Quiet sequence is suitable, and perhaps preferable, for many auditory studies. However, its applicability depends on the specific brain regions of interest.

Introduction

During echoplanar imaging (EPI), rapid switching causes the gradient coils to oscillate, resulting in the significant acoustic noise typically associated with fMRI scanning (Price et al., 2001; Ravicz et al., 2000). This acoustic noise, which can exceed 100 dBA, presents several serious challenges for studying the processing of auditory stimuli (Moelker and Pattynama, 2003). First, and perhaps most obviously, the high sound levels generated can render experimental stimuli unintelligible. Second, even if it is possible to hear the stimuli, listeners must separate them from the noise of the scanner, adding additional perceptual and cognitive demands to the task—that is, increases in listening effort (Davis and Johnsrude, 2003). Such task demands are likely to differentially affect participants with difficulties in auditory processing due to hearing impairment or normal aging (Grimault et al., 2001; Peelle and Wingfield, 2005; Wingfield et al., 2006). Finally, the acoustic noise of the scanner itself will activate auditory cortex (Bandettini et al., 1998), which may diminish effects induced by experimental manipulations (Elliott et al., 1999; Gaab et al., 2007). Thus, standard EPI sequences are sub-optimal for auditory tasks, and may make any results difficult to interpret. In addition, the reduction of acoustic noise may also be desirable for participant comfort, particularly when dealing with children or other special populations.

One common solution to the challenge posed by acoustic scanner noise is to use a sparse imaging procedure in which the repetition time (TR) of a sequence is longer than its acquisition time (TA), clustering slice acquisition in time in order to provide a silent period between the acquisition of consecutive volumes (Edmister et al., 1999; Hall et al., 1999; Scheffler et al., 1998). Auditory stimuli can then be presented during these silent periods without disruption from echoplanar scanner noise; the delay in the hemodynamic response to a stimulus enables the BOLD signal changes associated with these stimuli to be measured by the next volume of data acquired. Because of the longer TR, for a constant amount of scanning time, fewer images are acquired in a sparse imaging paradigm than in a continuous paradigm. This approach therefore reduces the temporal resolution of the data and, due to there being fewer observations, potentially reduces the accuracy of the parameter estimates (although this may be offset by higher overall

---

* Corresponding author.
E-mail address: jonathan.peelle@mrc-cbu.cam.ac.uk (J.E. Peelle).

[1] Current address: University of Aberdeen, Aberdeen, Scotland.

1053-8119/$ – see front matter. Crown Copyright © 2010 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.05.015

levels of signal due to the absence of spin history effects). Some modifications to sparse imaging paradigms have been developed to compensate for these shortfalls by collecting multiple volumes following a silent period (e.g., Schwarzbauer et al., 2006). Nonetheless, for a given amount of scanning time, sparse imaging approaches are fundamentally limited in the number of volumes that can be acquired relative to continuous sequences, and the extent to which differently-timed responses to a stimulus can be measured.

Another important consideration is the fact that responses in auditory regions are not only influenced by the amplitude of the scanner noise, but by other parameters, such as its perceived continuity: auditory cortex responds strongly to pulsed noises in the frequency ranges associated with typical gradient switching, and thus typical EPI sequences make for particularly effective stimulation (Giraud et al., 2000; Harms and Melcher, 2002; Seifritz et al., 2003; Tanaka et al., 2000). Thus, a second approach for addressing some of the issues faced in auditory fMRI studies is to change the qualitative nature of the acoustic noise. This approach was taken by Seifritz et al. (2006), who developed an EPI sequence that emits continuous noise (rather than pulsed) by implementing a quasi-continuous gradient switching pattern. The authors presented audio recordings of the noise generated by both types of sequences to participants and recorded neural responses using a sparse imaging paradigm. They found that conventional EPI produced stronger responses than the continuous noise EPI. Additionally, responses to pure tones in auditory cortex were greater when measured with continuous noise EPI relative to conventional EPI. These results emphasize that the nature or quality of acoustic stimulation from the scanner, and not just its average loudness, must be considered in auditory fMRI studies.

