NeuroImage 78 (2013) 295–304

Contents lists available at SciVerse ScienceDirect

NeuroImage

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

# Functional localization of the auditory thalamus in individual human subjects

Fang Jiang [a,*], G. Christopher Stecker [b], Ione Fine [a]

[a] Department of Psychology, University of Washington, Seattle, WA, USA
[b] Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA

ARTICLE INFO

*Article history:*
Accepted 8 April 2013
Available online 18 April 2013

*Keywords:*
Medial geniculate body
Auditory cortex
Functional magnetic resonance imaging

ABSTRACT

Here we describe an easily implemented protocol based on sparse MR acquisition and a scrambled 'music' auditory stimulus that allows for reliable measurement of functional activity within the medial geniculate body (MGB, the primary auditory thalamic nucleus) in individual subjects. We find that our method is equally accurate and reliable as previously developed structural methods, and offers significantly more accuracy in identifying the MGB than group based methods. We also find that lateralization and binaural summation within the MGB resemble those found in the auditory cortex.



## Introduction

Compared to vision, a great deal of auditory processing occurs sub-cortically, within areas such as the brainstem, midbrain, and thalamus (Ehret and Romand, 1997; Jones, 2003). The medial geniculate body (MGB) plays a key role in this pathway. Besides providing a thalamic relay between the inferior colliculus (IC) and the auditory cortex (AC), subregions of the MGB are involved in multiple ascending and descending auditory and multisensory pathways (see Winer et al., 2005, for a review). One difficulty in neuroimaging the human MGB is that it has proved quite difficult to reliably identify, either structurally or functionally, within individual subjects.

Individual thalamic nuclei are not distinctively revealed in typical T1- or T2-weighted structural scans. Devlin et al. (2006) described two methods for identifying the MGB anatomically in individual subjects: high-resolution proton density weighted scanning optimized for sub-cortical gray–white contrast, and tractography based on diffusion weighted imaging scans. Both methods can identify the MGB with reasonable reliability, but remain technically challenging. More recently, susceptibility weighted imaging (Haacke et al., 2009) has been proposed, but not yet validated, as a method of identifying the MGB. However SWI acquisition and analyses are challenging to implement.

When relying on functional data to localize a given area, there is a continuum of possible approaches, ranging from using a separate condition as a functional localizer to predetermine the area of interest at one extreme, to carrying out whole brain analyses for the contrast of interest and identifying focal activity in an appropriate region as MGN responses (see Friston and Henson, 2006; Friston et al., 2006; Saxe et al., 2006, for a discussion of the strengths and weaknesses of these two approaches across different paradigms).

However, to date, there is no validated technique for functionally localizing the MGB, and as a result most (though not all; Noesselt et al., 2010) papers examining MGB responses have relied heavily on group averaged responses to identify the MGB (Giraud et al., 2000; Griffiths et al., 2001; Krumbholz et al., 2005). This has been the case even when individual subject responses were of interest and an individual functional localizer ROI approach might therefore have been more optimal (Diaz et al., 2012; von Kriegstein et al., 2008). As described in more detail below, when comparing results across individuals using a group ROI, if a generous group ROI is chosen, then responses within individual subjects are likely to be averaged across an ROI that contains a high proportion of noise voxels. On the other hand, a more stringent choice of group ROI has the potential to underestimate MGB responses within individuals whose MGB falls outside the group ROI.

There are a number of potential reasons why functionally identifying the MGB may have proved so difficult. First, subcortical structures located near the brain stem can suffer from pulsatile motion effects. Factoring out these pulsatile motion effects using cardiac gating has been shown to improve signal to noise, but not sufficiently to allow for reliable imaging of the MGB within individuals (Guimaraes et al., 1998). However, BOLD modulation within the neighboring LGN can be imaged reliably without the need for cardiac gating (Schneider and Kastner, 2009; Wunderlich et al., 2005). Second, the MGB is small, with a volume of approximately 90 $mm^3$ in humans (5 mm wide, 4 mm deep, and 4–5 mm long, Winer, 1984). However visual responses within *subdivisions* (Haynes et al., 2005; Schneider et al., 2004) of the neighboring lateral geniculate nucleus have been measured. Thus it seems likely that neither cardiac motion nor the small size of the MGB fully explains why it is so difficult to obtain reliable individual responses within the MGB.

Another possibility is that MGB responses to commonly used experimental stimuli, such as simple tones, noises, or dynamic spectral ripples, might be relatively small, especially compared to the strong acoustic transients generated by environmental scanner noise. In

* Corresponding author.

1053-8119/$ – see front matter 
http://dx.doi.org/10.1016/j.neuroimage.2013.04.035

non-human primates, MGB neurons are responsive to natural and artificial sounds that vary along a diverse range of spectral and temporal feature dimensions (Allon and Yeshurun, 1985; Bartlett and Wang, 2011; Symmes et al., 1980). Moreover, the MGB has been demonstrated to show strong responses to complex speech-like stimuli (Diaz et al., 2012; von Kriegstein et al., 2008). Here we examined whether clearer responses within the MGB might be elicited by using stimuli containing spectrotemporally complex features (such as transients, broadband content, tonal elements and a low degree of predictability) while minimizing the effects of scanner noise by using sparse imaging.

Three experiments were carried out: in Experiment 1 we measured responses in the MGB to scrambled music in a passive listening task. In Experiment 2 we measured responses in the MGB to scrambled music while subjects carried out a one-back task where they had to identify when a music segment was consecutively repeated multiple times. In Experiment 3 we measured responses in the MGB while subjects passively listened to a dynamic ripple stimulus. These experiments demonstrate that a complex stimulus (scrambled music) containing complex elements (transients, broadband content, tonal elements and a low degree of predictability) can reliably identify the MGB within individuals when combined with a sparse acquisition protocol.

