The Neuroscience of Groove: Neural Mechanisms Marrying Music and Movement

Gabriel A. Nespoli

Ryerson University

# Abstract

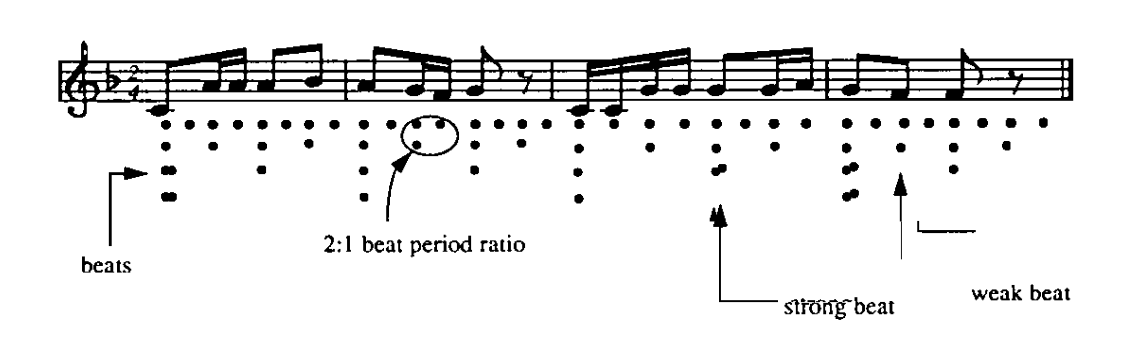
Music has a long history of being associated with movement synchronization such as foot-tapping or dance. These behaviours are easier with some music compared to others, and the reasons for this are not well understood. “Groove” is a quality of music that compels synchronous movement in the listener, and certain acoustic and musical features have been identified that contribute to a sense of groove. Neurons have been found to entrain to the beat of rhythms and music as well. Combining these two ideas, it could be predicted that neural populations related to movement (i.e., motor or premotor areas) would entrain more to high-groove versus low-groove music. The current proposal seeks to explore some acoustic and musical features of music that contribute to neural entrainment in motor areas of the brain. Study 1 investigates the effects of groove on premotor entrainment, using stimuli that have been rated on extent of groove in a previous study. Study 2 will extract acoustic features from those stimuli that have previously been found to be related to groove. Using a regression model, we will try to predict neural entrainment and behavioural synchronization ability on the basis of these acoustic features. Study 3 will investigate the musical feature of syncopation—which has previously been found to be associated with sense of groove—on extent of premotor entrainment and behavioural synchronization ability.

# The Neuroscience of Groove: Neural Mechanisms Marrying Music and Movement

What is it about music that makes us want to move? Why is it sometimes difficult to resist the urge to move when you hear a good beat? Whether it is bobbing your head, walking with a bounce in your step, or full-out dancing, synchronizing movement with music is often effortless and automatic. The quality of music to elicit a desire to move is referred to as “groove” [@Madison2006], and several studies have investigated features of music that contribute to groove [@Madison2011; @Davies2013; @Sioros2014; @Witek2014; @Stupacher2016] as well as its behavioural [@Janata2011] and neural [@Stupacher2013] consequences.

TODO explain meter with an example

Music comprises notes that are temporally arranged into various patterns of duration called *rhythmic groups* [@London2012]. Based on the time differences between the onsets of each note in a rhythm it is possible to abstract away an underlying isochronous[[1]](#footnote-1) pulse or beat. *Meter* is a psychological construct that organizes auditory events in time by aligning them onto a grid, whose main subdivision is the beat. These subdivisions can be conceptualized as nested levels of a hierarchy [@Lerdahl1983]. Each level consists of a set of equally spaced beats, and levels are related to one another by simple integer ratios. Points in time where many beat levels align are considered points of metrical accent, and tend to give rise to a single level that is considered to be the main beat (see Figure 1). When you tap your foot along with a piece of music you are usually synchronizing your taps with this main beat level.



*Figure 1.* Meter and Rhythm. The temporal location of notes can be organized into a metrical heirarchy [@Large1994]. Each row represents a beat level, and points in time where many beat levels align are points of metrical accent.

Different levels of the metric hierarchy are usually related to one another by factors of 2 or 3 (or 1/2 or 1/3), which correspond to *duple* and *triple* meter, respectively.

So, whereas *rhythm* concerns the physical location in time of musical events, *meter* concerns the psychological creation of an underlying structure with which we can attempt to understand and predict the rhythm as it unfolds. Meter can be referred to as a “mode of attending,” because each successive note will be perceived in the context of the meter which is currently perceived. This creates a set of rhythmic expectancies where notes are more likely to fall onto subdivisions of the grid than in between. The interaction between the sensation of rhythm and the expectancies of meter are central aspect of the perception of music.

## Behavioural Synchronization to Music

Humans are good at synchronizing their finger tapping with metronomes and music [@Repp2005; @Repp2013], and tapping along has been found to improve the perception of timing perturbations [@Manning2013; @Manning2015]. During full-body synchronization to music, motion capture studies have found that the movement of certain body parts will synchronize with different metrical levels ….

*Groove* is a high-level musical feature referring to music that compels synchronous movement [@Madison2006; @Janata2011; @Witek2014]. Madison [-@Madison2006] defined groove as “wanting to move some part of the body in relation to some aspect of the sound pattern” (p. 201). When asked to rate various statements about the definition of groove, Janata et al. [-@Janata2011] found that participants consistently endorsed statements containing words emphasizing movement and rhythm (e.g., move, dance, beat, rhythm) or compulsion (e.g., feel, make, want). For example, all[[2]](#footnote-2) participants agreed with the statement that “the groove depends on the extent to which the music makes you want to move” (p. 4).

Groove has been found to support behavioural synchronization, such as the ability to accurately tap along to the beat of the music. Janata et al. [-@Janata2011] created a corpus of 148 musical excerpts (20 seconds in duration) and had participants rate each excerpt on “the extent to which the music ‘grooves’” (p. 5). Based on these groove ratings they selected a subset of excerpts representing high-, mid-, and low-groove categories, and asked a separate group of participants to tap along with each excerpt. They found that when participants’ were asked to tap along to the beat (i.e., isochronous tapping), their taps were better synchronized with the beat for high-groove excerpts. Additionally, when participants were given freedom to tap as they liked (i.e., free-form tapping), the periodicities present in their tapping was also better synchronized with the periodicities present in the excerpt itself. This suggests that groove supports a person’s ability to move along in synchrony with the music (i.e., the beat and meter of the music).

