

Top-Down Control of Rhythm Perception Modulates Early Auditory Responses

John R. Iversen,^a Bruno H. Repp,^b and Aniruddh D. Patel^a

^a*The Neurosciences Institute, San Diego, California, USA*

^b*Haskins Laboratories, New Haven, Connecticut, USA*

Our perceptions are shaped by both extrinsic stimuli and intrinsic interpretation. The perceptual experience of a simple rhythm, for example, depends upon its metrical interpretation (where one hears the beat). Such interpretation can be altered at will, providing a model to study the interaction of endogenous and exogenous influences in the cognitive organization of perception. Using magnetoencephalography (MEG), we measured brain responses evoked by a repeating, rhythmically ambiguous phrase (two tones followed by a rest). In separate trials listeners were instructed to impose different metrical organizations on the rhythm by mentally placing the downbeat on either the first or the second tone. Since the stimulus was invariant, differences in brain activity between the two conditions should relate to endogenous metrical interpretation. Metrical interpretation influenced early evoked neural responses to tones, specifically in the upper beta range (20–30 Hz). Beta response was stronger (by 64% on average) when a tone was imagined to be the beat, compared to when it was not. A second experiment established that the beta increase closely resembles that due to physical accents, and thus may represent the genesis of a subjective accent. The results demonstrate endogenous modulation of early auditory responses, and suggest a unique role for the beta band in linking of endogenous and exogenous processing. Given the suggested role of beta in motor processing and long-range intracortical coordination, it is hypothesized that the motor system influences metrical interpretation of sound, even in the absence of overt movement.

Key words: rhythm; meter; beat; magnetoencephalography; MEG; music; beta; gamma; ERP; ERF; evoked; auditory; cortex; metrical interpretation; subjective accent

Introduction

Our perception of the world is shaped not only by extrinsic physical aspects of stimuli, but also by intrinsic interpretive processes. How do stimulus and brain interact to form perception? This question is commonly studied in the visual domain, using ambiguous stimuli with multiple perceptual interpretations (e.g., Necker cube, binocular rivalry, face/vase, etc.). Roles have been proposed for both bottom-up and top-down influences on what is perceived, but the

locations and means of interaction remain under active investigation.¹

In the current work, we examine this question in the context of auditory rhythm perception. In particular, we focus on the role of intrinsic interpretation in shaping the perceptual experience of rhythms, and the brain mechanisms that might mediate the influence of interpretation on perception. While rhythm refers simply to the organization of events in time, the perception of meter is an interpretation of the rhythm by the listener. Meter involves “hearing the beat” of a rhythm, the regular pulse that serves as a temporal anchor around which other events are organized. In much of the world’s music, meter is a fundamental aspect of how rhythms are organized

Address for correspondence: John R. Iversen, 10640 John Jay Hopkins Drive, San Diego, CA 92121. Voice: +858-626-2068; fax: +858-626-2099. iversen@nsi.edu

perceptually. In metrically perceived rhythms, the mental framework of regularly timed beats often involves multiple hierarchical levels of beat strength. The current research focuses on a basic aspect of metrical interpretation, namely, the location of the perceived beat.

In music, the location of the beat is often suggested by physical cues, such as accent, grouping, and phrasing or melody. However, the beat is ultimately a cognitive interpretation overlaid on top of the physical input. This can be seen in the ability to hear a beat even in syncopated rhythms, in which there is often no physical input at the time of the beat, as well as the phenomenon of subjective rhythmization, in which meter is perceived in sequences of physically identical sound events, perhaps due to the creation of subjective accents.²

Most importantly for the present study, the metrical interpretation of a rhythm can be modified at will: the location of the beat can be shifted voluntarily to different points within a rhythm.^{3–5} This is akin to a temporal version of mental rotation, and is an important ability in music perception and performance. Such metrical reinterpretation can have dramatic effects on the perceptual experience of a physically unchanged rhythm. We use this voluntary control of metrical interpretation as a model system for the study of the flexible cognitive organization of perception.

The focus of the present work is on human brain responses to auditory rhythms and on what these responses may reveal about the brain mechanisms underlying metrical interpretation in rhythm perception. The approach uses a single invariant rhythm that can be heard with a variety of metrical interpretations. The invariant stimulus is a simple auditory rhythm consisting of two tones and a rest (henceforth TT0 (Fig. 1A); see Refs. 3 and 6). The listener is instructed to hear the rhythm using one of two metrical interpretations: with the beat on either the first or second tone. The perceptual experience of the rhythm under the two interpretations is markedly different.

The tones of the rhythm evoke brain responses, which are interpreted as follows: If these responses reflect only the physical aspects of the stimulus, then they should not be affected by a listener's metrical interpretation. On the other hand, any differences in brain response with metrical interpretation can be attributed to mechanisms of metrical interpretation. Thus, we ask "How is the conscious interpretation of a rhythm reflected in brain activity?"

Specifically, the question is operationalized as asking how the evoked response to a sound is modulated by whether or not it is heard as the metrical beat. A second experiment asks whether the "imagined beat" effect on brain responses is similar to the effect of a physical stimulus accent, as would be expected if part of the mechanism of hearing the beat involved creating a subjective accent. Taken together, the two experiments examine the contribution of endogenous, voluntary interpretive factors and exogenous, physical factors on the rapid dynamics of brain processing of rhythm. As described below, we found that specific brain responses are selectively enhanced when a sound is heard as the beat, providing evidence for the modification of early auditory sensory processing by endogenous control. The control experiment suggests that this enhancement is related to the simulation of a subjective accent.

Brain Studies of Metrical Interpretation

A few prior studies have addressed the question of whether metrical interpretation is reflected in brain activity. We focus here on prior studies using EEG measures of brain response, which are most comparable to the current study because of their emphasis on short-timescale dynamics of brain response. Brochard and colleagues⁷ examined brain responses using a metronomic sequence that is typically perceived with subjective accents on every other tone.² They found that the brain response to occasional deviant tones was sensitive to whether

