Published in final edited form as:

Magn Reson Med. 2015 August ; 74(2): 462–467. doi:10.1002/mrm.25408.

Less noise, more activation: multi band acquisition schemes for auditory fMRI

Federico De Martino^{1,2}, Michelle Moerel², Kamil Ugurbil², Elia Formisano¹, and Essa Yacoub²

¹ Department of Cognitive Neurosciences, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, 6229 ER Maastricht, The Netherlands. ² Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, 2021 sixth street SE, 55455 Minneapolis, Minnesota, USA.

Abstract

Purpose—To improve acquisition in fMRI studies of audition by using multiband (MB) gradient-echo echo planar imaging (GE-EPI).

Methods—Data were acquired at 3T (Siemens Skyra) with a 32-channel head coil. Functional responses were obtained by presenting stimuli (tones and natural sounds [voices, speech, music, tools, animal cries]) in silent gaps between image acquisitions. Two-fold slice acceleration (MB2) was compared to a standard GE-EPI (MB1). Coverage and sampling rate (TR=3s) were kept constant across acquisition schemes. The longer gap in MB2 scans was used to present: 1) sounds of the same length as in conventional GE-EPI (type1; 800 ms stimuli); 2) sounds of double the length (type 2; 1600 ms stimuli).

Results—Functional responses to all sounds (i.e. main effect) were stronger when acquired with slice acceleration (i.e. shorter acquisition time). The difference between voice and non-voice responses was greater in MB2 type1 acquisitions (i.e. same length sounds as GE-EPI but presented in a longer silent gap) than in standard GEEPI acquisitions (interaction effect).

Conclusion—Reducing the length of the scanner noise results in stronger functional responses. Longer “silent” periods (i.e. keeping the sound length the same as in standard acquisitions) result in stronger response to voice compared to non-voice stimuli.

Keywords

Multiband GE-EPI; auditory fMRI; scanner noise

Introduction

The presentation of auditory stimuli for functional magnetic resonance imaging (fMRI) studies is challenging (1). Echo-planar imaging (EPI) sequences (e.g. gradient echo-echo

planar imaging [GE-EPI]), the most common approach for fMRI, are particularly loud due to gradient switching during the EPI readout. As a result, particular care has to be taken in designing auditory fMRI experiments (i.e. in presenting auditory stimuli) in order to reduce contamination from auditory responses arising from the scanner noise.

Several acquisition schemes have been suggested for reducing acoustic noise (2-6), however, the most often used scheme is to present auditory stimuli in silent gaps between the acquisitions of two subsequent volumes (7-9) (lengthening the repetition time (TR)). While not immune from effects of scanner noise, this procedure takes advantage of the delayed peak of the hemodynamic response function (HRF) (~ 6 s) and enables the perception of the stimuli against a silent background. Variants of this include: 1) lengthening the TR to as long as 20 – 25 seconds and acquiring only one time point per sound presentation (i.e. sparse sampling) (7); 2) maintaining relatively short TRs (3 s) while limiting the effective acquisition time (TA), (8,9) or 3) the use of interleaved silent steady state (ISSS) (4,10). Sparse sampling approaches severely reduce the statistical power of the design. Presenting sounds in relatively short gaps between TRs, while being more statistically powerful, puts limits on spatial coverage (i.e. maximum number of slices) and spatial resolution. The ISSS technique overcomes some limitations of conventional sparse sampling, by allowing for a more flexible placement of image acquisition and silent periods. However, while it improves sampling of the BOLD signal, it does not reduce the acquisition time or gradient noise while also limiting the total number of slices. Further, the resulting time series are irregularly sampled, needing subsequent interpolation during preprocessing. Finally, “silent” acquisition schemes have been proposed based on sinusoidal gradients (3,5,6) and developing readout schemes with continuous rather than pulsed noise (2). While such an approach has advantages, such sequences are not widely available on standard clinical systems.

