Feeling the beat: premotor and striatal interactions in musicians and non-musicians during beat perception

Jessica A. Grahn¹ and James B. Rowe¹,²,³
¹MRC Cognition and Brain Sciences Unit, Cambridge UK
²Cambridge University Department of Clinical Neuroscience, Cambridge, UK
³Behavioural and Clinical Neurosciences Institute, Cambridge, UK

Abstract

Little is known about the underlying neurobiology of rhythm and beat perception, despite its universal cultural importance. Here we used functional magnetic resonance imaging to study rhythm perception in musicians and non-musicians. Three conditions varied in the degree to which external reinforcement versus internal generation of the beat was required. The ‘Volume’ condition strongly externally marked the beat with volume changes, the ‘Duration’ condition marked the beat with weaker accents arising from duration changes, and the ‘Unaccented’ condition required the beat to be entirely internally generated. In all conditions, beat rhythms compared to nonbeat control rhythms revealed putamen activity. The presence of a beat was also associated with greater connectivity between the putamen and the supplementary motor area (SMA), the premotor cortex (PMC) and auditory cortex. In contrast, the type of accent within the beat conditions modulated the coupling between premotor and auditory cortex, with greater modulation for musicians than non-musicians. Importantly, the putamen's response to beat conditions was not due to differences in temporal complexity between the three rhythm conditions. We propose that a cortico-subcortical network including the putamen, SMA, and PMC is engaged for the analysis of temporal sequences and prediction or generation of putative beats, especially under conditions that may require internal generation of the beat. The importance of this system for auditory-motor interaction and development of precisely timed movement is suggested here by its facilitation in musicians.

Keywords

Rhythm; beat perception; basal ganglia; fMRI; timing; effective connectivity

Appreciation of musical rhythms is an important feature of human culture. A key feature of rhythm is an underlying regular beat: a perceived pulse that marks equally spaced points in time (Cooper and Meyer, 1960; Lerdahl and Jackendoff, 1983). Perception of the beat enables temporal intervals marked by stimulus onsets to be encoded as multiples or subdivisions of the beat, rather than as unrelated intervals, and improves rhythm reproduction and discrimination (Drake and Gerard, 1989; Ross and Houtsma, 1994; Hebert and Cuddy, 2002; Patel et al., 2005). Beat perception can feel automatic, and occurs without musical training even in young children, but little is known about the neural mechanisms.
Humans often move to the beat (Drake et al., 2000), suggesting motor networks may be important. Neuroimaging has confirmed activity in ‘motor areas’ during production and perception of rhythm (Schubotz et al., 2000; Schubotz and von Cramon, 2001; Grahn and Brett, 2007; Chen et al., 2008b). However, no consensus exists on the non-motor roles these areas play in rhythm or beat perception.

To perceive the beat, several cues are used, usually involving accents. An accent is an increase in salience when an event differs from surrounding events (Cooper and Meyer, 1960; Povel and Okkerman, 1981; Parnicutt, 1994). Accents can be ‘dynamic’ involving volume changes (Chen et al., 2006) or ‘temporal’ involving duration changes (Essens and Povel, 1985; Palmer and Krumhansl, 1990). ‘Subjective’ accents are also perceived even without external accents (for example, in a sequence of tones identical in volume, duration, pitch, etc.) as listeners internally emphasize certain tones in the sequence (Bolton, 1894; Temperley, 1963; Brochard et al., 2003). Listeners’ perceptions appear to be influenced by musical training in some studies (Yee et al., 1994; Drake et al., 2000) but not others (Snyder and Krumhansl, 2001). Musical training effects may arise from extensive engagement in encoding, memorizing, and performing rhythms with a beat.

In two complementary experiments, we used functional magnetic resonance imaging (fMRI) to study beat perception for rhythms with different accent types in musicians and non-musicians (see Figure 1). We used rhythms with regular volume accents (strong external beat emphasis), duration accents (weaker external beat emphasis) or no accents (no external beat emphasis). We present analyses of both regional activation and connectivity observed for beat perception in all three conditions, and also compare the effects of internal versus external beat generation.

We made four predictions. First, basal ganglia activity would increase for beat versus nonbeat rhythms, extending findings of basal ganglia involvement in duration beat rhythms (Grahn and Brett, 2007; 2009). Second, internal beat generation would modulate basal ganglia activity, consistent with the basal ganglia’s role in internally generated movements (Freeman et al., 1993; Mushiake and Strick, 1995; van Donkelaar et al., 1999). Third, beat rhythms would be associated with greater coupling between basal ganglia and cortical rhythm areas. Finally, musical training would enhance connectivity within the neural networks mediating beat perception, as it does for auditory-motor synchronization (Chen et al., 2008a).