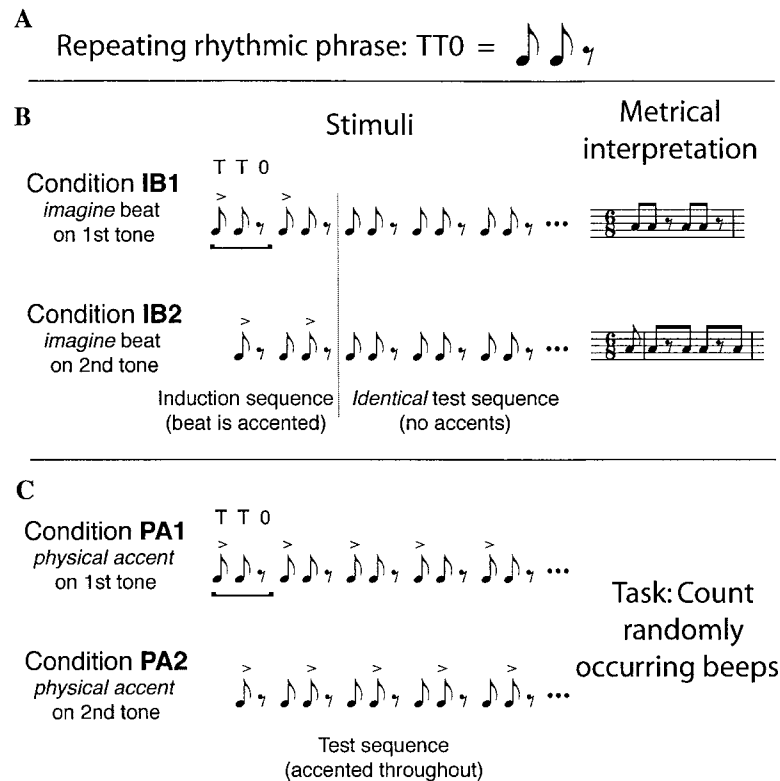


Figure 1. Description of stimuli. **(A)** The basic stimulus was a repeating series of two tones followed by a rest (TT0). **(B)** Metrical interpretation conditions studied in this experiment, in which the beat was heard (imagined) either on the first tone (imagined beat on 1 = IB1) or the second tone (IB2). Each sequence was preceded by a 10-beat induction sequence in which the desired beat was indicated by a physical accent. After the induction sequence, the stimulus was identical in the two conditions. Brain responses were measured only during the identical portions of the stimuli. **(C)** Control condition in which the beat was marked by a physical accent throughout, either on the first tone (physical accent on 1 = PA1) or the second tone (PA2). The task was not meter related, but involved counting the number of randomly occurring high-pitched beeps.

the deviance occurred on tones at hypothetically strong versus weak positions, based on predicted subjective accents. The most pronounced difference was in later parts of the evoked response (>300 ms post stimulus), and this was taken to reflect top-down modulation, perhaps of attention.

Snyder and Large⁸ measured brain responses to an isochronous sequence as well, but in this case every other tone was physically accented, which would be expected to give rise to a strong metrical interpretation with a beat on the accented tone. They found that the brain reacted differently to the occasional omis-

sion of strong versus weak tones. In particular, bursts of induced activity still occurred at the time of the omitted accented tone, providing evidence for the sensitivity of brain responses to metrical interpretation. This was taken to suggest a process of focal temporal expectancy that reflected the metrical interpretation of the rhythm, according to the dynamic attending theory of Jones, Large, and colleagues.^{9,10}

The preceding studies examined different aspects of brain responses to sound. Brochard and colleagues examined the traditional evoked response (ERP), which is calculated by averaging the raw brain responses aligned to the

tone onset.⁷ The ERP (and its magnetic counterpart, the evoked response field [ERF]) is dominated by relatively long timescale aspects of brain response, complicating the investigation of the fine temporal processing of successive events occurring in rapid succession. In contrast, Snyder and Large measured *oscillatory* activity, which differs from the traditional ERP/ERF in that it reflects higher frequencies of brain activity (in their case, 20–60 Hz).⁸ One advantage of such an approach is that oscillatory responses are temporally more focal, and thus responses to successive tones do not overlap as they do for the lower-frequency evoked responses.⁸ Oscillatory activity is often divided into multiple frequency ranges, such as beta (~15 to 30 Hz) and gamma (30 to 100 Hz), although there is not universal agreement on the exact boundaries or the functional significance of these ranges. Two ways of describing oscillatory responses are often used: evoked and induced responses. The evoked response reveals activity that is precisely time-locked to the stimulus onset, such as the ERF and evoked oscillatory responses. The induced response is a more inclusive measure that reflects changes in the power of oscillatory activity regardless of whether the oscillations are time-locked to the stimulus.

The present study was designed to extend past work in three ways. First, it attempts to reveal neural sensitivity to metrical structure explicitly by directly manipulating the location of the perceived beat of an invariant stimulus, as opposed to using occasional deviant stimuli to probe the brain's implicit reflection of a (presumed) metrical interpretation. The current design thus takes direct control of metrical interpretation by making it an experimental variable. Second, since the rhythmic sequence is not isochronous, behavioral verification of the metrical interpretation is straightforward, since the perception of the rhythm changes substantially under different interpretations. Third, combining the approaches of previous studies, the current work measures a wide spectrum of brain responses including ERF and oscillatory

responses, which are further subdivided into beta and gamma ranges.

Methods

Task and Stimulus

The basic stimulus consisted of a repeating sequence of two tones followed by a rest (symbolized TT0), and is similar to that used by Repp^{3,6} (Fig. 1A). The tones were 45-ms duration 1 kHz pips with inter-onset intervals of 200 ms (tempo: 100 bpm) and were delivered to participants through tubephones at a comfortable level, as determined prior to the experiment for each participant. This stimulus is the simplest one beyond an isochronous metronome, being formed by omitting every third event of a metronome (as omitting every second would yield another isochronous sequence). Yet this simple manipulation affords a significant increase in the variety of metrical percepts, as well as an important temporal reference (the gap) for the analysis of brain data, which enables us to be sure of where in the sequence the listeners are placing the beat.

Data were collected for two experimental tasks. The first (the “imagined beat” condition; Fig. 1B) tested whether voluntary metrical interpretation modulates brain responses. This examines the effect of endogenous processes in shaping brain responses. In different trials, listeners were instructed to adopt one of two metrical interpretations of the TT0 phrase by mentally placing the beat on either the first or second tone. These two conditions will be referred to as IB1 or IB2, for “imagined beat” on tone 1 or tone 2. Trials lasted 90 s (~150 repetitions of the TT0 phrase) and started with a 10-beat induction sequence, in which the tone that was intended to be heard as the beat was physically accented (2× amplitude). The induction sequence started on the beat (i.e., with tone 1 for IB1 and tone 2 for IB2). After the induction sequence, there was no further accenting,

tone onset.⁷ The ERP (and its magnetic counterpart, the evoked response field [ERF]) is dominated by relatively long timescale aspects of brain response, complicating the investigation of the fine temporal processing of successive events occurring in rapid succession. In contrast, Snyder and Large measured *oscillatory* activity, which differs from the traditional ERP/ERF in that it reflects higher frequencies of brain activity (in their case, 20–60 Hz).⁸ One advantage of such an approach is that oscillatory responses are temporally more focal, and thus responses to successive tones do not overlap as they do for the lower-frequency evoked responses.⁸ Oscillatory activity is often divided into multiple frequency ranges, such as beta (~15 to 30 Hz) and gamma (30 to 100 Hz), although there is not universal agreement on the exact boundaries or the functional significance of these ranges. Two ways of describing oscillatory responses are often used: evoked and induced responses. The evoked response reveals activity that is precisely time-locked to the stimulus onset, such as the ERF and evoked oscillatory responses. The induced response is a more inclusive measure that reflects changes in the power of oscillatory activity regardless of whether the oscillations are time-locked to the stimulus.

The present study was designed to extend past work in three ways. First, it attempts to reveal neural sensitivity to metrical structure explicitly by directly manipulating the location of the perceived beat of an invariant stimulus, as opposed to using occasional deviant stimuli to probe the brain's implicit reflection of a (presumed) metrical interpretation. The current design thus takes direct control of metrical interpretation by making it an experimental variable. Second, since the rhythmic sequence is not isochronous, behavioral verification of the metrical interpretation is straightforward, since the perception of the rhythm changes substantially under different interpretations. Third, combining the approaches of previous studies, the current work measures a wide spectrum of brain responses including ERF and oscillatory

responses, which are further subdivided into beta and gamma ranges.