With the introduction of multi-band (MB) for EPI based fMRI, multiple 2D GE-EPI slices can be collected simultaneously (11-14), thereby permitting much shorter volume TRs. MB-EPI is becoming widely available and is being routinely used by many investigators and large scale projects around the world, such as the NIH's Human Connectome Project (HCP) (<http://www.neuroscienceblueprint.nih.gov/connectome/>). Here, we propose to use the increase in temporal efficiency provided by MB-EPI, not to reduce the volume TR, but to shorten the acquisition time TA, allowing for longer silent gaps in silent acquisition schemes (i.e. reducing the number of gradient ‘noises’ in the *same* TR). While there are different pulse sequence strategies for auditory fMRI (described above), the goal of this work is not to compare the different approaches, but rather to evaluate advantages of using MB-EPI over conventional single band EPI for the purposes of presenting auditory stimuli using the “sound-in-gap” approach. We compare the functional responses elicited by auditory stimuli of different categories (tones and natural sounds) and quantify the specific benefit of the proposed method by evaluating changes in specificity or sensitivity of voice selective responses in the human temporal lobe.

Methods

Six healthy volunteers with normal hearing participated in the experiment. All data were acquired on a Siemens Connectome Skyra (3T) using a multi-channel RF coil (1 Tx channel, 32 Rx channels) at the Center for Magnetic Resonance Research (Minneapolis, Minnesota, USA). The Institutional Review Board (IRB) of the University of Minnesota approved the imaging protocol used in this study.

Data Acquisition and Analysis

T₁ weighted (1 mm isotropic) scans were acquired using a Magnetization Prepared Rapid Acquisition Gradient Echo (MPRAGE) sequence (TR = 2400 ms; TI = 1000 ms; TE = 3.09 ms; flip-angle = 8 degrees; GRAPPA = 2; FOV = 256×256 mm; 256 slices).

Functional responses were measured using standard GE-EPI (voxel resolution 2×2×2 mm³; matrix = 96×96; TE = 33 ms; flip angle = 90 deg.; bandwidth = 2265; RF excitation length = 5 ms; TA = 1.5 s; TR = 3 s; 20 slices). A multiband acquisition was used to shorten the acquisition time from 1.5 s (MB1) to 750 ms (MB2). The resulting peak sound pressure level was of 106 db for both the MB1 and the MB2 acquisition. The sampling rate (TR) and total scan time were kept constant across all acquisition in order not to bias the subsequent statistical comparison. Since the MB accelerations *do not* 'under-sample', as in conventional phase encode accelerated imaging, there is no SNR loss from under-sampling. The end result was an MB2 image with nearly identical SNR as a conventional single band EPI image (12,13) (see figure 1, top). To quantify this effect we compared voxel-wise temporal signal to noise (tSNR; calculated as the temporal mean over standard deviation) between the MB1 and MB2 acquisitions (Figure 1, bottom). The results indicate that while a small (and expected) advantage of MB1 was present, overall MB2 provided comparable tSNR. Such accelerated acquisitions (i.e. with a high amount of retained SNR) are possible when using high channel counts in combination with *controlled* (slice) aliasing accelerations, which results in significantly reduced g-factors and residual aliasing effects (13,14).

Within the silent gaps, sounds (tones, voices, animal cries, nature sounds) were presented with durations of either 800 ms (for MB1 or MB2) or 1.6 s (MB2) (figure 2). The experimental design followed a standard block design in which each sound category (low frequencies, middle frequencies, high frequencies, voices/speech, animal cries, natural sounds, tools) were presented in blocks of 18 seconds followed by rest periods of 12 seconds. Two functional runs per scan type (MB1, MB2 with short sounds and MB2 with long sounds) were acquired for each subject. Sounds were presented using an MRI compatible ear-bud system (Sensimetrics). The order of acquisition was randomized in each subject. After the acquisition, subjects were asked if they perceived clear differences between acquisition types.