## Methods

### Participants

36 participants (21 male, 15 female, age range 18-41 years, mean 29) took part in experiment one and experiment two (in that order) on the same day after providing written informed consent. Nineteen had musical training defined as five or more years of formal musical training and continuing musical activities. Seventeen had no musical training, defined as no current or previous formal musical training or musical performance activities. All participants reported having normal hearing.

### Stimuli

For experiment one, volume accented and duration accented rhythmic stimuli were used. The stimuli were between 14 and 18 s long. There were four rhythm types in a 2×2 factorial design, with the factors Beat (Beat/Nonbeat) and the dimension that varied to indicate the beat (Volume/Duration). See Figure 1 for schematic depictions of the stimuli. Thus, the first rhythm type (Volume accented with Beat) consisted of 81 tones, in which every 4th tone was louder by 8.5 dB, in order to give rise to the perception of a regular beat (occurring 21 times
per trial). For each trial, the tone length was chosen from a range of 180 to 228 ms (in 8 ms steps) so that a new beat would be induced in each trial, not simply carried over from a previous trial. Accordingly, the beat occurred at a rate of 720 to 912 ms. The second rhythm type (Volume accented with No beat) also had 81 tones. However, the tone lengths were not isochronous, so no regular beat could be fit to the rhythm. The Volume Nonbeat rhythms were created from each Volume Beat rhythm. One-third of the original intervals were shortened by 30%, another third were lengthened by 30%, and the remaining third stayed the same. All the intervals were randomly reshuffled, and 21 tones were randomly chosen to receive a volume accent (to be comparable to the Volume Beat condition). Thus, the number of intervals, overall rhythm length, number of volume accents, and overall RMS intensity was equivalent between the Volume Beat and the Volume Nonbeat conditions.

The Duration Beat condition was constructed from twenty patterns used in previous experiments (Grahn and Brett, 2007) that rely on the temporal context (i.e. the relative durations) to give rise to the perception of a regular beat (Povel and Okkerman, 1981; Grahn and Brett, 2007). See Supplementary Table 1 for a list of the patterns used. Five of these patterns were randomly selected for each trial, and each trial consisted of all five patterns played without a break (as one continuous rhythm). Each pattern was used three times across the experiment. For each Duration Beat trial, an equivalent Duration Nonbeat trial was created. First, all the intervals from the Duration Beat trial were divided into their respective categories (4 different interval lengths were present in each trial). Then, one-third of the intervals in each category were shortened by 30%, one-third remained the same, and one-third of the intervals were lengthened by 30%. All the intervals were then randomly reshuffled. Thus, the number of intervals, overall rhythm length, and RMS intensity was equivalent between the Duration Beat and the Duration Nonbeat conditions.

For all conditions, 500 Hz sine tones (rise/fall times of 8 ms) sounded for the duration of each interval, ending 40 ms before the specified interval length to create a silent gap that demarcated the intervals. The sequences used filled intervals, as filled intervals provide the benefit of attenuation of environmental noise (e.g., that experienced during MRI). The tones in the Duration conditions were approximately 6 dB softer than accented tones in the Volume condition and 2.5 dB louder than the unaccented tones in the Volume condition. The overall RMS intensity was equated between the Volume and Duration conditions, so that the average level of auditory stimulation was the same.

In Experiment Two, the Volume Beat and Volume Nonbeat conditions were as above, but shortened to 61 tones each. In addition, unaccented versions of each condition also were created, by removing amplitude modulation (see Figure 1). The intervals in the unaccented conditions (Unaccented Beat and Unaccented Nonbeat) were therefore temporally identical to the Volume Beat and Volume Nonbeat conditions. Unaccented condition sequences were presented at the same dB level as the Duration condition sequences.

**Behavioural Design**

Before scanning, participants heard a randomly selected sample of each of the rhythm types. After listening to each rhythm, they answered three questions about the beat, using rating scales of 1 to 10. The first question was “How much did this sound have a beat?” Pilot studies indicated that many non-musicians were not confident when answering this question, as it was perceived to be an objective property of the rhythm that they were unqualified to judge. Therefore, two more subjective questions were included: “How easy was it to feel a beat?” and “How clear was the beat?”.
fMRI Experimental Design

Experiment one was designed to enable both regional analysis of activation differences between conditions and connectivity analyses between areas. Experiment two was initially conceived of as a supplementary experiment that controlled for differences in temporal complexity. The Volume and Duration conditions in experiment one differ in the degree of external accent present, but they also differ in temporal complexity. However, in experiment two, the Volume and Unaccented conditions are balanced in terms of temporal complexity, while still differing in the degree of external reinforcement of the beat. Thus, any activation differences observed between Volume and Duration conditions in experiment one would be unlikely due to differences in complexity if they were also present in contrasts between Volume and Unaccented conditions in experiment two. Concerns about time constraints and participant fatigue led us to keep experiment two brief, focusing on the analysis of differences in regional activation associated with different accents without difference in temporal complexity. As a result, experiment 2 was not optimized for analyses of connectivity.