Methods

Task and Stimulus

The basic stimulus consisted of a repeating sequence of two tones followed by a rest (symbolized TT0), and is similar to that used by Repp^{3,6} (Fig. 1A). The tones were 45-ms duration 1 kHz pips with inter-onset intervals of 200 ms (tempo: 100 bpm) and were delivered to participants through tubephones at a comfortable level, as determined prior to the experiment for each participant. This stimulus is the simplest one beyond an isochronous metronome, being formed by omitting every third event of a metronome (as omitting every second would yield another isochronous sequence). Yet this simple manipulation affords a significant increase in the variety of metrical percepts, as well as an important temporal reference (the gap) for the analysis of brain data, which enables us to be sure of where in the sequence the listeners are placing the beat.

Data were collected for two experimental tasks. The first (the “imagined beat” condition; Fig. 1B) tested whether voluntary metrical interpretation modulates brain responses. This examines the effect of endogenous processes in shaping brain responses. In different trials, listeners were instructed to adopt one of two metrical interpretations of the TT0 phrase by mentally placing the beat on either the first or second tone. These two conditions will be referred to as IB1 or IB2, for “imagined beat” on tone 1 or tone 2. Trials lasted 90 s (~150 repetitions of the TT0 phrase) and started with a 10-beat induction sequence, in which the tone that was intended to be heard as the beat was physically accented (2× amplitude). The induction sequence started on the beat (i.e., with tone 1 for IB1 and tone 2 for IB2). After the induction sequence, there was no further accenting,

and importantly, the stimuli were physically identical across the conditions. Brain responses were analyzed only for these periods of identical stimuli. Listeners completed four to five trials of each condition.

During the recording, listeners were instructed not to move or use motor imagery. They were, however, instructed to press and hold a button during any period they felt they had lost the intended metrical organization. Such periods were excluded from further analysis, but were infrequent (half the listeners never reported losing the beat). Those that did lose the beat lost it on average 10% of the time (somewhat more often while trying to imagine the beat on the first tone (IB1), which was generally reported to be harder).

In the second task (called the “physical accent” condition; Fig. 1C), the same stimuli were used, but with a physical accent (doubling of amplitude) on either the first or second tone that continued throughout the trial. (In effect, the induction sequences from the imagined beat conditions continued throughout the trial.) This tests the contribution of exogenous factors in shaping brain responses. These two physical accent conditions will be referred to as PA1 and PA2, for physical accent on tone 1 or tone 2. As the task was meant to assess the effect of physical accent, but not metrical interpretation, listeners were not told about the beat or meter, but were told only to pay attention to a stream of randomly and infrequently occurring high-pitched tone pips (2 kHz). Listeners were instructed to count the number of these high-pitched tones, which formed a separate perceptual stream from the TT0 rhythm. The TT0 rhythm was not mentioned in the instructions. This method was used in order to keep attention on the auditory modality, but not explicitly invoke metrical perception. This control task was always collected before the “imagined beat” task, in an attempt to assess the effect of physical accent in a way as uncontaminated as possible with metrical interpretation. In analysis, data around the high-pitched pips was not included (the TT0 sequence containing the

pip as well as the two following the pip were excluded).

Twenty-nine persons participated in the experiment. They all had some musical performance experience, as it was considered important that they be able to understand the concept of meter and of “hearing the beat” on a certain note within a rhythm, and also to be able to accurately report when they lost the beat. The degree of musical experience varied from 1 to 40 years (mean [SD]) = 14.1 [11.7] years). The first 12 participants completed only the “imagined beat” task, while the remaining 17 also completed the “physical accent” task (in a separate test session prior to the imagined beat task). In the initial group of 12 participants who did only the imagined beat task, the amplitude of the first tone was 10% higher than the second in an attempt to eliminate the expected perceptual bias of hearing the second tone as accented because of rhythmic grouping accents.¹¹ After the experiment, subjects reported being unaware of this amplitude difference. In the second group, this amplitude difference was eliminated from the IB stimulus and the two tones were identical.

Analysis of Brain Responses

During each trial, brain activity was measured continuously by magnetoencephalography (MEG) using a 148-magnetometer array (Magnes 2500WH; 4D Neuroimaging, San Diego, CA), with additional channels for eye blinks, saccades, and EKG. Data were collected at 508 Hz, after bandpass filtering from 1 to 100 Hz. The raw data was noise-reduced using a recording of room noise in order to minimize environmental contamination. Participants sat in a semi-reclined position within a magnetically shielded room and heard stimuli through tube-phones. Subject responses were collected through a button box pushed with their right index finger (all participants were right-handed).

Stimulus-evoked responses were computed by averaging MEG signals in epochs aligned to the first tone of the repeated TT0 phrase

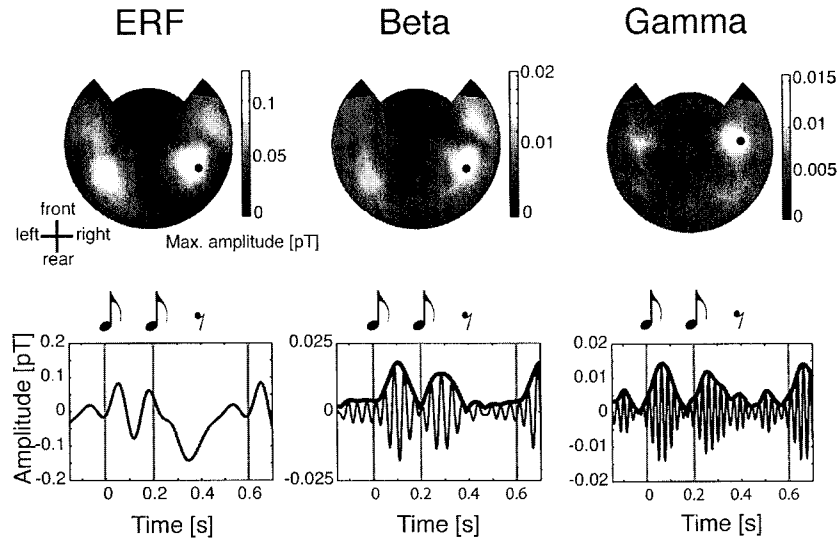


Figure 2. Example of auditory evoked responses. Stimulus-evoked response for a typical participant is shown in three frequency bands: ERF (1–10 Hz), beta (20–30 Hz), and gamma (30–50 Hz). Shown are responses to condition IB1 (imagined beat on first tone averaged over 640 epochs). Top row: Topographic map of maximum response amplitude over the head showing patterns typical of an auditory source. Lighter colors correspond to higher peak amplitudes. Bottom row: Stimulus-locked averages of the strongest channel (marked by a dot in the upper panels). Further analysis of beta and gamma responses use the amplitude envelope (thicker line). Note the phasic response to each tone in beta and gamma, in contrast to the overlapping responses of the ERF. In the evoked potential plots in this and following figures, thin gray vertical lines at 0 and 200 ms mark the onset of tones 1 and 2, respectively, and the vertical line at 600 marks the onset of tone 1 in the next tone pair. (In color in *Annals* online.)