Janata et al. [-@Janata2011] also video recorded participants during a no-tapping condition, where they were instructed to simply listen to the music. They coded these videos for spontaneous movements of different body parts (i.e., head, trunk, foot, hand), and found that there was more spontaneous movement while listening to high-groove music. This strengthens the notion that groove *compels* movement; it is able to create a desire in the listener to move along with the music.

## Neural Entrainment to Music

Music has been found to the ability to entrain the firing patterns of neurons. Many studies have found that neurons will entrain their firing to both isochronous [@Nozaradan2011; @Snyder2005; @Fujioka2012; @Arnal2015] and non-isochronous [@Nozaradan2012; @Nozaradan2013;] rhythms. These studies have used various methods including electro- and magnetoencephalography (EEG and MEG), transcranial magnetic stimulation (TMS), and functional magnetic resonance imaging (fMRI). Features of the E/MEG that have been investigated include the steady-state evoked potential (SSEP), broad-band spectral activity in the delta, beta, and gamma frequency ranges, and inter-trial phase coherence. These methods and results are presented in the following sections.

### Steady-state evoked potential (SSEP)

Nozaradan et al. [-@Nozaradan2011] has published a series of articles using a novel method for studying spectral activity in the EEG, called the SSEP [see @Nozaradan2014 for a review]. To calculate the SSEP, multiple trials are averaged together in the time domain in order to discard non-phase-locked activity. A fast-Fourier transformation (FFT) is performed on each channel of EEG data, and then the spectra across all channels are averaged together. The resulting averaged spectrum is then inspected at frequencies of interest in order to determine the extent of entrainment.

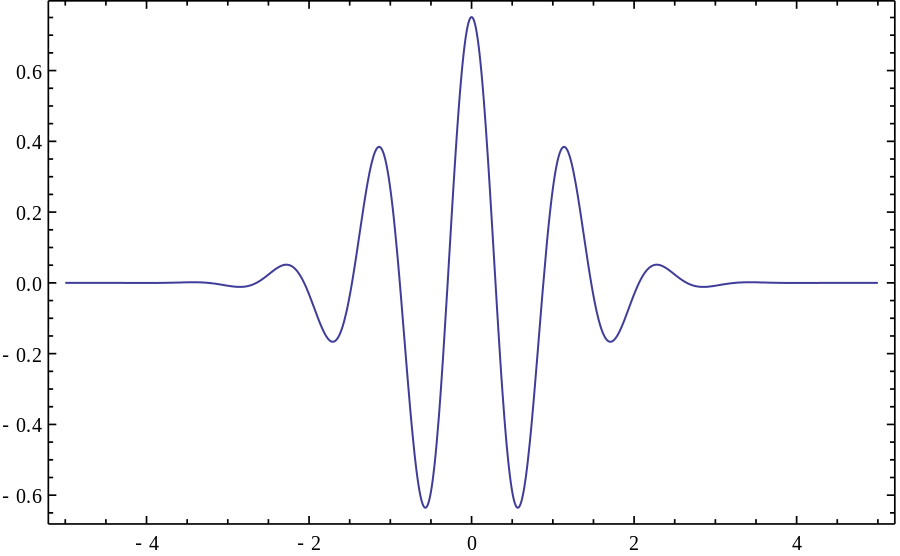
In a first study Nozaradan et al. [-@Nozaradan2011] had participants listen to a pure tone that had been amplitude modulated to have a steady pulse of 2.4 Hz. Looking at the EEG, there was a significant peak in the SSEP at 2.4 Hz, suggesting that neural activity was being entrained by the stimulus. Participants were also instructed to imagine duple or triple meter (i.e., by perceptually grouping the beats in twos or threes). This resulted in additional peaks at frequencies associated with these meters, suggesting that the SSEP does not capture entrainment that is solely stimulus-induced. When participants biased their perception of the stimulus, this was reflected in the EEG.

The SSEP is also sensitive to meter. When participants listened to rhythms which contained more than one periodicity, the SSEP reflected this with peaks at beat- and meter-related frequencies [@Nozaradan2012]. It is important to note that the frequencies being considered here are not necessarily those that are most prominent in the physical stimulus, but rather more “musical” ones that would be used for finding the beat and tapping your foot. Interestingly, SSEP activity at these frequencies was selectively enhanced, irrespective of the spectrum of the stimulus. Thus, there does not seem to be a one-to-one mapping between the spectrum of the stimulus and the SSEP, indicating that a sense of beat and meter are (at least partly) endogenously generated.

The SSEP has also been found to track the beat and meter in natural music, and this tracking is disrupted in the presence of contradictory rhythmic cues [@Tierney2014].