Rhythms were presented diotically over headphones with 30 dB attenuation of scanner noise by insert earplugs (3M 1100 earplugs, 3M United Kingdom PLC, Bracknell, UK). None of the participants reported difficulty in hearing the rhythms. Head fixation used foam pads.

For experiment 1, a 14-18s rhythm was presented on each trial, followed by a 12.4 s rest period. Eleven trials of each rhythm type (Volume Beat, Volume Nonbeat, Duration Beat, and Duration Nonbeat) and a subsequent rest period were presented in permuted order. To ensure participants' attention to the rhythms, they completed a pitch change detection task, pressing a key if they heard a tone that was different in pitch to all the other tones (450 Hz, presented 20 times during the session). Piloting indicated this frequency difference was distinguishable by both musicians and non-musicians, but did not lead to performance at ceiling. Frequency detection was chosen as it was unlikely to encourage any tapping or subvocalizing of the rhythm. In addition, the relative rarity and unpredictability of the deviant meant the observed effects of interest could be unconfounded with response preparation. Participants were told that they would be listening to rhythms similar to those rated before scanning, and that their task was to press the button under their index finger each time they heard a change in the pitch of the rhythm. They were also told that the aim of the study was to examine brain activity when ‘feeling the beat’, and to focus on the beat in addition to listening for pitch changes. Finally, they were instructed not to move any part of the body during scanning, apart from the index finger to press the button when a deviant tone was heard. Participants were visually monitored throughout to ensure no tapping or other movement occurred.

For experiment 2, trials of Volume Beat, Volume Nonbeat, Unaccented Beat, and Unaccented Nonbeat rhythms were played without deviant tones. Participants were again instructed to focus on feeling the beat, but not to move. Eight 11-13.5 s rhythms of each type were played alternating with 4 seconds of rest, in a balanced permuted order.

Image Acquisition

A 3T Siemens Tim Trio MRI scanner was used to collect 580 echoplanar imaging (EPI) volumes in experiment one followed by 261 EPI volumes in experiment two. All EPI data had 36 slices, matrix size of 64 × 64, TE = 30 ms, TR = 2.19 s, FOV = 19.2 × 19.2 cm, flip angle = 78°, slice thickness 3mm, interslice distance of 0.75 mm, and in-plane resolution of 3 × 3 mm. High-resolution MPRAGE anatomical images (TR = 2250 ms, TE = 2.99 ms, flip angle = 9°, IT 900 ms, 256×256×192 isotropic 1 mm voxels) were collected for anatomic.
localization and coregistration. The total time each participant spent in the scanner was approximately 40 minutes.

**Image analysis**

SPM5 was used for data analysis (SPM5; Wellcome Department of Imaging Neuroscience, London, UK). Images were sinc interpolated in time to correct for acquisition time differences and realigned spatially with respect to the first image using trilinear interpolation. The coregistered MPRAGE image was segmented and normalized using affine and smoothly nonlinear transformations to the T1 template in Montreal Neurological Institute (MNI) space. The normalization parameters were then applied to the EPIs and all normalized EPI images were spatially smoothed with a Gaussian kernel of full-width half-maximum 6 mm.

Subject-specific first level models included epochs representing the four conditions and a transient event for the deviant tone/button press (experiment one only), convolved by the canonical hemodynamic response function. EPI volumes associated with discrete artifacts were included as covariates of no interest (including volume displacement >4 mm or spikes of high variance in which scaled volumes to volume variance was 4 times greater than the mean variance). Autocorrelations were modeled using an AR(1) process and low-frequency noise was removed with a standard high-pass filter of 128 seconds.

Relevant contrast images (voxel-wise differences in parameter estimates) from single participant models were entered into second-level random effects analyses for group inference (Penny and Holmes, 2003). In experiment one, results are presented at threshold FDR p < 0.05 for whole brain comparisons. As experiment two contained fewer scans than experiment one, and consequently less power, a reduced search volume was defined from significant voxels obtained in the Beat – Nonbeat contrast in experiment one (FDR corrected p < 0.05) and used for the Beat – Nonbeat contrast in experiment two. In effect, therefore, experiment two used the Beat-Nonbeat contrast from experiment one as a localizer. This was motivated by the restricted principal aims of experiment two (a control for temporal complexity in beat perception and replication of key effects from experiment one), and in recognition of reduced statistical power of experiment two.