## Methods

### *Participants*

A total of 11 young adults (4 males; 1 left-handed; 27.4 ± 4.7 years old) participated across all three experiments. Nine subjects carried out Experiment 1 (passive listening) and nine subjects carried out Experiment 2 (1-back task): there was an overlap of 7 subjects between Experiments 1 and 2. The order of experiments was counterbalanced for subjects who participated in Experiments 1 and 2. Four of the subjects who participated in Experiment 1 also carried out Experiment 3 (dynamic ripple).

All participants reported normal hearing and no history of neurological or psychiatric illness. Written and informed consent was obtained from all participants prior to the experiment, following procedures approved by the Institutional Review Board of the University of Washington Human Subjects Division of the University of Washington.

### *Auditory stimuli*

Auditory stimuli were delivered via MRI-compatible stereo headphones (Sensimetrics S14, Malden MA) and sound intensity was adjusted to each individual participant's comfort level. The intensity in the binaural condition was scaled by −6 dB in each ear relative to the monaural case to equate the total sound amplitude across monaural and binaural conditions. Equating the amplitudes in this way reduces differences in loudness across conditions, but unfortunately did not allow us to measure binaural interactions.

#### *Scrambled music*

For both Experiment 1 and Experiment 2, auditory stimuli consisted of scrambled musical segments extracted from popular music including "God shuffled his feet" (Crash Test Dummies), "Will o' the wisp" (Miles Davis), and "Saeta" (Miles Davis). Both Miles Davis tracks consisted of music only, and Crash Test Dummies track contained lyrics. For each scan, only one sound file (i.e., one song) was used, and the order of the files was kept the same for all participants. Overall sound levels were scaled to equate amplitude across three songs. Scrambling was done by reading song files into MATLAB (Mathworks, MA), subdividing these files into 900 ms segments, and then presenting these 900 ms segments in a scrambled order. Each 8 s stimulus presentation interval consisted of 8 randomly selected 900 ms segments, separated by 100 ms silent intervals, presented either monaurally or to both ears.

#### *Dynamic ripple*

For Experiment 3, the auditory stimulus was dynamic ripple, a spectrally and temporally modulated complex broadband stimulus that has been used to study BOLD responses in the auditory cortex (Langers et al., 2003; Schonwiesner and Zatorre, 2009). Following Lanting et al. (2008), we used a dynamic-ripple stimulus that consisted of temporally and spectrally modulated noise, with a frequency range of 125–8000 Hz, a spectral modulation density of one cycle per octave, a temporal modulation frequency of two cycles per second, and a modulation-amplitude of 80%. This dynamic-ripple stimulus was presented either binaurally or monaurally in a single session, using analogous methods as for the passive scrambled music experiment (Experiment 1).

### *Procedure*

All participants were instructed to close their eyes and pay attention to the auditory stimulus. In *Experiment 1* and *Experiment 3* there was no task (mimicking the passive localizer stimuli traditionally used to identify the LGN, Schneider and Kastner, 2009). In *Experiment 2*, we included a one-back task, where participants were required to press the response button when they detected a consecutively repeated 900-ms segment. Such repetitions occurred randomly, 3–4 times during each scan.

During a stimulus presentation interval, scrambled musical segments were delivered either to both ears (*binaural condition*), to the right ear (*monaural right condition*), or to the left ear (*monaural left condition*). The monaural right condition is described as the *contralateral* condition for regions in the left hemisphere and as the *ipsilateral* condition for regions in the right hemisphere, and vice versa for the monaural left condition. We also included a 4th condition in which no sound was delivered during the stimulus presentation interval (*silence condition*). Conditions were presented in a fixed order (binaural, monaural right, monaural left, and silence) across all experiments. Each condition was repeated 8 times in a scan for a total of 32 8 s auditory stimulus presentation intervals (each followed by 2 s MR acquisition). Each scan therefore lasted for 320 s (10 s × 4 conditions × 8 reps). Each subject carried out six scans, which resulted in a total of 48 repetitions per condition over the course of scanning for each of the three experiments.

### *MRI scanning*

Blood oxygenation-level dependent (BOLD) functional imaging was performed with a 3 T Philips system at the University of Washington Diagnostic Imaging Sciences Center (DISC). The scan protocol consisted of 2.75 × 2.75 × 3 mm voxels; repetition time, 10 s; echo time, 16.5 ms; flip angle, 76°; field of view, 220 × 220; and 32 transverse slices. Three-dimensional (3D) anatomical images were acquired at 1 × 1 × 1 mm resolution using a T1-weighted MPRAGE (magnetization-prepared rapid gradient echo) sequence.

A sparse echo planar imaging pulse sequence was used so that stimulus presentation was uninterrupted by acoustic MRI scanner noise (Hall et al., 1999). 2 s volume acquisitions were preceded by an 8 s delay period during which there was no scanner noise and the auditory stimuli were delivered (Fig. 1). Because of a hemodynamic delay of about 4–5 s to peak response within auditory cortex (Inan et al., 2004; Jancke et al., 1999), each volume acquisition measures BOLD response to stimulation during the middle of the stimulus presentation period, with relatively little contribution from the acoustic scanner noise of the previous acquisition. It is worth noting that the longer delay between acquisitions (which allows for more time to restore magnetic equilibrium) results in a higher signal-to-noise ratio