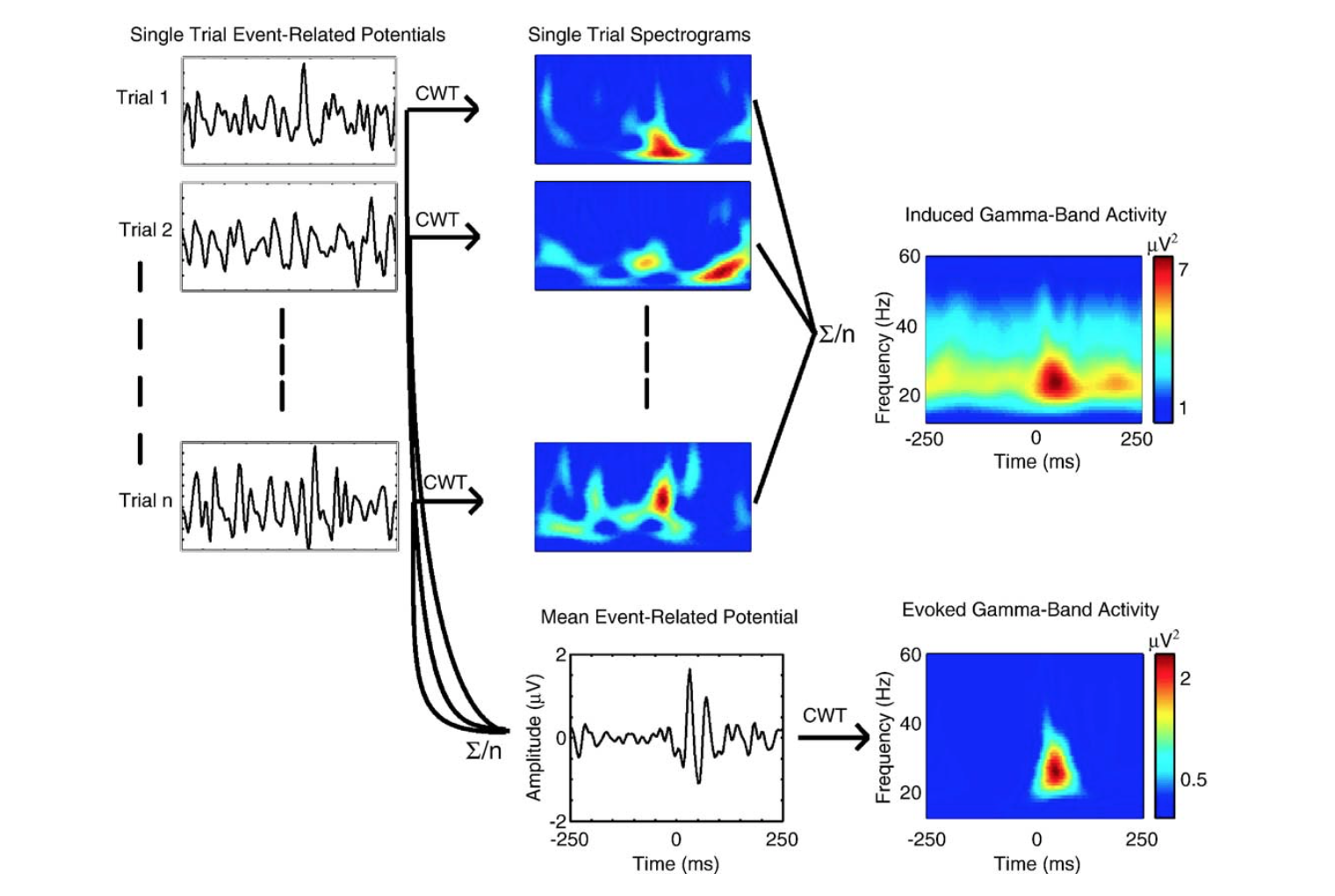
### Frequency bands

EEG (and MEG) activity has long been categorized into different frequency bands, defined loosely as follows: delta (< 4 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (> 30 Hz). In studies on dynamic attention and beat perception, the bands that have received the most attention are beta [@Fujioka2012] and gamma [@Snyder2005]. Activity in different bands is commonly extracted using a continuous wavelet transformation (CWT). A wavelet is a short waveform that begins and ends at the zero crossing, and is defined a priori. For example, the complex-valued Morlet wavelet (Figure 2) is commonly used in research on human perception. Wavelets can be parametrically adjusted to have different frequencies (i.e., the distance between successive peaks) or number of cycles (i.e., the number of peaks). Then, the wavelet is fit to the data at each time point and at each frequency scale. The number of cycles in the wavelet is usually scaled with frequency, such that the wavelet fit to lower frequencies has fewer cycles than the one used for higher frequencies.

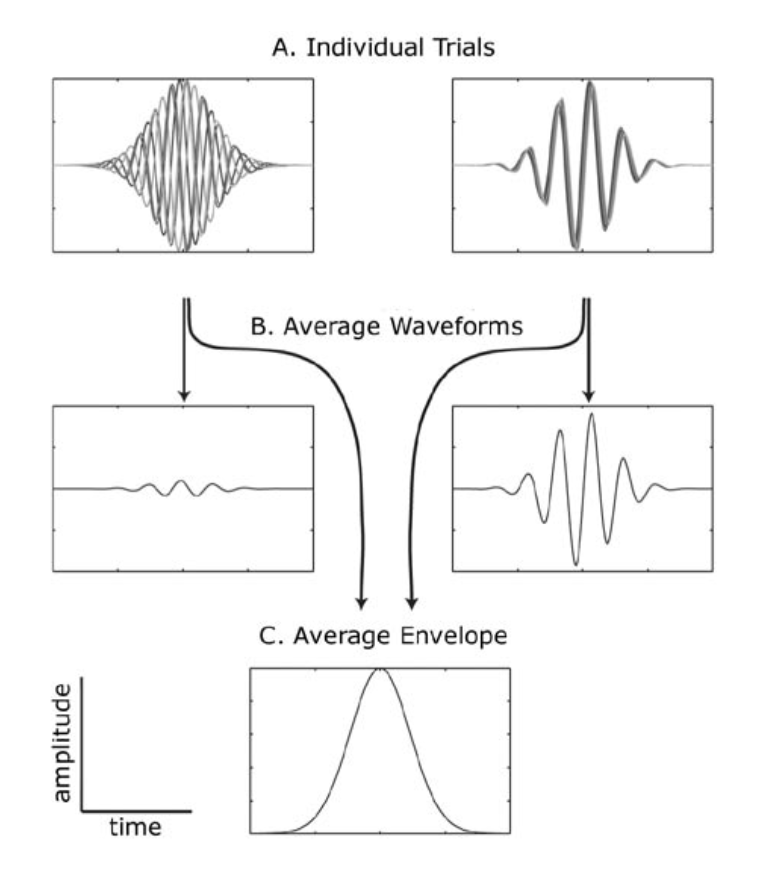


*Figure 2.* The Morlet wavelet.

Snyder and Large [-@Snyder2005] investigated the entrainment of gamma-band activity to isochronous sequences, and calculate two different kinds of activity using a CWT, called evoked and induced activity. The difference between evoked and induced activity concerns the order in which a) trials are averaged together, and b) the CWT is performed (see Figure 3). For evoked activity, trials are averaged together (i.e., in the time domain) and then a CWT is performed on the resulting waveform. For induced activity, the CWT is performed on each trial, and the resulting spectra are averaged together. These processes are done for each channel of the EEG before averaging across channels. The main difference is that evoked activity only captures phase-locked activity. For example, many trials could consist of the same amount of energy at certain frequency, but if the phase of this oscillation is not time-locked between trials, then a time-domain average will attenuate these oscillations. In this case, the evoked calculation will not show much power at that frequency, whereas the induced calculation will (Figure 4).



*Figure 3.* Evoked versus induced activity [@Snyder2005]. The evoked calculation captures only phase-locked activity by first averaging trials in the time domain before computing the spectrum. For induced activity, the spectrum is calculated separately for each trial before averaging.



*Figure 4.* Phase locked versus non-phase-locked activity [@Zanto2006]. Given several trials of non-phase-locked (panel A, left) and phase-locked (panel A, right) activity, the average waveforms (panel B) will differ significantly. In contrast, the average envelope (panel C) does not.

For example, gamma-band (20-60 Hz) activity, calculated over fronto-central electrodes, has been found to synchronize with an isochronous sequence of loud and soft tones [@Snyder2005]. Additionally, non-phase-locked gamma-band peaks have been found to persist during tone omissions whereas phase-locked peaks arrived late [@Zanto2006], suggesting that the non-phase-locked activity is predictive of upcoming tones.