**Effects of stimulus rate**

Additional analyses were conducted to determine if activity in areas that responded to the beat also correlated with the rate, or tempo, of the stimuli. This was to ensure no rate confound could arise between our conditions. Please see Supplementary methods for details of this analysis.

**Psychophysiological Interaction Analysis**

The physiological interaction between two regions may be modulated by the experimental or ‘psychological’ context, representing a psychophysiological interaction (PPI). PPI analysis by general linear models provides an anatomically unconstrained method to analyse changes in connectivity between regions in a network (Friston et al., 1997).

For the first experiment, we analysed of the differential connectivity of the left and right anterior and posterior putamen foci under beat vs. nonbeat conditions. As both the left and right putamen showed similar results (no differences at p < 0.5 FDR) the results are collapsed across both putamen foci. To construct the model, the deconvolved time courses of activity in the foci were extracted from the F-contrast of all effects of interest. The regions of interest were spheres (8-mm-diameter) centred at: R anterior putamen: x = 26, y = 8, z = 8; L anterior putamen: x = −26, y = 2, z = 6; R posterior putamen: x = 30, y = −16, z = 8; L
posterior putamen: x = -32, y = -18, z = 0. These activation time courses constituted the first regressor in the PPI analyses. Then three experimental context variables were entered as the second, third, and fourth regressor. The first experimental context variable examined the two Beat conditions: comparing beat in the context of the Duration condition versus the context of the Volume condition (Duration Beat = 1, Volume Beat = -1, Duration Nonbeat = Volume Nonbeat = 0). The second experimental context variable examined the two Nonbeat conditions, again in the context of Duration accents versus Volume accents (Duration Nonbeat = 1, Volume Nonbeat = -1, Volume Beat = Duration Beat = 0). The third experimental context variable was the effect of Beat versus No beat, collapsed across accent type (Duration Beat = Volume Beat = 1, Duration Nonbeat = Volume Nonbeat = -1). The 

product of the neural time course with each psychological factor was calculated then reconvolved by the hemodynamic response function to create three PPI terms. The effects of these interaction terms were investigated for each subject and each putamen region and entered into a second-level random-effects model. A regions-of-interest image was created from the areas showing significant rhythm-related activity (8-mm diameter spheres on peak voxels of superior temporal gyri, premotor and supplementary motor cortices, and the cerebellum in the all rhythms – rest contrast). This image was used as a compound region of interest (with small volume correction, SVC) in the PPI analyses.

Secondary PPI analyses were conducted on the cortical areas showing a specific modulation of connectivity with the putamen during Beat versus No beat conditions. The source regions of interest for these PPIs were spheres (8-mm diameter) centred at the peaks of activity in the all rhythms – rest contrast. These regions included the right PMC (x = 54, y = 0, z = 48), left PMC (x = -52, y = -10, z = 50), right SMA (x = 6, y = 0, z = 66), and left SMA (x = -6, y = -4, z = 66). The PPI analyses were conducted as described above, using the same SVC image as above.

A similar connectivity analysis was conducted for experiment two, although this second study was underpowered for PPI analysis due to fewer scans and shorter rest intervals. We specifically tested if the pattern of connectivity observed for Duration and Volume Beat versus Nonbeat conditions in experiment one was also present for the Unaccented Beat versus Nonbeat condition in experiment two. The bilateral anterior putamen foci from experiment one were used as seed regions. Peaks within 1 cm of PPI-related peaks observed in experiment one are reported with uncorrected p-values.

Results

Behavioral Results

The three behavioural ratings of beat (how much was a beat present, how easy was it to feel the beat, how clear was the beat) were highly correlated across subjects within each condition (0.70 < r < .99, all p < .001). Therefore, an average of the three ratings for each subject was taken as the measure of subjective experience of beat. Analyses were also conducted on the individual ratings measures in addition to an averaged measure. No differences in the pattern of results were found, so, for conciseness, only statistics for the average measure are reported.

Beat ratings are illustrated in Figure 1. The ratings underwent a mixed ANOVA which revealed significant main effects for Beat (F(1,34) = 115.51, p < .001) and Accent type (F(1,34) = 19.70, p < .001), but not for Musical training (F < 1). There was also an interaction between Beat and Musical training (F(1,34) = 4.34, p = .045), due to musicians rating beat conditions higher and nonbeat conditions lower than non-musicians. There was also a significant interaction between Beat and Accent type (F(2,68) = 3.15, p = .049), reflecting that ratings for Volume Beat > Unaccented Beat > Duration Beat, but Volume...
Nonbeat = Unaccented Nonbeat > Duration Nonbeat. No other significant effects were found. Behavioural data for the pitch-deviant detection collected during the scanning session is also presented in the supplemental data.