after resampling the raw data to a sample rate of 500 Hz. Responses in several frequency bands were then found (using FIR forward–reverse filtering): ERF (1–10 Hz), beta (high: 20–30 Hz) and gamma (30–50 Hz). The justification for these specific bands goes beyond their conventionality: examination of wavelet spectra of single-subject brain responses typically revealed multiple peaks of activity centered around 25 and 40 Hz, suggesting two separate response components. Before averaging, epochs containing flux-jump artifacts or those in which the subject lost the beat were excluded. In the PA condition three epochs following the occurrence of each high-pitched tone were omitted. For each condition approximately 500 epochs were averaged per participant. Participants with fewer than 400 epochs were excluded, as well as those for

whom beta or gamma responses following the tones were no greater than background (defined as the mean amplitude in the 100 ms preceding the first tone of TT0), yielding 23 participants whose data were analyzed (10 in the “imagined beat”-only group and 13 who completed both the “imagined beat” and “physical accent” conditions).

Figure 2 shows examples of the evoked response in the three frequency bands analyzed: ERF, beta, and gamma for one participant (in the IB1 condition). The top row of plots shows the topography of maximal response amplitude across the sensor array. The response pattern is consistent with auditory sources of evoked neural activity. The bottom row shows the evoked waveform measured at the channel with highest amplitude (marked by a dot in the upper panels). Tone onsets are marked by vertical

lines. Visible in the ERF is the N100m response around 100 ms after each tone onset. Unlike the ERF, for both beta and gamma the response to each tone was clearly differentiated, and thus could be quantified separately. Analysis of these oscillatory responses used the amplitude envelope, computed by Hilbert transform (heavy line). To quantify the strength of response to each tone, the mean power (amplitude²) envelope of the 15 channels with highest power was computed. The mean power in the 200-ms window following each tone onset was then computed. The effect of the experimental manipulation was quantified as the proportional increase of response power when the tone was heard as the beat versus not the beat (or physically accented versus not accented). This is called the “imagined beat effect” (IBE). For example, for the first tone, the IBE is $[\text{Power}(\text{IB1}) - \text{Power}(\text{IB2})] / \text{Power}(\text{IB2})$, where $\text{Power}(\text{IB1})$ is the power of the response to tone 1 when it is imagined to be the beat, and $\text{Power}(\text{IB2})$ is the power of the response to tone 1 when it is not imagined to be the beat. Grand mean waveforms across participants were also computed. Prior to averaging across participants, the mean beta and gamma envelopes for each participant were separately normalized by scaling the envelopes so that the peak value across both conditions was one. As both conditions were scaled equally, this did not alter response relationships within an individual participant, but made comparisons between participants more equal.

Results

Effect of Metrical Interpretation on Brain Responses

In the first experiment, brain responses to the TT0 stimuli were measured as listeners heard the rhythm using one of two metrical interpretations that placed the imagined beat either on the first (IB1) or second tone (IB2). Figure 3 shows the across-participant ($n = 10$) grand mean responses for the ERF, beta, and gamma

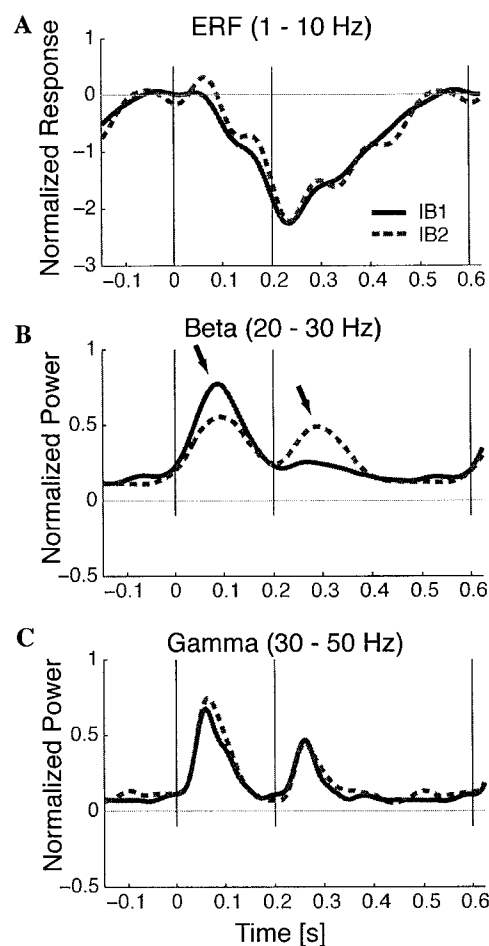


Figure 3. Across-participant grand means of normalized evoked responses for the two imagined beat conditions ($n = 10$; solid blue line: beat imagined on tone 1; dashed red line: beat imagined on tone 2). Grand averages are shown for three frequency bands: (A) ERF (1–10 Hz), (B) beta (20–30 Hz), and (C) gamma (30–50 Hz). For beta and gamma, the mean power envelopes were averaged across individuals after first normalizing each individual’s peak power across both conditions to one. The largest difference is in the beta response, where the response to both tones 1 and 2 is larger when that tone is imagined to be the beat (arrows). (In color in *Annals* online.)

frequency ranges for this experiment. IB1 responses are drawn with solid lines and IB2 responses with dashed lines. No significant difference is observed between the ERF responses as a function of metrical interpretation (Fig. 3A). The beta and gamma responses are

phasic, with a peak following each tone onset, but only the beta response shows a difference on account of metrical interpretation (Fig. 3B). There is a clear increase in evoked beta amplitude when a tone is imagined to be the beat (indicated by arrows) compared to when it was not the beat. Specifically, the response to the first tone is larger in the IB1 condition, in which it was imagined to be the beat (solid line), than in the IB2 condition, in which it was not the beat. Similarly, the response to the second tone was greater in the IB2 condition, in which it was heard as the beat (dashed line). In contrast, there is little difference in the gamma responses across conditions. Quantification of these effects is presented below.

Another interesting comparison concerns the response to imagined beat and nonbeat tones within a condition. In this case, in the IB1 condition (solid line) the imagined beat response to tone 1 is larger than the response to tone 2, which is not imagined to be the beat. In the IB2 condition, however, there is little difference between responses to tone 2 (imagined to be the beat) and tone 1. The simplest explanation for this asymmetry is a pronounced first-tone advantage in the beta response, namely, that the first tone of the TT0 has a larger beta response than the second (perhaps due to adaptation) and this effect is additive to the effect of imagining a beat. A second possibility is that the size of the difference in response between beat and nonbeat tones might be enhanced (perhaps also by suppression of the response to the non-beat tone) for the more difficult IB1 condition, as it was harder to maintain the beat on the first tone.