Using MEG, the modulation of beta-band (13-25 Hz) activity has been found to synchronize with isochronous sequences at various rates [@Fujioka2012]. Specifically, there is a consistent decrease in beta-band power after each tone, reaching a minimum at 200 ms post-stimulus onset. This decrease is followed by an increase in power that peaks concurrently with the onset of the next tone, regardless of the repetition rate. Beta-band activity is possibly of particular interest to the current dissertation due to its association with movement [@Engel2010].

### TMS and Motor Cortex Excitability

Transcranial magnetic stimulation (TMS) has been used to investigate neural response to rhythm. Using TMS, the primary motor cortex can be rhythmically stimulated with an isochronous pulse while the participant is listening to music. Electromyography (EMG) can then be used to measure peripheral muscle activation corresponding to the region of stimulation in the motor cortex. Stupacher et al. [-@Stupacher2013] found that when the TMS pulses are presented on the beat of the music, the resulting EMG activity is greater than when the pulses are presented off the beat. This finding suggests that motor cortex is entraining to the beat of the music and is more responsive to TMS pulses at those locations in time. During on-beat TMS pulses, neurons are already firing at this time and so there is a large increase in EMG activity seen at the distal muscle. By contrast, during off-beat TMS pulses neurons are [mostly] not firing, and so the increase in EMG activity is not as large.

### Source Localization

What parts of the brain are involved in neural entrainment to the beat of music or other auditory stimuli? Auditory regions have been implicated, but also many other regions including motor, premotor, basal ganglia, and visual regions. This is potentially surprising that so many non-auditory regions would be attuned to an auditory stimulus. One potential explanation is that if one sensory modality detects some regularity in the world, other brain regions would be interested in this so that they can interact with the world appropriately. For example, motor cortex would need to know the temporal location of beats in order to synchronize one’s movements with it.

Grahn and Brett [-@Grahn2007] used fMRI to study the brain regions involved in the perception of a beat in musical rhythm. They measured participants’ ability to tap along with different rhythms, and found that performance was better for rhythms that had a stronger sense of meter. These same rhythms elicited a greater BOLD response in the basal ganglia and supplementary motor area (SMA). Grahn and Rowe [-@Grahn2009] found that the presence of a beat was associated with greater connectivity between basal ganglia, premotor areas, and auditory cortex. While these studies were not able to investigate entrainment of neural populations per se, it seems that the perception of a beat recruits areas of the brain related to movement.

## Neural Mechanisms

*Neural resonance theory* [@Large2009] offers a potential neural mechanism for the ability of neurons to synchronize with the explicit and implied meter of auditory stimuli. The theory posits that neurons act as frequency-tuned oscillators that will automatically synchronize their firing with the incoming sensory stream. Different neurons will entrain to different metrical levels, creating a neural representation of the different patterns of oscillatory activity in the stimulus. Critical for the theory is that neurons are connected in multiple hierarchical levels, where the firing of higher levels will automatically entrain to the firing patterns in lower levels. At points of metrical accent, many neurons will be firing at the same time, and will give rise to stronger synchronization at higher levels. In this way, a sense of meter can emerge as a natural consequence of the periodicities present in the stimulus.

This representation of meter in higher-level populations of neurons can be considered a *model* of the stimulus, different from the more “verbatim” representation present in primary sensory cortices. As such, these neurons can make predictions about the stimulus, for example the temporal location of future beats. This arrangement is similar to *predictive coding theory* [@Friston2010], which states that the brain creates internal models of sensory input based on experience. By continually comparing the prediction with actual sensory input, the brain is able to evaluate its performance and update the model accordingly. Thus, all sensory input is percieved in the context of models which already exist based on one’s previous experiences with similar stimuli. In the case of beat perception, rhythms will be perceived based on the statistical regularities present in music [@Vuust2014]. For example, since most Western music is in duple meter (where beats are grouped in twos), an isochronous sequence without any physical metrical accents will often be perceived as having a duple meter.

Because of the “automatic” nature of neural resonance, it seems likely that any two networks in the brain could spontaneously fire together provided that a) they are physically connected in some way and b) the preferred patterns of firing are similar. Taken together with predictive coding, this suggests that a model that was developed for one input source could resonate with input from another source. For instance, it has been proposed that motor planning regions draw on their experience with locomotion in an effort to predict auditory sensory input [@Patel2014]. Put into a neural resonance framework, existing oscillatory motor programs (e.g., for walking) that match the auditory sensory input will increase the amplitude of their oscillation, which will feed back to auditory regions to create an expectancy. What is interesting about this case is that a motor program that ostensibly developed to support locomotion falls into the service of auditory perception.

## Hypotheses

What is it about music that makes us want to move? What are the properties of music that support neural resonance in motor regions? The current proposal seeks to investigate the effects of musical and acoustical features of music on neural entrainment in premotor cortex. Additionally, how premotor entrainment is related to movement and synchronization ability will also be examined. Study 1 concerns the effects of groove on neural entrainment in premotor cortices, using a subset of stimuli that have been rated on amount of perceived groove [@Janata2011]. Study 1 has been completed, and found that neural entrainment to the beat in premotor cortices was greater for high-groove versus low-groove music. Critically, this difference across groove categories was not present in auditory or visual cortices, reinforcing the association between groove and movement. See Appendix A for methods and results.

Study 2 concerns acoustical features of music that have been found to contribute to groove—the salience of the beat, for example—and investigates how they modulate neural entrainment and tapping ability. A regression analysis will be used to figure out which acoustical features are most implicated in driving neural entrainment. Study 3 concerns the musical feature of syncopation, which is related to ratings of groove [@Witek2014]. Drum breaks with varying levels of syncopation will be played for participants to investigate its effects on neural entrainment.

Addressing the question of localized neural entrainment requires a brain imaging technique that has good spatial resolution to identify activity from specific regions, as well as good temporal resolution to measure entrainment at specific frequencies. MEG and spatial filtering analysis has been used for this purpose given the high density of sensors in MEG [@Fujioka2010]. Using a this method, Fujioka et al. [@Fujioka2012; described above] searched for brain regions whose beta-band oscillations were phase coherent with auditory cortex. They found the strongest coherence in the cingulate gyrus, and posit that this is likely related to activity in SMA.