**Regional changes in brain activation**

**Experiment one**

Listening to rhythms (collapsed across all types and compared to rest) activated the supplementary motor area (SMA), bilateral premotor cortex (PMC), superior temporal gyri (STG), insula, putamen, cerebellum, and right middle frontal gyrus (BA 45/46), supramarginal gyrus, and inferior frontal gyrus (BA 44/45). See Supplementary Table 2 for exact coordinates.

A main effect of Beat (Volume Beat + Duration Beat – Volume Nonbeat + Duration Nonbeat) was found most strongly in the putamen bilaterally (see Table 1 and Figure 2a). A main effect of Volume accents (Volume Beat + Volume Nonbeat – Duration Beat + Duration Nonbeat) was found most strongly in auditory cortex, but also in the right frontal cortex (Brodmann areas 6, 8, 9, 44, 45, 46), left frontal cortex (Brodmann areas 8, 43, 44) and bilateral cerebellum (Crus II). See supplementary Table 3 for exact coordinates.

Within the Duration-accented conditions (Duration Beat-Duration Nonbeat), the beat condition significantly activated the putamen bilaterally. No significant activations were found for the reverse contrast (Duration Nonbeat-Duration Beat). These findings replicate previous results (Grahn and Brett, 2007). Between the Volume-accented conditions, however, a different pattern was observed. The beat minus the nonbeat condition (Volume Beat-Volume Nonbeat) revealed no significant activations. However, the nonbeat minus the beat (Volume Nonbeat-Volue Beat) condition activated the bilateral pre-SMA, PMC (BA6), STG, insula, inferior frontal gyrus (BA 44/45), and the cerebellum (lobules VI and VIII on the left, lobule VI on the right). See supplementary Table 4 for detailed coordinates.

Several areas were sensitive to the interaction between Beat and Accent type ((Duration Beat-Duration Nonbeat) – (Volume Beat-Volume Nonbeat)). In some areas the interaction was due to greater activity in the Duration Beat – Duration Nonbeat contrast than in the Volume Beat – Volume Nonbeat contrast. In other areas, including the STG and cerebellum, the interaction is due to significantly decreased activity in the Volume Beat – Volume Nonbeat contrast compared to the Duration Beat – Duration Nonbeat contrast (Supplementary Figure 1 and Supplementary Table 5 illustrate the interaction effects in more detail). There were no significant differences in regional activation associated with musical training.

**Experiment two**

The comparison of Beat versus No beat ((Unaccented Beat + Volume Beat) – (Unaccented Nonbeat + Volume Nonbeat)) revealed activation of putamen bilaterally (see Table 1 for a list of maxima, and Figure 2a: results corrected FWE p<0.05 within a reduced search volume defined from the FDR corrected equivalent contrast in the first experiment). The average activity of individual basal ganglia structures (ROIs of the putamen, pallidum, and caudate (Tzourio-Mazoyer et al., 2002)) is shown in Figure 2b. The graph depicts signal intensity extracted from the Beat – Nonbeat contrast images for each condition (Volume and Duration from experiment one, and Unaccented from experiment two).

We compared activation patterns across experiments one and two for an equivalent contrast common to both experiments (Volume Beat - Volume Nonbeat). In every ROI reported (all bilateral basal ganglia structures, cerebellum, premotor cortex, SMA, insula, right
supramarginal gyrus, and right middle frontal gyrus), no significant differences were found in activity between experiment 1 and experiment 2 for this contrast. Thus, the slightly shorter stimuli and absence of the monitoring task in experiment two did not significantly affect activation in rhythm network areas.

**Effects of stimulus rate**

Activation within the basal ganglia was not significantly correlated with rate in any of the rhythm conditions. See Supplementary results for more detail.

**Changes in connectivity between regions**

The first set of PPI analyses compared the connectivity patterns of the anterior and posterior putamen to other brain areas for experiment one. During the Volume and Duration Beat conditions compared to the Volume and Duration Nonbeat conditions, the anterior putamen showed increases in connectivity with the PMC, SMA, and right STG (p < .05, SVC see Methods) and a similar trend in the left STG and right cerebellum (p < .1, SVC). This is illustrated in Figure 3, top. For experiment two, similar, although less robust, results were observed for the Unaccented Beat – Unaccented Nonbeat condition. All reported peaks are within 1 cm of peaks for Experiment one. The anterior putamen showed increases in connectivity with the bilateral PMC (Right: x = 58, y = 2, z = 40, t = 1.82; Left: x = -56, y = 2, z = 40, t = 1.74), cerebellum (Right: x = 58, y = 2, z = 40, t = 1.88; Left: x = -32, y = -60, z = -32, t = 1.66), at p <.05 uncorrected. Trends were observed in bilateral SMA (Midline: x = 0, y = -2, z = 60; t = 1.44, p = .07 uncorrected; Left: x = -4, y = -6, z = 60, t = 1.56, p = .06 uncorrected) and left STG (x = -50, y = -20, z = 2, t = 1.28, p = .1 uncorrected). Effect sizes were similar between experiment one and experiment two for putamen connectivity with motor areas (R PMC: .052 vs .060; L PMC: .074 vs .048; R SMA: .048 vs .060; L SMA: .040 vs .056; L cerebellum: .027 vs .026; R cerebellum: .036 vs .032). This suggests that lack of significance at a corrected threshold was due to increased variability in the underpowered design, and that the Beat-Nonbeat connectivity patterns between putamen and motor areas across experiment one and two were qualitatively similar.