All data were analyzed with BrainVoyager QX (Brain Innovation, Maastricht, The Netherlands). Anatomical data were normalized in Talairach space and automatic segmentation was used to delineate the white matter gray matter boundary. Individual surface reconstructions of both hemispheres were used for cortex based alignment (CBA) (15). Functional data were pre-processed (slice scan time correction, motion correction, high

pass temporal filtering, temporal smoothing [two consecutive data points] and spatial smoothing [FWHM = 3 mm]), aligned to individual anatomical data and resampled in the normalized (Talairach) 3D volume at a resolution of 2 mm isotropic. A fixed effect general linear model (GLM) analysis was used to estimate functional responses for each condition, separately for each sequence type (MB1, MB2 with short sounds and MB2 with long sounds) in the 3D volume space. Correction for multiple comparisons was based on False Discovery Rate (FDR) (16) or cluster level correction (17). Resulting functional maps were sampled on the average cortical surface obtained after CBA.

Results

In addition to equal peak sound pressure level generated by the two acquisition types, the subjects did not report any perceptual difference between the acquisition types despite the MB2 TA being half the MB1 TA within the same TR. The only occasional report (2 out of the six subjects) was that in they perceived speech as more understandable (possibly due to the longer sounds in the MB2-long sounds acquisitions). All acquisition types resulted in robust activation of the temporal lobe bilaterally in each individual subject. In single subjects (figure 3) and at the group level (figure 4), the MB2-long sounds acquisition resulted in significantly ($q[FDR]=0.05$) stronger responses to natural sounds compared to the standard GEEPI (MB1) acquisition. At group level differences were located in the left posterior planum temporal (PT), right Heschl's Gyrus (HG), and the left anterior superior temporal gyrus (STG) (figure 4, see figure 3 top for results in a representative single subject). While in some subjects we detected significant differences in the responses between MB2-long sounds acquisition and MB1 when comparing the responses to simple tones (figure 3 top), these differences were not significant at group level (figure 4). On the other hand, responses to both tones and natural sounds were significantly ($q[FDR]=0.05$) stronger in MB2-short sounds acquisitions compared to MB1 in both single subjects (figure 3) and in the group (figure 4). At group level, differences were located on STG and PT bilaterally and right posterior STG for the comparison between responses to simple tones and STG bilaterally and left PT for the comparison between natural sounds (figure 4; see figure 3 for a representative single subject). In all single subjects and at group level we observed no significant differences when comparing the BOLD responses to sounds acquired with the two MB2 acquisitions.

When analyzing the differences between acquisition types with respect to voice selective responses (i.e. interaction between acquisition type and the contrast voices vs. non-voices) only the MB2-short sounds acquisition resulted in a significantly higher specificity (cluster size corrected $p<0.05$) than the MB1 acquisition in both single subjects (figure 3) and the group (figure 5). At group level, differences were located in classical voice selective regions (18) in the right medial and anterior superior temporal sulcus (STS) and left posterior STS (figure 5; see figure 3 for a representative single subject).

Conclusions

Our results show that reducing the acquisition time (i.e. decreasing the occurrences of scanner noise and increasing the silent gap within the same TR) produces significantly

stronger responses to simple and complex sounds in the auditory cortex when using sparse acquisition GE-EPI. In particular, stronger overall responses to tones can be detected if the longer gap is used to increase the quiet time (MB2-short sounds) while stronger responses to natural sounds can be obtained whether the longer gap is used to increase the quiet time or to increase the duration of the auditory stimuli (MB2-short sounds and MB2-long sounds). Interestingly, and contrary to the subjects feedback which indicated a behavioral advantage for the MB2-long sounds acquisition, only increasing the quiet time (MB2-short sounds) resulted in a significant interaction effect between acquisition types (MB1 and MB2-short sounds) and stimulus types (voices vs. non-voices). In other words, when using MB2 and short sounds the longer gap resulted in a stronger activation for voices and speech compared to animal cries, tools, and nature sounds in voice selective regions. This benefit in category selective responses goes beyond a simple increase in the overall response that has been previously demonstrated with other “quiet” acquisition schemes (2,3). This result leads us to speculate that given the choice of a repetition time, the use of slice acceleration can be beneficial in auditory studies especially when the temporal efficiency is used to lengthen the silent gap.