Source analysis with EEG is becoming increasingly feasible with the advent of high-density systems.Independent components analysis (ICA) with dipole fitting can be used to localize activity measured via EEG [@Makeig2004; @Onton2006]. ICA is a method of identifying the activity of many sources whose activity has been linearly mixed onto a number of sensors [@Bell1995]. In the case of EEG, there are many sources of electrical activity in the brain, whose activity is measured at all scalp electrodes in different combination [@Makeig1996; @Bell1995]. ICA decomposes the sensor data into a number of possible independent components (ICs) that model the time-series activity of possible independent sources. To localize the ICs, a dipole model can be fit to its scalp topography. Since brain sources are dipolar [@Delorme2012], ICs with well-fitting dipole models (i.e., dipole models with low residual variance) are likely to originate from a single source in the brain. The location of the center of the dipole model is considered to be the location of the source of activity. The accuracy of this method has been corroborated in EEG studies replicating sources found in functional magnetic resonance imaging (fMRI) studies [@McGarry2012], as well as simultaneous fMRI-EEG studies [@Yin2016].

## Study 1: Groove Preferentially Modulates Neural Entrainment in Premotor Cortices

See Appendix A for the methods and results from Study 1.

## Study 2: Premotor Entrainment and Acoustic Features

What physical aspects of an auditory stimulus might affect the ability to synchronize with it? There are many acoustic features that can be extracted from musical recordings, and many have been used for research in behavioural synchronization and groove perception. Features that have been often used in this work are described below. These features can be easily extracted from audio files using MATLAB (The Mathworks) and the MIR Toolbox [@Lartillot2008]; function names for calculating these features are also given below.

*Tempo* is a measure of the frequency of the beat. There are often multiple periodicities present in an excerpt, and so the metrical level chosen for the canonical tempo is largely subjective. However, because metrical levels are related to one another by simple integer ratios (usually 2), it is still possible to match excerpts based on this tempo “profile”.

*Pulse clarity* (mirpulseclarity) is a measure of how easily a listener can perceive the underlying rhythmic structure of the music [@Lartillot2008a]. There are various ways to calculate this feature [@Lartillot2008a; @Madison2011; @Burger2012], and Stupacher et al. [-@Stupacher2016] has found that that different methods are able to predict groove ratings better than others. Madison et al. [-@Madison2011] refers to this as *beat salience*.

*Event density* (mireventdensity) is a measure of the density of note events over time. There are at least two different ways of calculating event density, which differ in their ability to predict groove ratings [@Stupacher2016]. In MIR Toolbox, event density is defined as the average frequency of events and is reported as the number of note onsets per unit time [@Lartillot2008a]. In Madison et al. [-@Madison2011], it is defined as the variance over time of a signal that represents the note onsets in an excerpt.

The *distribution of RMS energy* (mirlowenergy) is a way of quantifying the moment-to-moment changes in RMS energy of an excerpt. It is calculated by finding the number of time frames which are below the average RMS level for the whole excerpt. Higher values indicate that more frames are below the average RMS; therefore, the frames which are above the average must have a greater magnitude difference from the average. Perceptually, this means that louder frames of the excerpt are particularly salient compared to other frames.

*Percussiveness* is a feature used by Burger et al. [@Burger2012], and is defined as the average slope of the amplitude envelope of note onsets. A steeper slope would mean a more percussive sound. For example, compare the sharp attack of a snare drum with the smooth onset of a cello.

*Spectral flux* (mirflux) is a measure of the moment-to-moment changes in an excerpt’s spectrum. It can be averaged across the whole spectrum, or calculated for specific frequency bands (called sub-band spectral flux). For example, Stupacher et al. [-@Stupacher2016] investigated overall as well as sub-band spectral flux in the following bands: 0-50 Hz, 50-100 Hz, 100-200 Hz, 200-400 Hz, 400-800 Hz, 800-1600 Hz, 1600-3200 Hz, 3200-6400 Hz, 6400-12800 Hz, and 12800-22050 Hz. Spectral flux in lower frequency bands (i.e., 0-50 Hz, 50-100 Hz, and 100-200 Hz) has been previously associated with more consistently-timed movements [@Burger2012].

### Acoustic Features and Behavioural Synchronization

Stupacher et al. [-@Stupacher2016] investigated the effects of frequency on the ability to accurately and consistently tap along with music. They experimentally manipulated the frequency of the bass drum and the bass and asked participants to tap along to the beat as synchronously as possible. They found that tapping was less variable and more forceful when the bass instruments had lower frequencies. Thus, features like low-frequency energy are likely related to neural entrainment in premotor cortices. Interestingly, groove ratings were also higher for these excerpts.

Burger et al. [-@Burger2012] has used motion capture techniques to elucidate the relationship between acoustic features of music and spontaneous full-body movement. Participants listened to 30-second excerpts of music and were asked to move in a way that felt natural, and were encouraged to dance if they wanted. Correlation analyses revealed that music with a clear pulse and with high percussiveness elicited movements with low spatial variation[[3]](#footnote-3), indicating that participants’ movements were simpler and exhibited fewer degrees of freedom. Temporal regularity—movement at the same frequencies as the meter—was related to these features as well as low-frequency spectral flux. In a second study, they found that pulse clarity was related to regularity in movements of all body parts, whereas spectral flux and percussiveness were more related to specific body parts such as the head and hands [@Burger2013].

### Acoustic Features and Groove

Various acoustic features of musical stimuli have been previously associated with groove. Madison et al. [-@Madison2011] found that event density and pulse clarity were both associated with groove ratings. Faster tempos have also been associated with groove [@Janata2011], although Study 1 found that slower tempos were associated with premotor entrainment. By analyzing the corpus of musical examples used in Janata et al. [-@Janata2011], Stupacher et al. [-@Stupacher2016] identified RMS[[4]](#footnote-4) variability, spectral flux, and event density as features that were predictive of groove ratings. Additionally, they found that songs with a drum kit differed significantly from songs without a drum kit on all features studied, suggesting a role of instrumentation.

### Acoustic Features and Neural Entrainment

ERP work has found that lower frequencies are more attended to for timing information. Hove et al. [@Hove2014] presented participants with an isochronous sequence of two simultaneous tones at different frequencies: one high tone and one low tone. When the timing of the low tone was perturbed, they found larger mismatch negativity responses than when the timing of the high tone was perturbed. This suggests that lower frequency information is relied upon more heavily for understanding rhythm.

TODO (B) study 2 methods

### Aims and Considerations

Study 2 aims to discover how different physical aspects of a musical stimulus will preferentially entrain motor regions of the brain. How this neural entrainment might be related to behavioural measures such as movement synchronization (i.e., tapping) are also investigated.

The main goal will be to develop a regression model that can predict entrainment ability and/or premotor neural entrainment on the basis of acoustic features of excerpts. To this end, a large stimulus set will be used.