There were no differences in cortico-subcortical coupling for musicians and non-musicians. When comparing the Volume Beat to the Duration Beat condition, no significant differences in connectivity were observed between anterior putamen and other areas. The posterior putamen PPI yielded no significant changes in connectivity for either Beat versus No Beat, or Volume Beat versus Duration Beat.

Brain areas whose activity shows greater coupling with the putamen in beat conditions could subsequently interact differently with each other depending on how the beat was perceived. Therefore, secondary PPIs were determined from the SMA and PMC bilaterally. The SMA bilaterally and left PMC showed increased coupling with the bilateral STG and right PMC during the Duration Beat condition compared to the Volume Beat condition, as shown in Figure 3, bottom.

Several of the increases in coupling between the cortical motor areas and the auditory cortex appear to be driven by the musicians in the study. The mean PPI coefficients for musicians and non-musicians of the coupling between the bilateral SMA and left PMC to the bilateral STG are shown in Figure 3, bottom. Musicians show significantly greater coupling than non-musicians in the Duration Beat condition than in the Volume Beat condition, whereas non-musicians tend to show more similar levels of coupling for both conditions (For exact PPI coordinates, see Supplementary Tables 6 and 7).
Discussion

The stimulus manipulations were successful in modulating the participants’ perception of the beat (Figure 1). High ratings of beat perception occurred even in the absence of external accents (Unaccented Beat condition), corroborating other work showing that internal subjective accents are generated when listening to unaccented isochronous rhythms (Temperley, 1963; Brochard et al., 2003). Moreover, the ratings indicate that these internal accents were as effective as external duration accents at inducing beat perception.

The putamen, pallidum, and caudate responded more to beat rhythms than nonbeat rhythms (Figure 2b). Only the basal ganglia (the putamen most robustly) responded to beat presence per se, showing activity increases for all beat conditions compared to nonbeat conditions. Interestingly, putamen activation did not parallel the behavioral beat ratings. Participants rated beat presence as Volume Beat > Unaccented Beat > Duration Beat, whereas putamen activity was Unaccented Beat > Duration Beat > Volume Beat. Why might this be the case? A critical difference between conditions is the requirement for internal beat generation - this is unnecessary in the Volume Beat condition, and essential in Unaccented Beat condition. Internal generation may modulate the basal ganglia response, with beat perception driving activity to a certain degree, but internal generation driving it as well. Importantly, the putamen response was not merely due to temporal complexity, as complexity was matched in the Unaccented and Volume Beat conditions. Also, it is unlikely that the response was influenced by potential differences in the perceived beat rate between conditions. Other studies have searched for correlations between basal ganglia activity and tapping or vocalization rates and found none (Jenkins et al., 1997; Riecker et al., 2003; Lutz et al., 2005; Riecker et al., 2006). Here also, basal ganglia activity did not significantly correlate with stimulus rate. However, as we did not specifically manipulate beat rate independently of stimulus rate, this remains to be fully tested.

Turning to other rhythm-responsive areas, greater activity was generally shown for the most unpredictable, temporally complex condition, the Volume Nonbeat condition, in which all tone onsets and accents are unpredictable (Figure 2c). In these regions, including dorsal premotor cortex, prefrontal cortex, inferior parietal lobule, and cerebellum, activity is associated with temporal complexity in motor tasks (Catalan et al., 1998; Lewis et al., 2004; Chen et al., 2008a). Their activation in a purely perceptual task suggests functional overlap between neural networks involved in the perception of temporal complexity and the organization and sequencing of temporally complex movements (Penhune et al., 1998).

Perhaps more surprisingly, several areas (right supramarginal and middle frontal gyri, bilateral insula) responded strongly to the ‘simplest’ condition—the Unaccented Beat condition. For this condition, some participants reported that, in addition to a regular beat, they perceived more complex patterns of accents, giving rise to a rhythmic pattern similar to the duration condition. Thus, even though the stimuli are quite simple, what subjects ‘do’ with their perception may not be so simple. The Volume Beat condition is also simple, but external accents force a particular interpretation of the sequence.