Figure 4 shows plots of the IBE on beta and gamma for all individuals included in the grand means of Figure 3. (The IBE is not computable for the ERF response because there were not separate peaks for tone 1 and tone 2.) The IBE is the fractional increase in response to a tone because of the fact of that tone's being heard as the beat within a metrical interpretation (compared to when it was not heard as the beat). Panels A and C are scatterplots

of the IBE for the first versus second tone for beta and gamma, respectively. Each dot represents a single listener. The reliability of the beta effect at the individual level is clearly visible (Fig. 4A), with all participants showing a positive IBE for at least one of the tones: all points lie in or near the upper right quadrant. The mean IBE for tone 1, tone 2, and both tones combined is greater than 50% (Fig. 4B). The mean IBE is significant in all three cases, although it is stronger for tone 2 (tone 1: 51%, $P = 0.037$; tone 2: 72% $P = 0.006$; combined: 61% $P < 0.001$, Wilcoxon signed-rank tests). The IBE for gamma-band responses (Fig. 4C and D) is smaller (mean $< 15\%$; note different scale), and not consistent across individuals, yielding no significant effect of imagining the beat on gamma responses (all $P > 0.19$).

A second group of participants, who completed the physical accent task (described below), also completed the imagined beat task in a later session. Their results demonstrated a similar group effect of the imagined beat in increasing beta response, but only on the first tone: The IBE for tone 1 was 45.2% ($P = 0.05$; compared with 51% for the initial group), but that for tone 2 was not significantly different from zero (-10.7% , $P = 0.17$). As with the initial group, there were no significant differences in the ERF or gamma-band responses (mean IBE for gamma was 8.5%, $P = 0.37$). There were two primary differences between this group and the first that might be relevant to the lack of an observed effect on tone 2: First, the degree of experience with the stimuli was very different, as participants in the second group had already heard the physically accented version of the stimuli continuously over a 2-h session before performing the imagined beat task. It is possible that this experience may have altered the task, perhaps making the imagining of a beat on tone 2 less demanding, which might therefore have reduced the need for auditory modulation. The second difference is the absence of the 10% amplitude accent on tone 1, which also would have made imagining a beat on tone 2 easier.

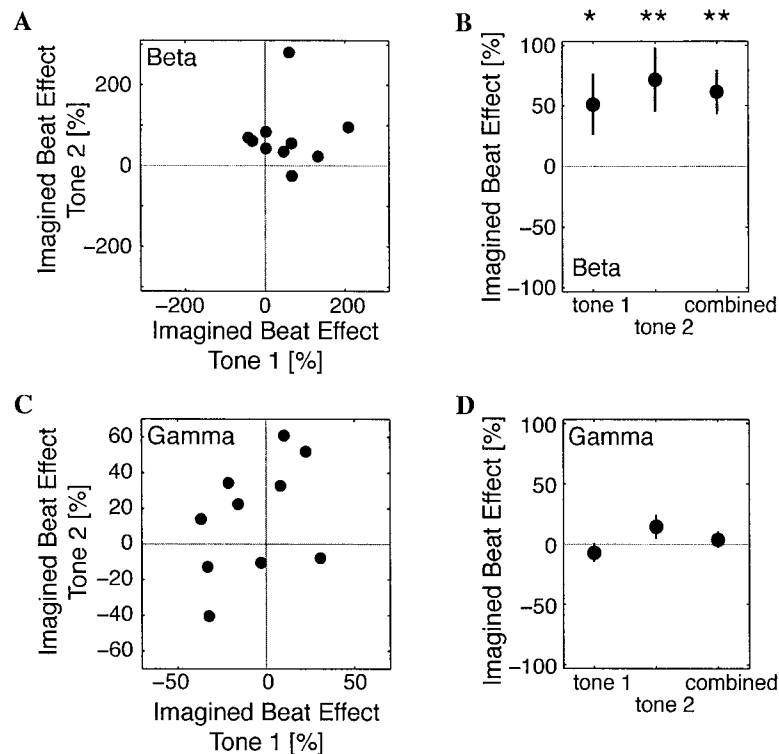


Figure 4. Quantification of the effect of imagining a beat on tone evoked responses in beta (**A**, **B**) and gamma (**C**, **D**) frequency ranges. The imagined beat effect (IBE) is the proportional change in response to a tone when it was imagined to be the beat versus when it was not. The left panels show the IBE for tones 1 and 2 for each of the participants ($n = 10$) underlying the means of Figure 3. For beta (**A**), the IBEs of all participants are in or near the top right quadrant, meaning that the beta response was consistently increased when that tone was imagined to be the beat. In contrast, gamma IBE is small and scattered (**C**). The right panels show means (bars: SE) of the effect of imagining a beat for tones 1 and 2 separately and both tones combined. Imagining the downbeat on a particular tone significantly increased the beta response to that tone, compared to when the same tone was not imagined to be the beat (**B**). In contrast, gamma responses were insensitive to imagining the beat (**D**). (Significance [Wilcoxon signed-rank]: * $P < 0.05$, ** $P < 0.01$.)

Effect of Physical Accents on Brain Responses

The second task examined the effect of physically accenting the first or second tone in the T10 pattern. This tested the hypothesis that an aspect of imagining a beat is the creation of a subjective accent, which might be mediated by enhancement of tone-evoked beta-band responses. Specifically, it was asked whether enhancements of auditory evoked beta caused by imagining a beat mimic those seen with actual stimulus accents. The conditions paralleled

those in the first experiment (see Methods), in that responses to the PA1 stimulus (accent on the first tone) are compared to those of the PA2 stimulus (accent on the second tone).

Figure 5 shows the grand mean ($n = 13$) evoked responses for the physical accent conditions in a format identical to that of Figure 3, but with the solid curves instead showing PA1 responses and the dashed curves showing PA2 responses. Three points should be noted. First, the enhancement of beta (Fig. 5B) by the physical accent is very similar to the enhancement of beta by imagining a beat (Fig. 3B). Second,

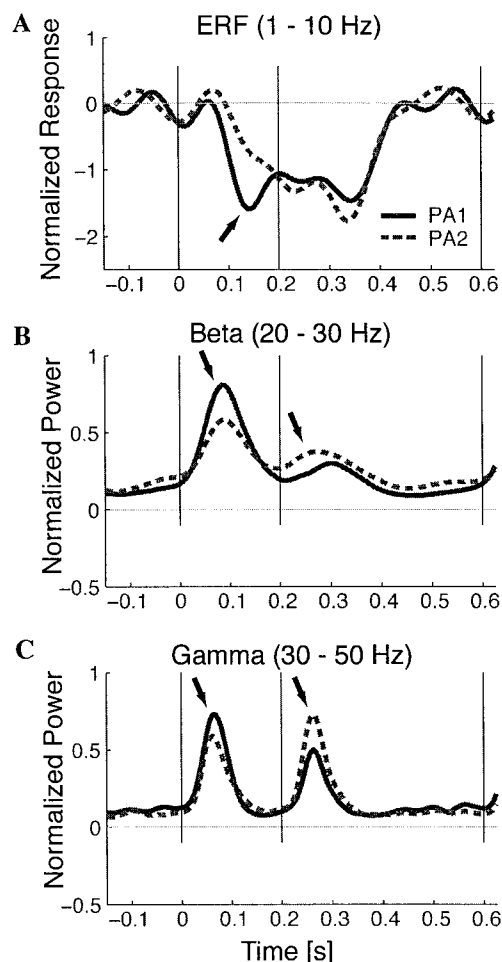


Figure 5. Grand mean responses for physically accented stimuli ($n = 13$). The format is the same as that of Figure 3, with solid blue lines showing response to the PA1 (accent on 1) condition, and the dashed red line showing the response to the PA2 (accent on 2) condition. Beta responses (**B**) are modulated by physical accents in a manner similar to the imagined beat (Fig. 3B), as indicated by arrows. In contrast to the imagined beat, physical accents also modulate the ERF (**A**) and gamma-band responses (**C**), also shown by arrows. (In color in *Annals* online.)

unlike the imagined beats, physical accents also affect the ERF and gamma-band responses, yielding stronger responses in these frequency ranges as well. Finally, like the imagined beat case, the beta response within a condition shows a first-tone advantage, again suggestive of adaptation of the beta response to successive tones.