Given the large effect of tempo in Study 1, it is possible that the effect of tempo will swamp the effects of other features in the regression. So, an attempt will be made to keep the excerpts in the stimulus set to a narrow range of tempos.

Using the complete stimulus set (approximately 150 excerpts) from Janata et al. [-@Janata2011], songs without a drum kit will be removed, and various acoustic features will be extracted from the remaining excerpts. A stimulus set will then be constructed that attempts to equate approximately 30 excerpts on all of the relevant features. Except where noted, they will be extracted using MIR Toolbox [@Lartillot2007], a MATLAB toolbox developed for this purpose.

Post-hoc exploratory analyses of the stimuli in Study 1 revealed clues as to what acoustic features of high-groove music might be driving the enhanced premotor entrainment. First, it is notable that while almost all of the high- and mid-groove songs included a drum kit, only one of the low-groove songs did. Additionally, entrainment to the beat frequency was associated with music that had larger loudness peaks throughout. Musical features have been

**Groove Ratings.**

**Behavioural synchronization.**

**Neural entrainment.**

TODO (A) study 3 intro

## Study 3: Premotor Entrainment and Musical Features (Syncopation)

Syncopation is a feature of music that occurs when some notes are played on metrically weak beats instead and not on metrically strong beats. Longuet-Higgins and Lee [-@Longuet-Higgins1984] define syncopation as a note occurring on a metrically weak beat that precedes a rest (no note) on a metrically strong beat.

define syncopation. Syncopation is a feature of music that has been linked to groove.

Witek et al. [-@Witek2014] played drum breaks for participants that varied in their degree of syncopation.

Many definitions of syncopation, but it is usually described as a shift of rhythmic emphasis from metrically strong to metrically weak beats. - Fitch, W. T., and Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. Music Percept. 25, 43–58. doi: 10.1525/mp.2007.25.1.43 - Longuet-Higgins & Lee 1984 computational model of rhythmic syncopation

Also get participants to tap along to the various syncopated rhythms. Nozaradan et al. [-@Nozaradan2013] found that harmonic peaks in the SSEP went away during actual tapping as compared to when just listening.

We would expect that for rhythms which are associated with more accurate tapping would show better entrainment in premotor areas.

Stimuli will consist of a subset of the 50 drum breaks used in Witek et al. [-@Witek2014]. As was done in that study, drum breaks will be created in Garageband (Apple, Inc.) using MIDI. The musical score and syncopation indices are available in the supplementary materials of Witek et al. [-@Witek2014]. All drum breaks will be presented at 120 beats per minute (2 Hz) for 32 seconds (16 bars; 16 repetitions of the drum break). If you find an effect of tempo in study 2, you might choose to manipulate tempo in study 3.

Following each drum break, participants will be asked to rate the excerpt on a 5-point Likert scale for amount of groove and amount of enjoyment (as in Witek et al. [-@Witek2014]). In the first block of trials, participants will listen to the drum breaks as in Studies 1 and 2. In the second block of trials, participants will be instructed to tap along with the drum breaks.

**Groove ratings.** It is expected that the groove ratings will replicate from Witek et al. [-@Witek2014], where groove ratings and syncopation index will have an inverted-U relationship. High and low levels of syncopation should elicit the lowest groove ratings.

**Behavioural synchronization.** It is also expected that tapping accuracy and variability will demonstrate an inverted-U relationship with syncopation index.

**Neural entrainment.** Premotor entrainment is expected to be greater for optimally-syncopated drum breaks.

TODO discussion section

## General Discussion

This discussion section is really just notes at this point. Does there need to be an extensive discussion section for a proposal?

Another theory of interest to this discussion is predictive coding theory [@Friston2010], which suggests that the brain creates internal models of sensory input based on experience. This model is used to generate predictions about future events. By continually comparing the prediction with actual sensory input, the brain is able to evaluate its performance. Any prediction errors are used to update the model. Thus, the brain’s goal is to minimize the prediction error. In the case of beat perception, rhythms are perceived based on one’s previous experiences [@Vuust2014], and so the statistical regularities present in music will be used to inform any new sensory input. For example, an isochronous sequence without any metrical accents will often be perceived as having a duple meter (where beats are grouped in twos), since most Western music is based on duple meter.

### Embodied Cognition

Burger et al. [-@Burger2013; -@Burger2014] discuss their findings in terms of embodied cognition. i.e., synchronizing movements with music is a method of coming to perceive or understand it.

### An Addiction Model of Groove

* addiction model of beat perception and groove?
  + syncopated rhythm is initially rewarding due to predictive error and updating the model
  + this is so enjoyable that you start to crave it
  + so you come to expect the hit of DA
* groove creates a desire to move by activating premotor programs (system is primed)
  + because system is primed, the nigro-striatal pathway is also primed (volitional movement, PD is from cell death in SN)
  + VTA and SN are very close to one another
  + perhaps SN activation (in prep for movement) entails VTA activation
  + VTA will activate mesolimbic reward systems (also mesocortical?)

### Future Directions

**Motor Cortex Entrainment in Parkinson’s Disease** - clinical population - jessica grahn rhythms that will help walking in parkinsons - optimization of musical parameters - lab in singapore has software for tracking gait with fitbit or similar - this could be a single chapter in the dissertation

**Study x: Speech (and kids books?)** - Optimal level of “syncopation” for good oration - optimal level of sync for entraining people (music) and engaging people (speech) - entrainment vs simulation - if too boring, entrainment washes away - look at alpha entrainment and suppression in same study.. connect to mu work - what does groove have to do with a great speech? quite a lot. it’s all about entrainment (to create predictability and expectation; inform the listener when to expect important events) and desynchronization (have to keep things interesting otherwise the regularity will be too predictable, and people will just tune out: “i’ve got this, i don’t need to keep entraining”).

# Appendix A

## Study 1: Groove Preferentially Modulates Neural Entrainment in Premotor Cortices

Groove is a feature of music that compels movement, and Janata et al. [-@Janata2011] have created a large corpus of popular musical excerpts for which they have obtained ratings of groove from 215 participants. Study 1 uses a subset of these stimuli to investigate the effects of groove on the ability of neurons to entrain to the beat. Since groove compels movement, we predicted that premotor cortices would entrain more to the beat of high-groove excerpts, compared to low-groove ones. This would provide a neural basis for the association between a sense of groove and the desire to move.

## Methods

### Participants

Eleven participants with self-reported normal hearing were recruited from the Ryerson University psychology student participant pool. This study was approved by the Ryerson Research Ethics Board and all participants completed informed consent prior to the study.