Although changing the acoustic characteristics of the scanner noise effectively boosts BOLD responses in auditory areas for some stimuli, it still leaves open the possibility of interference by energetic masking, and may also lead to extra challenges of listening effort, especially for more complex (e.g., linguistic) stimuli. One way to mitigate these effects is to use active noise control to minimize the effects of scanner noise (Chambers et al., 2007; Hall et al., 2009). Here we adopt an alternate approach to reducing the impact of acoustic noise by using an EPI sequence that is sufficiently quiet to allow participants to easily perceive auditory stimuli, even in the presence of pulsed scanner noise (Schmitter et al., 2008). In this sequence, acoustic noise is minimized by using a constant phase encoding gradient and a sinusoidal readout echo train to produce a narrow-band acoustic frequency spectrum. The scanner-specific frequency response function can be measured using an MR-compatible microphone placed inside the magnet bore. It is then possible to choose a readout gradient switching frequency (within the limits imposed by BOLD fMRI) that results in a lower acoustic response based on this frequency response function. In addition, the clicking noise of the slice-selection gradient is reduced by choosing a lower slew rate.

This modified gradient switching scheme can reduce the acoustic noise of EPI by up to 20–30 dB compared to trapezoidal EPI using the same imaging parameters. However, it may also influence data quality. For example, the longer EPI readout duration required by using a slower gradient switching frequency would be expected to exacerbate susceptibility effects near tissue boundaries, such as in inferior temporal and orbital frontal regions (Devlin et al., 2000; Ojemann et al., 1997). In addition, the nonuniform sampling of *k*-space requires an adaptation of standard image reconstruction software, and because of the sinusoidal readout gradient, the resulting images are also smoother than those from a standard sequence.

The primary aim of the current study is to evaluate the data provided by this new sequence relative to existing EPI sequences. We chose to do so using auditory stimuli that result in robust and replicable patterns of activation in well-known regions of cortex based on several previous studies (Davis et al., 2007; Mummery et al., 1999; Rodd et al., 2005).

Method

Participants

Six healthy right-handed adults aged 20–26 years (3 females) participated in this study. All were native English speakers with self-reported normal hearing and no history of neurological problems. Written consent was obtained from all participants on a protocol approved by the local ethics committee.

Materials

The materials consisted of a set of unambiguous sentences created for Rodd et al. (2005). Sentences ranged in duration from 1.14 to 3.58 s and contained simple declarative statements (e.g., “The police returned to the museum.”). The 120 sentences used were divided into 4 groups of 30, matched for duration, naturalness, imageability, and number of words using Match software (Van Casteren and Davis 2007), available from http://www.mrc-cbu.cam.ac.uk/people/maarten.van-casteren/mixandmatch.html. For each sentence, a probe word was generated for use in a behavioral task. Half of these probe words were semantically related to the sentence, and half were unrelated.

For a baseline condition, sentences matched in duration to the experimental sentences were used to create signal-correlated noise (SCN) (Schroeder, 1968). These stimuli have the same overall amplitude envelope and spectral profile as the original sentences but are lacking spectral detail, and are therefore entirely unintelligible. They thus provide a good control for the acoustic stimulation and temporal pattern of the sentences without conveying any linguistic information.

Procedure

Participants were instructed to attend to each stimulus, and after each sentence respond to the probe word on the screen. In the case of the real sentences, the probe word had a 50% probability of being semantically related to the just-heard sentence; participants indicated whether or not the probe word was semantically related to the sentence using a button-press response. For example, for the sentence “There were beer and cider on the kitchen shelf,” a related probe word might have been “drink”. For the SCN trials, the word “left” or “right” appeared on the screen, and participants were instructed to press the appropriate button. Participants were situated in the magnet and familiarized with this procedure through a short practice session, during which we also ensured the sentences were being presented at a comfortable listening level. Participants were informed that they would be performing the same task four times, but that due to different imaging parameters the scanner noise would differ. In all cases they were instructed to ignore the scanner noise as much as possible and concentrate on listening to the auditory stimuli.

Image acquisition

All images were acquired on a Siemens 3 T Tim Trio scanner (Siemens Medical Systems, Erlangen, Germany). The four EPI sequences used to collect data on the experimental paradigm are listed in Table 1 and described below. In addition to EPI data, we cquired a T1-weighted structural image for each participant using an MPRAGE sequence ($TR = 2250$ ms, $TE = 2.99$ ms, $TI = 900$ ms, flip angle $= 9°$, $FOV = 256\ \text{mm} \times 240\ \text{mm} \times 160\ \text{mm}$, voxel size $= 1\ \text{mm} \times 1\ \text{mm} \times 1\ \text{mm}$). In addition, field map data (2D structural and phase difference images) were acquired using a standard double echo GE sequence ($TE1/TE2 = 5.19/7.65$ ms; $TR = 400$ ms; flip angle $= 60°$, slice thickness $= 3$ mm; matrix size $= 64 \times 64$; in-plane resolution $= 3 \times 3$ mm; total acquisition time $= 54$ s). The phase difference images were unwrapped and converted into magnetic field maps (Jenkinson, 2003).