The temporal efficiency of MB-GE EPI, while previously shown to be advantageous in reducing the TR for fMRI (11-14), can significantly increase the sensitivity of auditory fMRI experiments without decreasing the TR. We have limited here the slice acceleration to a factor of two (MB2) as this was expected to not result in any significant decrease in SNR or temporal SNR (14,19), as confirmed by our analysis (figure 1). While higher MB factors are now routinely used (e.g. the 3T HCP protocol uses MB8) the effect of high slice acceleration with limited (z-direction) slice field of views (~ 4 cm here), as is typically used in auditory fMRI, versus whole brain protocols (~ 14 cm), remains to be evaluated. Of note, however, is that despite the small z-FOV the band separation in the protocol used here is similar (~ 2 cm) to the MB8 protocol, thus, in addition to insignificant SNR losses from g-factors and no SNR hit from TR shortening, the residual aliasing (leakage) effects are also expected to be small (14). Higher MB factors are expected to provide additional gains, but likely more incremental, compared to cutting the acquisition time in half with no noticeable loss in SNR, as is achieved in the MB2 acquisition employed here. Future studies will investigate the possibility of using higher acceleration factors as well as the application of multi-band acquisition to shortening the overall TR, which will permit higher resolutions, larger volume coverage, or more time to present more sounds in auditory fMRI experiments.

Acknowledgements

This work was supported in part by the National Institute of Health (NIH) Human Connectome Project (NIH U54MH091657), National Institute of Neurological Disorders and Stroke (NINDS) P30 NS076408, Biomedical Technology Resource Centers (BTRC) National Center for Research Resources (NCRR) P41 RR08079 and the National Institute of Biomedical Imaging and Bioengineering (NIBIB) P41 EB015894, the W.M. Keck Foundation, and MIND institute. The 7 T magnet purchase was funded in part by National Science Foundation DBI-9907842 and National Institute of Health S10 RR1395. EF was funded by NWO VICI grant 453-12-002.

References

1. Di Salle F, Esposito F, Scarabino T, Formisano E, Marciano E, Saulino C, Cirillo S, Elefante R, Scheffler K, Seifritz E. fMRI of the auditory system: understanding the neural basis of auditory gestalt. *Magnetic Resonance Imaging*. 2003; 21(10):1213–1224. [PubMed: 14725929]
2. Seifritz E, Di Salle F, Esposito F, Herdener M, Neuhoﬀ JG, Scheffler K. Enhancing BOLD response in the auditory system by neurophysiologically tuned fMRI sequence. *NeuroImage*. 2006; 29(3):1013–1022. [PubMed: 16253522]
3. Peelle JE, Eason RJ, Schmitter S, Schwarzbauer C, Davis MH. Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *NeuroImage*. 2010; 52(4):1410–1419. [PubMed: 20483377]
4. Schwarzbauer C, Davis MH, Rodd JM, Johnsrude I. Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI. *NeuroImage*. 2006; 29(3):774–782. [PubMed: 16226896]
5. Schmitter S, Diesch E, Amann M, Kroll A, Moayer M, Schad LR. Silent echo-planar imaging for auditory FMRI. *MAGMA*. 2008; 21(5):317–325. [PubMed: 18716815]
6. Zapp J, Schmitter S, Schad LR. Sinusoidal echo-planar imaging with parallel acquisition technique for reduced acoustic noise in auditory fMRI. *J Magn Reson Imaging*. 2012; 36(3):581–588. [PubMed: 22585371]
7. Formisano E, Kim DS, Di Salle F, van De Moortele P-F, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron*. 2003; 40(4):859–869. [PubMed: 14622588]
8. Staeren N, Renvall H, De Martino F, Goebel R, Formisano E. Sound categories are represented as distributed patterns in the human auditory cortex. *Current biology*. 2009; 19(6):498–502. [PubMed: 19268594]
9. Formisano E, De Martino F, Bonte M, Goebel R. “Who” is saying “what”? Brain-based decoding of human voice and speech. *Science (New York, NY)*. 2008; 322(5903):970–973.
10. Mueller K, Mildner T, Fritz T, Lepsien J, Schwarzbauer C, Schroeter ML, Möller HE. Investigating brain response to music: A comparison of different fMRI acquisition schemes. *NeuroImage*. 2011; 54(1):337–343. [PubMed: 20728550]
11. Moeller S, Yacoub E, Olman CA, Auerbach EJ, Strupp JP, Harel N, Ugurbil K. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med*. 2010; 63(5):1144–1153. [PubMed: 20432285]
12. Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Glasser MF, Miller KL, Ugurbil K, Yacoub E. Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS ONE*. 2010; 5(12):e15710. [PubMed: 21187930]
13. Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice Echo Planar Imaging with reduced g-factor penalty. *Magn Reson Med*. 2012; 67(5):1210–1224. [PubMed: 21858868]
14. Xu J, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, Yacoub E, Ugurbil K. Evaluation of slice accelerations using multiband echo planar imaging at 3T. *NeuroImage*. 2013; 83:991–1001. [PubMed: 23899722]
15. Goebel R, Esposito F, Formisano E. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping*. 2006; 27(5):392–401. [PubMed: 16596654]
16. Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*. 2002; 15(4):870–878. [PubMed: 11906227]
17. Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med*. 1995; 33(5):636–647. [PubMed: 7596267]
18. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature*. 2000; 403(6767):309–312. [PubMed: 10659849]

19. Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton CE, Zsoldos E, Ebmeier KP, Filippini N, Mackay CE, Moeller S, Xu J, Yacoub E, Baselli G, Ugurbil K, Miller KL, Smith SM. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*. 2014; 95C:232–247. [PubMed: 24657355]

Figure 1. Single EPI slices acquired in one subject with no slice acceleration (MB1,) and a slice acceleration factor of two (MB2). Scatter plot of voxel-wise temporal signal-to-noise ratio (tSNR) of MB1 and MB2 acquisitions (bottom).

Figure 2.

Schematic of the acquisition (black rectangles) and sound presentation (red rectangles) schemes. With no slice acceleration (MB1, top) sounds (800 ms long) are presented within a silent gap of 1500 ms. A slice acceleration of two (MB2) allows increasing the gap to 2250 ms while keeping the total repetition time unaltered (3000 ms). Two different sound presentation schemes are considered in the MB2 case: 1) short sounds (800 ms) (middle); 2) long sounds (1600 ms).

Figure 3.

Main effect (top; FDR corrected $q=0.05$) of sequence type (MB1 and MB2) in the overall response to tones and natural in a representative single subject and interaction effect (bottom; cluster level corrected $p<0.05$) between acquisition scheme and voice selective responses. Functional maps are superimposed to the partially inflated cortical surface of both hemispheres. The white dotted line indicates Heschl's Gyrus, the location of the planum temporale (PT), planum polare (PP) and superior temporal gyrus (STG) is highlighted in all hemispheres.

Figure 4.

Group (fixed effects analysis; FDR corrected $q=0.05$) difference in overall response to tones (top) and natural sounds (bottom) between MB2 (short sounds on the left; long sounds on the right) and MB1 acquisitions. Functional maps are superimposed to the partially inflated average cortical surface (after cortex based alignment) of both hemispheres. The white dotted line indicates Heschl's Gyrus, the location of the planum temporale (PT), planum polare (PP) and superior temporal gyrus (STG) is highlighted in all hemispheres.

MB2 - short sounds (voice selective) > MB1 (voice selective)

Figure 5.

Group (fixed effect; cluster level corrected $p < 0.05$) interaction effect between acquisition scheme and voice selective responses (i.e. MB2 short sounds [voice-nonVoice] > MB1 [voice-nonVoice]). Functional maps are superimposed to the partially inflated average cortical surface (after cortex based alignment) of both hemispheres (top) and to three transversal slices of the average anatomy in Talairach space (bottom). The white dotted line indicates Heschl's Gyrus, the location of the planum temporale (PT), planum polare (PP) and superior temporal gyrus (STG) is highlighted in both hemispheres.