### Stimuli

Thirty pieces of music representing high- medium- and low-groove were selected on the basis of ratings provided by Janata et al. [-@Janata2011] Excerpts were restricted to the first 30 seconds of each piece on the basis of previews available on the iTunes Music Store (Apple, Inc.). The tempo of each excerpt was extracted manually by the first author. Seven excerpts had tempi that changed over the course of the excerpt and were thus excluded from analysis. See Table A1 for a complete list of the excerpts used.

### Procedure

Participants were seated at a computer in a sound-attenuated chamber. Stimulus presentation was controlled using Presentation software (Neurobehavioral Systems) running on a PC. Excerpts were played over Logitech Z130 loudspeakers in two blocks. Each block contained one presentation of each excerpt. The order of presentation within blocks was individually randomized. Participants pressed the “space” bar on the keyboard to initiate each excerpt. Participants were asked to listen to each excerpt carefully but to refrain from moving.

EEG data were recorded using a 64-channel BioSemi ActiveTwo system and were digitized at 512 Hz. Analyses were carried out using a combination of custom MATLAB (version R2016a; The Mathworks, Inc.) scripts and EEGLAB [version 13.5.4b; @Delorme2004].

**Preprocessing and Independent Components Analysis.** To improve the quality of the ICA decomposition, various preprocessing steps were executed [@Winkler2015]. Low-frequency drift is a common non-stationary artifact, which has the potential to detract from the reliability and dipolarity of the ICs [@Debener2010]. Low-frequency drift can be removed with a high-pass filter at 1 Hz or above [@Winkler2015]. However, since we were interested in beat frequencies near 1 Hz, two preprocessing pipelines were necessary: one with a high-pass filter at 1 Hz to obtain the ICA weights, and a second with a high-pass filter at 0.1 Hz to obtain activity within each electrode. The ICA weights from the first pipeline were then used with the preprocessed electrode activity from the second pipeline in order to obtain IC activity that was free of low-frequency drift.

For the first pipeline, raw EEG data were high-pass filtered at 1 Hz using the EEGLAB function pop\_eegfiltnew. Bad channels were identified using the clean\_rawdata plugin. This plugin works by first windowing the data into 1-second non-overlapping windows. In each window, each channel is interpolated based on all other channels, and this interpolation is correlated with the actual data from that channel. If this correlation is lower than 0.8, the channel is flagged for the given window. Any channel with flags in more than half of the windows is rejected. Data were then rereferenced to the average, and line noise was removed (pop\_cleanline). Sections of data where no music was presented (i.e., while entering responses on the computer) were removed (pop\_epoch), and the remaining data were subjected to the extended version of the Infomax ICA [@Bell1995] as implemented in EEGLAB (pop\_runica).

In the second pipeline, raw EEG data were high-pass filtered at 0.1 Hz. Channels that were identified as bad in the first pipeline were removed. All other channels were referenced to the average and epoched. Next, the ICA weights and sphering matrices (i.e., EEG.icaweights and EEG.icasphere) obtained from the first pipeline were used with the electrode data from the second pipeline to compute IC activity.

Regions of interest were defined based on Brodmann areas (BA): the premotor cortex region comprised BA6, the auditory cortex region comprised BA41, BA42, and BA22, and the visual cortex region comprised BA17, BA18, and BA19. In order to localize components with respect to these target brain regions, equivalent dipole models were fit to each IC (pop\_dipfit). For each region of interest, an IC was selected if its Talairach coordinates were within 11 mm[[5]](#footnote-5) of the target brain region [@Lancaster1997; @Lancaster2000] and the residual variance of its corresponding dipole model was less than 15%.

**Entrainment Analysis.** The following steps were implemented to determine extent of entrainment in each IC:

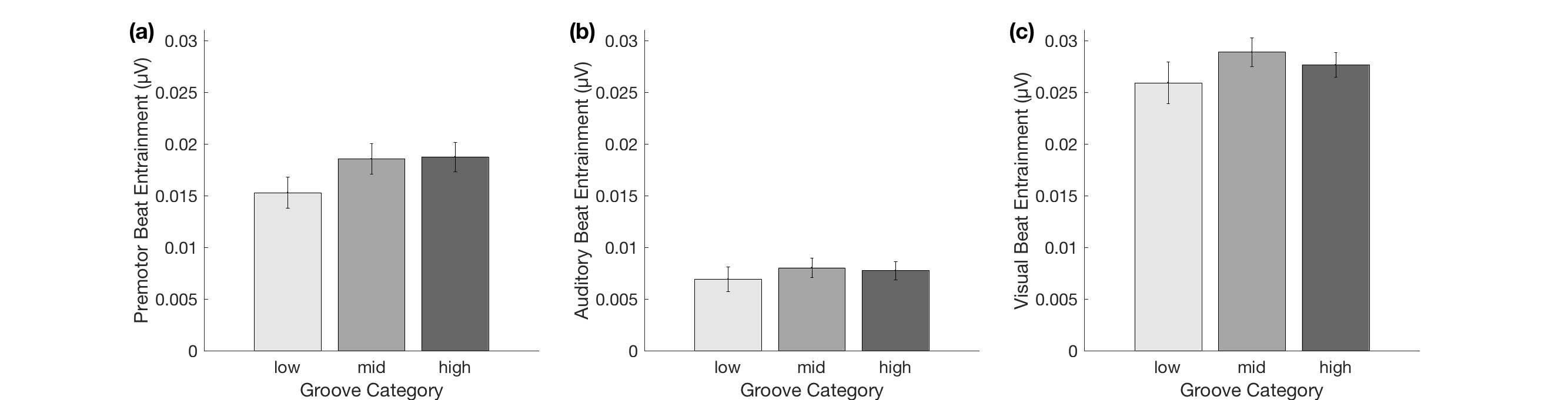
1. The activity in the two presentations of each excerpt was averaged in the time domain and subjected to a fast Fourier transform with bin width of 0.0312 Hz.
2. Since there was no temporal baseline data with which to normalize each trial, the noise floor was removed from each bin by subtracting activity in the surrounding bins as in Nozaradan et al. [-@Nozaradan2011]. Specifically, two bins on either side of the target frequency that were two bins away from the target frequency were averaged together. The result was subtracted from the value in the target frequency bin. This procedure assumes that in the absence of any stimulation, there will be no peaks of activity significantly higher than the noise floor, thus the result would be zero.
3. The extent of entrainment was defined as the maximum value across the eleven frequency bins centered on the beat frequency. The maximum was used instead of the mean because a) the stimuli were not computationally generated, so we cannot assume the tempo was perfectly consistent throughout each excerpt, and b) the tempo was extracted manually for each excerpt and so we cannot assume they are exact.