The consequences of the putamen’s role in beat perception are suggested by the analyses of connectivity. Beat perception led to increased cortico-subcortical coupling of the putamen with bilateral SMA and PMC regardless of how the beat was determined (Figure 3, top). One interpretation of this increased coupling is that the putamen encodes information about beat timing that facilitates cortical motor areas in precise control of movement timing, required, for example, when movements are made in time with beats.

In contrast to cortico-subcortical coupling, the cortico-cortical coupling amongst the SMA, PMC and auditory cortex did depend on accent type: greater coupling was observed for the
Duration Beat than the Volume Beat condition. Moreover, the increase in coupling depended on musical training. Only for musicians was coupling between bilateral SMA and STG significant (Figure 3, bottom). These differences in coupling between musicians and non-musicians are seen despite the lack of significant differences using traditional analyses of regional activation. This is consistent with other work showing that connectivity measures may be more sensitive than regional analyses (Rowe et al., 2007; Sonty et al., 2007). Recent work has investigated auditory-motor connectivity during tapping tasks in which participants to synchronize to rhythms (Chen et al., 2006; 2008b). During tapping, auditory-motor connectivity was increased as volume accents were increased, likely due to the greater influence of the auditory accent structure on participants' motor response (Chen et al., 2006).

In a subsequent study, auditory-motor connectivity increased with musical training (Chen et al., 2008a), although tapping accuracy and regional activation levels also differed. Here we show that during a purely perceptual task, and without differences in regional activity, auditory-motor connectivity can be altered by musical training.

The difference in coupling between musicians and non-musicians was paralleled by the difference in behavioral ratings: musicians rated the Duration Beat condition as having more of a beat than non-musicians did, whereas the two groups rated the Volume Beat condition similarly. Arguably, the Duration Beat condition is most similar to the type of rhythms musicians spend extensive time learning and performing, thus musicians' ability to organize ('chunk') and anticipate onsets may be superior to that of non-musicians (Smith, 1983). This ability may influence expectations about what will be heard later in the sequence (e.g., expecting onsets to coincide with predicted beats in the Duration Beat condition). In the brain, this could be mediated by top-down influence from motor areas to auditory cortex. In contrast, anticipation of future onsets in the Volume Beat condition is trivial for both musicians and non-musicians, thus no connectivity differences between the groups occur.

Whether musically trained or not, beat perception occurs spontaneously in most people without great effort. Cognitive theories of beat perception propose that the beat in music is indicated by several accent types: volume, duration, melodic, harmonic, timbral, etc. When attempting to find a beat in a sequence, people generate hypotheses about the beat location based on the perceived accents (Drake and Botte, 1993; Toivanen and Snyder, 2003; Hannon et al., 2004), and predict that future accented events will occur 'on the beat'. Successful prediction leads to enhanced processing of stimulus features (Jones et al., 2002). The basal ganglia have been implicated in prediction of events (Doya, 2000; Tanji, 2001; Schultz, 2006). When accents occur in unpredicted locations (not on the beat), then the listener's current predictions will be incorrect, causing a prediction error that leads to adjusted future hypotheses (Dickinson and Schultz, 2000; Davidson and Wolpert, 2003). This prediction error and continual updating of an internal model of events may explain the large degree of activation seen for the Volume Nonbeat conditions in both experiments: strongly accented events, which are normally indicative of a beat, were occurring at unpredictable times. The salient but unpredictable volume accents could cue participants to search for a beat to a greater degree than less salient duration accents.

We propose that the role of the basal ganglia in rhythm perception, as in other domains, is prediction: when a detectable structure is present in the rhythm, predictions can be made about the timing of future onsets. Successful predictions can enhance the speed of perceptual organization of the sequence, reducing working memory load. With EEG, increases in induced gamma-band activity are observed in anticipation of expected beat locations (Snyder and Large, 2005; Zanto et al., 2006). Taken together, these results suggest a strong relationship between anticipation or prediction and beat perception. These results also link to models of musical expectancy (for example, Large and Jones, 1999), suggesting that prediction may a key process in the perception of musical rhythm.
The results are also of clinical significance, as the basal ganglia are compromised in Parkinson's disease. Parkinson's disease patients have decreased striatal dopamine release, affecting excitatory input to the putamen (Lewis et al. 2003). Studies in patients with Parkinson's disease have shown deficits in timing tasks (Artieda et al., 1992; O'Boyle et al., 1996; Harrington et al., 1998). In addition, Parkinson's disease patients have selective deficits in discriminating rhythms that have a weakly-indicated beat structure (Grahn and Brett, 2009), such as Duration Beat rhythms, but they are unimpaired in discriminating control nonbeat rhythms. Rhythmic signals with a strong external beat ameliorate gait problems in Parkinson's disease and Huntington's disease (McIntosh et al., 1997; Thaut et al., 2001a; Thaut et al., 2001b). Thus, rhythmic cueing therapy may depend on common neural systems underlying rhythm perception and movement, including the putamen.