Comparison of Effects of Metrical Interpretation and Physical Accents

The comparison of the effects of metrical interpretation and physical accent on brain responses is summarized in Figure 6, which shows the average enhancement of beta and gamma evoked response power in the imagined beat and physical-accent tasks, pooled across all participants (imagined beat: $n = 23$; physical accent, $n = 13$) and both tones. The average enhancement of beta by the internal manipulation of metrical interpretation (imagined beat) and the stimulus manipulation of accenting one tone (physical accent) is similar in magnitude and significance (imagined beat: 35%, $P = 0.006$, $df = 45$; physical accent: 46%, $P = 0.005$, $df = 25$; Wilcoxon signed-rank tests). The difference in the beta effect due to imagination versus physical accent is not significant. In the gamma range, the internal manipulation of beat has no significant effect (7%, $P = 0.33$), but physical accents do, with an effect size comparable to that seen for beta (40%, $P = 0.007$). Thus there is a dissociation of responses in the beta and gamma ranges. Interpretations of this finding are discussed below.

Discussion

Summary of Findings

The current study examined how the top-down interpretation of a rhythm is reflected in brain activity. It found that metrical interpretation of a simple rhythm has a large effect on sound-evoked brain responses, but only in the beta frequency range (20–30 Hz). By using a physically invariant rhythm for which listeners could voluntarily switch their metrical interpretation (i.e., the location of the beat), it was found that when a tone was imagined to be the beat, the evoked beta response power was increased by an average of 35%. This shows that the mechanisms involved in endogenous metrical

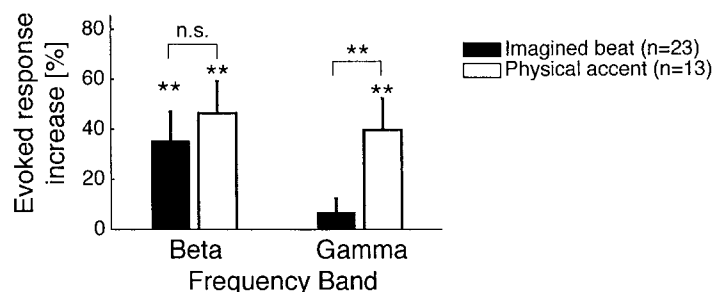


Figure 6. Summary of effects of imagined beat and physical accent: dissociation of beta- and gamma-band responses. Bars show the mean effect size (across all participants and across tones 1 and 2) of imagining the beat (solid) and physical accent (open) of increases in responses in beta and gamma frequency bands. Both imagined beats and physical accents positively modulate beta, whereas gamma is only modulated by physical accents. (Significance [Wilcoxon signed-rank test]: ** $P < 0.01$.)

interpretation have a strong modulatory effect on early auditory evoked responses. This modulatory effect is temporally precise, and specific in neural frequency: it was able to differentially modulate responses to tones only 200 ms apart, and only affected neural responses in the beta range (no significant changes were seen in the ERF (1–10 Hz) or gamma-band (30–50 Hz) responses.)

A second finding is that imagining a tone to be the beat increased beta in a similar way as if that tone had been physically accented. This suggests that the enhancement of beta is perceptually relevant, relating to the creation of subjective accents through a process that might neurally mimic aspects of responses seen to exogenous accents. A third finding is that while beta responses reflect both imagined beat and physical accent, ERF and gamma-band responses reflect only the physical accent.

Together, these results suggest a special role for brain activity in the beta range: evoked beta reflects both top-down and bottom-up processes, whereas the other evoked responses respond only to physical stimulus features. This suggests that neural activity in the beta range plays an important role in the interaction of endogenous and exogenous factors in shaping perception. Furthermore, it suggests that such interplay occurs early in the cortical auditory processing of sound.

What else is known about sound-evoked beta-band activity? While gamma-band responses to sound have been studied extensively,^{12,13} sound-evoked beta has received relatively less attention. Beta-band responses have been suggested to serve as a marker of stimulus novelty,^{14,15} and induced beta may indicate periodic expectation.⁸ Sound-evoked beta has also been suggested to modify the processing of subsequent sounds.¹⁶ The finding of beta-band involvement agrees with the findings of Snyder and Large.⁸ The highest power of both evoked and induced oscillatory activity, which they observed to be most strongly linked to metrical interpretation, was in the same 20–30 Hz range as the beta activity reported here. They found, as here, that evoked beta is sensitive to physical parameters of the stimulus, but it was not possible to separately examine the effect of endogenous meter on evoked beta because beats always coincided with physical accents. Their main finding concerned induced (nonphase-locked) beta, showing that it could anticipate tone onsets and occurred even when a tone was omitted.

Snyder and Large's⁸ finding that induced beta can anticipate tone onsets suggests a mechanism underlying the enhancement of evoked response to the beat observed in the present experiments: If an increase in induced beta coincides with the beat and precedes the tone onset, tones corresponding to the beat would arrive at

a time of increased beta activity, which might then yield a larger evoked response on account of a larger pool of responsive neurons. That is, the induced beta may set the stage for the modulation of evoked responses. This hypothesis will need to be tested using experiments designed to decouple the perceived beat from physical stimuli, preliminary results of which are presented below.

The fact that the effect of imagined beat on brain responses was observed in the beta range may have other implications for the mechanism of metrical interpretation, which are discussed below.

Mechanisms of Metrical Interpretation

Stepping back, the goal of this research is to understand the mechanism by which metrical interpretation is generated. How far have we progressed on this path? Repp³ lists several possible mechanisms by which metrical accents might be generated:

- (1) Dynamic modulation of attention, such that the location in time of the metrical beat receives greater attention, and thus becomes more salient (after Refs. 9 and 10).
- (2) Maintenance of an auditory image of a physical accent, such as the physical accents used in the induction sequence. This need not be an image of a previously heard sound, but could simply be created as a placeholder for the beat.
- (3) Establishment of a covert motor rhythm marking the beat, that is, the simulation of action without any physical output, which might have an auditory consequence.
- (4) A hybrid model including generation of a complete internal rhythm using simulated action, with imagined accents on the beat.

To what extent do the present results constrain the possibilities?