## Results

To ensure that the entrainment values were significantly higher than the spectral noise floor, a *t*-test was carried out across all excerpts for each participant and target brain region. All *t*-tests were significant. For participants with multiple sources localized in the same region, only the maximum entrainment value was subjected to analysis. For participants without any sources localized to a given region, the entrainment value was set to zero (i.e., no entrainment in that region).

A one-way ANOVA with Brain Region (Premotor, Auditory, and Visual) as the within-subjects factor revealed that entrainment differed across brain regions (*F*(2,20) = 15.875, *p* < 0.001). Visual had the highest level of entrainment (mean = 0.0266 µV) followed by Premotor (mean = 0.0168 µV) and Auditory (mean = 0.0075 µV). Because of the large differences in magnitude of entrainment across brain regions the effect of groove category was assessed using three one-way ANOVAs.

Groove Category had a significant effect in the Premotor Region (*F*(2, 20) = 5.156, *p* = 0.016), wherein entrainment was greater for high-groove music (mean = 0.0176 µV) and mid-groove music (mean = 0.0182 µV) than low-groove music (mean = 0.0145 µV; Figure 5a). Post-hoc tests revealed that there was no difference between high-groove and mid-groove entrainment (*p* = 0.534), and that high-groove entrainment was greater than low-groove entrainment (*p* = 0.047). Groove Category was not significant for Auditory (*F*(2,20) = 1.533, *p* = 0.240; Figure 5b) or Visual Regions (*F*(2,20) = 2.325, *p* = 0.124; Figure 5c).



*Figure A1.* (a) Entrainment to the beat frequency in premotor cortex (BA6) was significantly greater for high- and mid-groove songs than for low-groove songs. Entrainment in (b) auditory (BA41, BA42, BA22) and (c) visual (BA17, BA18, BA19) cortices did not differ across groove categories. Error bars represent standard error.

## Discussion

Study 1 found evidence that premotor cortices will preferentially entrain to high-groove music. In contrast, auditory and visual cortices did not exhibit a preference for groove, suggesting that this is specific to motor regions. This provides a neural basis for previous reports that a sense of groove is related to a desire to move along with the music [@Janata2011], and offers a neural mechanism by which motor output might be better synchronized to high-groove versus low-groove music. Additionally, Study 1 successfully used an emerging method of EEG analysis in order to study both temporal and spatial aspects of brain activity.

The relatively strong level of neural entrainment observed in visual cortex was somewhat unexpected. We might have expected greater levels of entrainment in auditory over visual areas given that our stimuli were auditory. We might have also expected greater levels of entrainment in premotor over visual areas because music compels movement. One possible explanation for the visual advantage may be the scale of neuronal populations involved; the visual cortex might simply have more neurons. Once entrained, the larger population of neurons may lead to a stronger response regardless of the driving modality. It is also possible that premotor areas were inadvertently suppressed by our instructions to refrain from moving.

# Tables

Table A1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Song | Artist | Groove Rating | Groove Category | Tempo (Hz) |
| Beauty of the Sea | The Gabe Dixon Band | 32.1 | low | 1.90 |
| Can’t Let Go | Lucinda Williams | 58.9 | mid | 1.58 |
| Citi Na GCumman | William Coulter and Friends | 35.2 | low | - |
| Dawn Star | Dean Magraw | 34.8 | low | - |
| Down With Love | Blossom Dearie | 57.0 | mid | 2.33 |
| Druid Fluid | Yo-Yo Ma, Mark O’Connor, and Edgar Meyer | 38.1 | low | 1.95 |
| Flandyke Shore | The Albion Band | 36.5 | low | 1.50 |
| Flash Light | Parliament | 105.1 | high | 1.77 |
| Fortuna | Kaki King | 32.6 | low | - |
| Freedom of the Road | Martin Sexton | 59.7 | mid | 2.45 |
| How High the Moon (1st take) | Ella Fitzgerald | 65.2 | mid | 1.88 |
| Hymn for Jaco | Adrian Legg | 29.3 | low | - |
| If I Ain’t Got You | Alicia Keys | 98.7 | high | 2.00 |
| In the Mood | Glenn Miller and His Orchestra | 96.9 | high | 2.73 |
| It’s A Wrap (Bye, Bye) | FH1 (Funky Hobo #1) | 105.9 | high | 1.53 |
| Lady Marmalade | LaBelle | 102.5 | high | 1.92 |
| Mama Cita (Instrumental) | Funk Squad | 101.6 | high | 1.58 |
| Music | Leela James | 101.1 | high | 1.63 |
| Ray Dawn Balloon | Trey Anastasio | 38.5 | low | 2.77 |
| Run / Immortal | Beth Hart | 60.8 | mid | 1.92 |
| Sing, Sing, Sing | Benny Goodman and His Orchestra | 97.4 | high | 1.80 |
| Space Oddity | David Bowie | 38.7 | low | 2.27 |
| Squeeze | Robert Randolph and the Family Band | 63.4 | mid | 1.95 |
| Summertime | Ella Fitzgerald and Louis Armstrong | 67.9 | mid | - |
| Superstition | Stevie Wonder | 108.7 | high | 1.67 |
| Sweet Thing | Alison Brown | 30.9 | low | - |
| The Child Is Gone | Fiona Apple | 62.3 | mid | 1.08 |
| The Girl From Ipanema | Astrud Gilberto, Joao Gilberto, and Stan Getz | 57.0 | mid | 2.15 |
| Up for the Downstroke | The Clinton Administration | 102.4 | high | - |
| What a Wonderful World | Louis Armstrong | 66.4 | mid | 1.23 |

# References

1. An isochronous sequence is a rhythm that consists of tones which are equally-spaced in time, such as a metronome. If all the tones in the sequence are identical (e.g., have the same pitch, loudness, etc.), then there is only one beat level present in it. [↑](#footnote-ref-1)
2. One participant responded “don’t know” [↑](#footnote-ref-2)
3. Spatial variation is a measure of the intrinsic dimensionality of a data set. More variables are required to represent more complex data sets, and would thus be considered to have higher spatial variation. For example, given a principle components analysis of some data, a measure of spatial variation could be the number of components required to explain a certain amount of variance. [↑](#footnote-ref-3)
4. RMS (root mean square) is a measure of the total energy present in a waveform. It is calculated by (a) squaring each value in the waveform, (b) averaging the values at all time points, and (c) taking the square root of the result. [↑](#footnote-ref-4)
5. The localization procedure creates a 3D cube centered on the given Talairach coordinates resolved to a 1 mm3 voxel in a standardized Talairach space. The cube extends 5 mm on all sides resulting in a total size of 11 mm3. [↑](#footnote-ref-5)