In conclusion, the basal ganglia show a specific response to the beat during rhythm perception, regardless of musical training or how the beat is indicated. We suggest that a cortico-subcortical network including the putamen, SMA, and PMC is engaged for the analysis of temporal sequences and prediction or generation of putative beats, especially under conditions that require internal generation of the beat. In these conditions, the coupling among cortical motor and auditory areas is facilitated for musically trained individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Medical Research Council (JAG, cost code U.1055.01.001.00001.01), the Wellcome Trust (JBR, WT077029), and the Betty Behrens Research Fellowship from Clare Hall, Cambridge (JAG).

References


---

J Neurosci. Author manuscript; available in PMC 2009 December 10.


*J Neurosci.* Author manuscript; available in PMC 2009 December 10.

Figure 1.
Schematic depictions of the auditory stimuli used in experiments one and two. On the left is a depiction of the auditory waveform, in the middle is a depiction in standard musical notation, and on the right are the mean ratings of beat presence given by participants for each condition (mus = musician, non = non-musicians). Red lines mark the onsets coinciding with the beat (applicable to beat conditions only), and correspond to the first note of each measure in the music. Volume accents are indicated by higher relative height (in left panel) or accent symbols (‘<’ in the right panel). The depictions represent an excerpt of approximately 2.5 seconds.
Figure 2.
SPM analyses. The Beat vs. Nonbeat contrasts overlaid on a template brain (a) and as extractions from regions of interest (b and c). Beat vs. Nonbeat SPM contrasts for experiment one and experiment two (for experiment two, activation is masked by same contrast from experiment one). Both experiments show significant bilateral activity in the putamen for this contrast. For experiment one, peak voxels are p < .05 whole-brain FDR corrected. For experiment two, peak voxels are p < .05 small volume corrected. Z refers to the level of the axial slice shown in stereotaxic MNI space. The graphs in parts b and c show mean activation (in arbitrary units) for the Volume Beat vs. Volume Nonbeat and Duration Beat vs. Duration Nonbeat contrasts (from experiment one), and Unaccented Beat vs Unaccented Nonbeat contrast (from experiment two). Part b) shows extractions from basal ganglia regions (as defined by Tzourio-Mazoyer et al., 2002) which were of a priori interest in this study. Part c) shows extractions from peak voxels in additional regions active in all rhythms vs. rest contrast in experiment one. * = p < .05 (significant difference between beat and nonbeat for that condition). SMG = supramarginal gyrus, MFG = middle frontal gyrus.
Figure 3.
Connectivity analyses in experiment one. Top panel displays regions showing increased coupling with the anterior putamen in Beat compared to Nonbeat conditions. Mean PPI coefficients (arbitrary units) for each of the target regions are shown in the graph. * = p < .05 SVC. Bottom panel displays regions with increased coupling during the Duration Beat condition versus the Volume Beat condition. Mean PPI coefficients (arbitrary units) from the target regions for each of the significant source → target pairs are shown in the graph (p < .05 SVC). Coefficients for musicians and non-musicians are shown: * = p < .05, significant difference between groups (independent samples t-test).
Table 1

Stereotaxic locations of peak voxels in Beat – No beat SPM contrast in experiments one and two.

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>Experiment one</th>
<th>Experiment two</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t score</td>
<td>x</td>
</tr>
<tr>
<td>L anterior putamen</td>
<td>4.87</td>
<td>−26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R anterior putamen</td>
<td>5.01</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L posterior putamen</td>
<td>5.2</td>
<td>−32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R posterior putamen</td>
<td>4.63</td>
<td>30</td>
</tr>
<tr>
<td>L caudate nucleus</td>
<td>4.29</td>
<td>−8</td>
</tr>
<tr>
<td>L ventral striatum</td>
<td>4.17</td>
<td>−18</td>
</tr>
<tr>
<td>R middle orbital gyrus</td>
<td>3.9</td>
<td>0</td>
</tr>
<tr>
<td>L inferior temporal gyrus</td>
<td>3.87</td>
<td>−50</td>
</tr>
<tr>
<td>L paracentral lobule</td>
<td>3.8</td>
<td>−18</td>
</tr>
</tbody>
</table>

This table shows the brain region, t values, and stereotaxic coordinates (in mm) of peak voxels for experiment one (p < .05 whole-brain FDR corrected) and experiment two (p < .05 SVC using mask generated from experiment one Beat – No beat contrast, thresholded at p < .05 FDR) in MNI space.