There are two parallel issues underlying these suggestions. The first concerns the *func-*

tion of the modulations of auditory response that we observed: Do they represent the modulation of attention, and/or the direct modification of the auditory image to create a subjective accent (which may then capture attention in subsequent processing)? The second issue concerns the anatomical *source* of the modulation of auditory responses, that is, whether it comes from frontal attention areas, from motor areas, or some other area. We treat these two issues below.

The current study showed what might be the ultimate consequence of an internal mechanism for generating metrical interpretation: the enhancement of evoked beta responses to tones coinciding with the beat. This is interesting because it suggests that voluntary top-down processes involved in metrical interpretation can act to modify early auditory responses. This, together with the observation that beta is enhanced by an imagined beat in much the same way as by a physical accent, is suggestive of an imagery-based explanation. That is, it is consistent with the idea that internal placement of the beat involves simulating a physical accent through the internal modulation of auditory responses.

However, an additional finding may temper the simple equation of beta response to perceptual consequences (e.g., stronger beta results in a louder perceived sound): the beta response shows a serial-order effect in addition to an effect of the imagined beat/physical accent. If beta and subjective accents were equated, it would predict a stronger accent on the first tone following a gap, which does not correspond to the standard finding that the tone preceding the gap should bear a grouping accent.^{6,11} There need not, however, be a single mechanism for subjective accenting, but grouping accent and metrical accent could use different mechanisms. In any event, the presence of such serial-order effects argues for the type of design used here, where the comparison is made between different metrical interpretations of the same physical tone, rather than between successive tones of the pattern.

An important question for future work is to explore the relation between the imagery-based model proposed above, and models proposing the modulation of attention or expectation. Distinguishing between these models is not simple because they need not be mutually exclusive, and if more than one mechanism is at play, the direction of causality still needs to be worked out. For example, could attention to a moment in time create a subjective accent at that point, or is it the creation of a subjective accent that draws attention to a moment in time?

A second key question for future work concerns the source of the modulation of auditory beta. For example, does this reflect the influence of an attentional network involved in generating dynamic pulses of attention?¹⁰ Or, is it input from the motor system that might influence auditory processing, perhaps as visual and somatosensory signals have recently been shown to do (by altering the underlying oscillatory phase to gate incoming signals)?^{17,18} The current results do not directly answer this question of source, or questions about the type of mechanism that generates the modulatory signal, be they neural oscillators or some other type of periodic pattern generation. The results do place temporal constraints on the modulation, showing that it must be temporally brief and precise, and be functionally able to modulate auditory responses.

A motor hypothesis is proposed below because of a variety of lines of evidence consistent with it, particularly regarding the role of beta-band responses, and the activation of motor areas by metrical rhythms. It must be noted that the current results are also consistent with an attention-based account as discussed above,¹⁰ although there are apparently fewer studies to date that have attempted to draw a link between explicitly attentional processes and evoked beta.^{8,14}

Motor Influence on Auditory Perception?

There are several reasons to suspect motor involvement in beat perception. The first, and

most obvious, is that a natural consequence of beat perception is rhythmic movement tied to the beat. The motor system is certainly able to produce regular outputs marking the beat at precise points in time, such as would be needed to selectively modulate auditory responses as observed in the present results. A second reason is that several studies have recently shown that the motor system, broadly speaking, is active in tasks involving rhythmic auditory perception, but no overt movement.^{19–21} Furthermore, certain parts of the motor system (basal ganglia and the supplementary motor area) are more active when listeners perceive highly metrical rhythms that induce a strong sense of beat.²⁰ These findings suggest the presence of an auditory–motor link that is active even in the absence of movement. The hypothesis proposed here takes an extra step in suggesting that the motor system may reciprocally affect the auditory system, possibly setting up precisely timed beat-related expectations or altering subjective perception (e.g., via subjective accentuation).

The fact that the effect of metrical interpretation was found in the beta frequency range is consistent with the motor–auditory coupling idea. Beta has been intimately associated with the motor system, in which it is involved in both overt movement and motor imagery.²² Beta has also been theoretically shown to be able to mediate longer-distance cortical coupling than gamma.²³ Consistent with this, beta has been suggested to play a role in the coupling of motor and sensorimotor networks during movement and motor imagery.^{24–27} More specifically, beta has been shown to be involved in tapping with a beat.^{27–30}

Several aspects of this motor–auditory coupling hypothesis remain to be addressed by future work. Conceptually, three aspects of metrical perception have been intertwined throughout this paper, and the literature: imagined beat (subjective accent), attention, and expectation. In theory, beat-related activity in the motor system could relate to all three. One approach to resolving this is to ask whether

the beta effect in the present work depends on there being a sound coincident with the beat, or whether beta responses can be observed for syncopated rhythms in which some beats are not marked by sounds. Prior work on responses to occasionally omitted tones has shown that induced beta is present at the time of an omitted tone,⁸ but it remains to be shown whether this marks the omission of an expected auditory event, or whether it is a marker of an internally generated beat or expectation. The use of syncopated rhythms, in which some beats are never marked by tones, and thus could not set up an expectation for a tone, could address this issue. A second test of the motor hypothesis would be to localize the source of sound-evoked beta, and to study intracortical interactions between auditory and motor areas (as done by Pollok *et al.*²⁸) or other networks, with an eye toward inferring causal relationships. Work along both fronts is in progress.

As an illustration of this work, Figure 7 shows preliminary results ($n = 7$) of a study of beat perception in syncopated rhythms. In the experiment, an identical rhythm is presented in one of three conditions (Fig. 7A), with the endogenously generated perceived beat either coinciding with a note (IB0) or with the perceived beat occurring before (IB+) or after (IB-) the note, yielding a syncopated rhythm. Figure 7B shows evoked beta responses to these three conditions. In all cases the evoked response follows the tone. The result is consistent with the current study, as the strength of the evoked response is greater when the tone is heard as the beat (IB0 condition). Figure 7C shows induced beta responses in the three conditions. Notably, the peak of induced beta coincides with (and anticipates) the time of the imagined beat (marked by vertical dashed lines 150 ms before and after the beat), not the tone. This finding is consistent with the results of Snyder and Large,⁸ and further suggests that modulation of induced beta represents beat-related processes that may be distinct from auditory expectation, since tones never occur on the analyzed beats. A corollary of these two findings

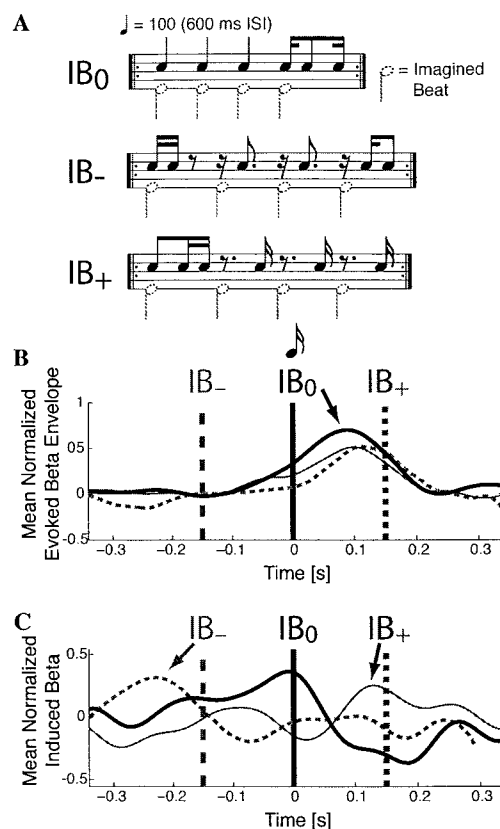


Figure 7. An experiment using syncopated rhythms. **(A)** Stimuli used in the study were the same four-beat rhythm, but with the imagined beat (dotted notes) falling at different points in the rhythm. In condition IB0, the beat always coincides with a note. In conditions IB- and IB+, the first beat is always marked by a note, but the next three beats occur at times where there is never a note. In IB- the beat occurs prior to a note, while in IB+ the beat occurs after a note. All notes were staccato tone pips as in earlier experiments, with the same brief (45 ms) duration. Brain responses surrounding the three isolated notes were computed by averaging using the same methods described in the Methods section. **(B)** Grand mean ($n = 7$) normalized evoked response envelope in the beta range (17-24 Hz) for B0 (solid line), B- (heavy dashed red line), and B+ (thin blue line). Responses are aligned to tone onset ($t = 0$), with the mean base line from -200 to -100 ms subtracted. In all conditions, the evoked response follows the note. It is largest for condition B0, when the beat and note coincide (arrow). **(C)** Grand mean normalized fluctuation of induced beta shows a peak of power that reflects the timing of the imagined beat, not the auditory input (arrows). Beta fluctuation was computed by subtracting the mean over the entire interval (-300 to 300 ms). (In color in *Annals* online.)

is that there is a positive correlation between the strength of induced beta at the time of tone onset and the size of the evoked beta response. This result supports the hypothesis proposed above that the process responsible for increasing induced beta may facilitate sound-evoked responses, possibly providing a mechanism for the creation of subjective accents.

In summary, an influence of endogenous perceptual organization of auditory sequences was revealed in beta-band responses to sound. The precisely timed modulation of beta is consistent with a role mediating subjective accents, and is suggestive of motor–auditory interactions playing a role in the perceptual organization of rhythmic sound.

Acknowledgments

We thank Lacey Kurelowich, Jovan Stewart, Heidi Moomaw, and Alex Leopardi for their help with the experiments. This work was supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute, where J.R.I. is the Karp Foundation Fellow and A.D.P. is the Esther J. Burnham Fellow.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Kornmeier, J., C.M. Hein & M. Bach. 2008. Multistable perception: when bottom-up and top-down coincide. *Brain Cognit.* **69**: 138–147.
- Bolton, T.L. 1894. Rhythm. *Am. J. Psychol.* **6**: 145–238.
- Repp, B.H. 2005. Rate limits of on-beat and off-beat tapping with simple auditory rhythms. 2. The roles of different kinds of accent. *Music Percept.* **23**: 165–188.
- Repp, B.H. 2007. Hearing a melody in different ways: multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *Cognition* **102**: 434–454.
- Repp, B.H., J.R. Iversen & A.D. Patel. 2008. Tracking an imposed beat within a metrical grid. *Music Percept.* **26**: 1–18.
- Repp, B.H. 2005. Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 1. Qualitative observations. *Music Percept.* **22**: 479–496.
- Brochard, R., D. Abecasis, D. Poter, *et al.* 2003. The “ticktock” of our internal clock: direct brain evidence of subjective accents in isochronous sequences. *Psychol. Sci.* **14**: 362–366.
- Snyder, J.S. & E.W. Large. 2005. Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cogn. Brain Res.* **24**: 117–126.
- Jones, M.R. 1976. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychol. Rev.* **83**: 323–355.
- Large, E.W. & M.R. Jones. 1999. The dynamics of attending: how we track time varying events. *Psychol. Rev.* **106**: 119–159.
- Povel, D.-J. & H. Okkerman. 1981. Accents in equitone sequences. *Percept. Psychophys.* **30**: 565–572.
- Pantev, C., S. Makeig, M. Hoke, *et al.* 1991. Human auditory evoked gamma-band magnetic fields. *PNAS* **88**: 8996–9000.
- Ross, B., T.W. Picton & C. Pantev. 2002. Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res.* **165**: 68–84.
- Haenschel, D., T. Baldeweg, R.J. Croft, *et al.* 2000. Gamma and beta frequency oscillations to novel auditory stimuli: a comparison of human electroencephalogram (EEG) data with in vitro models. *PNAS* **97**: 7645–7650.
- Kisley, M.A. & Z.M. Cornwell. 2006. Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clin. Neurophys.* **117**: 2549–2563.
- Hong, L.E., R.W. Buchanan, G.K. Thaker, *et al.* 2008. Beta (~16 Hz) frequency neural oscillations mediate auditory sensory gating in humans. *Psychophysiology* **45**: 197–204.
- Lakatos, P., G. Karmos, A.D. Mehta, *et al.* 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* **320**: 110–113.
- Schroeder, C.E., P. Lakatos, Y. Kajikawa, *et al.* 2008. Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* **12**: 106–113.
- Chen, J.L., V.B. Penhune, R.J. Zatorre. 2008. Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* **18**: 2844–2854.
- Grahn, J.A. & M. Brett. 2007. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**: 893–906.
- Lahav, A., E. Saltzman, G. Schlaug. 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* **27**: 308–314.

22. Schnitzler, A., S. Salenius, R. Salmelin, *et al.* 1997. Involvement of primary motor cortex in motor imagery: a neuromagnetic study. *NeuroImage* **6**: 201–218.
23. Kopell, N., G.B. Ermentrout, M.A. Whittington & R.D. Traub. 2000. Gamma rhythms and beta rhythms have different synchronization properties. *PNAS* **97**: 1867–1872.
24. Stancak, A. & G. Pfurtscheller. 1996. Event-related desynchronization of central beta-rhythms during brisk and slow self-paced finger movements of dominant and non-dominant hand. *Brain Res. Cogn. Brain Res.* **4**: 171–183.
25. McFarland, D.J., L.A. Miner, T.M. Vaughan & J.R. Wolpaw. 2000. Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topogr.* **12**: 177–186.
26. Brovelli, A., M. Ding, A. Ledberg, *et al.* 2004. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *PNAS* **26**: 9849–9854.
27. Mayville, J.M., A. Fuchs, M. Ding, *et al.* 2001. Event-related changes in neuromagnetic activity associated with syncopation and synchronization timing tasks. *Hum. Brain Map.* **14**: 65–80.
28. Pollok, B., J. Gross, K. Mueller, *et al.* 2005. The cerebral oscillatory network associated with auditorily paced finger movements. *NeuroImage* **24**: 646–655.
29. Boonstra, T.W., A. Daffertshofer, C.E. Peper & P.J. Beek. 2006. Amplitude and phase dynamics associated with acoustically paced finger tapping. *Brain. Res.* **1109**: 60–69.
30. Thaut, M.H. 2003. Neural basis of rhythmic timing networks in the human brain. *Ann. N. Y. Acad. Sci.* **999**: 364–373.