Neuro-oscillatory tracking of low- and high-level musico-acoustic features during naturalistic music listening: insights from an SEEG study.

Diana Omigie[ab], Katia Lehongre[cd], Vincent Navarro[cd], Claude Adam[cd], Severine Samson[ce].

[a] Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany.

[b] Goldsmiths, University of London, London, UK.

[c] AP-HP, GH Pitié-Salpêtrière-Charles Foix, Epilepsy Unit, F-75013, Paris, France.

[d] Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Brain and Spine Institute, ICM, F-75013, Paris, France.

[e] Laboratoire de Neurosciences Fonctionnelles et Pathologies, EA4559, Université Lille-Nord de France, Villeneuve d'Ascq, France.

Author Note: This work and dataset have not been previously disseminated elsewhere.

Correspondence concerning this paper should be addressed to: Dr. Diana Omigie, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main. Germany. Email: d.omigie@gmail.com

Abstract

Studies investigating the neural processing of musico-acoustic features have tended to do so using highly controlled musical stimuli. However, it is increasingly argued that failing to use naturalistic stimuli limits the extent to which findings from lab studies can be extrapolated to rich and varied real-world experiences. Here, we recorded electrical brain activity from 8 epileptic patients, implanted for pre-surgical evaluation with Stereo-encephalography (SEEG), while they listened to pieces from the western tonal music repertoire. We estimated the *sound intensity* and *key* and *pulse clarity* of the stimuli using a toolbox for automatic extraction of musico-acoustic features. We then used partial-correlation analyses to examine the patterns of neuro-oscillatory activity associated with the processing of these features. Our results showed clear tracking of sound intensity in high-gamma and alpha frequency bands in posterior superior temporal gyrus, reflecting neural firing and the transfer of auditory information from the thalamus to auditory cortices, respectively. Patterns of partial correlations, in line with our hypotheses, also suggested limbic and inferior frontal cortical tracking of tonal and rhythmic uncertainty, albeit without the robustness shown for sound intensity tracking in auditory areas. The study provides an important contribution to the existing literature in its adherence to the call for a greater use of ecologically valid stimuli in neuroscientific investigations of music listening. Our results, specifically, have implications for research on the neural processing of musical uncertainty and for future studies seeking to use intracranial EEG to examine naturalistic music processing.

Keywords: music, oscillations, auditory cortex, intracranial electroencephalography, uncertainty.

Neuro-oscillatory tracking of low- and high-level musico-acoustic features during naturalistic music listening: insights from an SEEG study.

Introduction

Failing to clarify the neural correlates of information processing in the context of naturalistic stimuli limits the extent to which findings from laboratory studies can be extrapolated to rich and complex real-world experiences (Hasson, Nir, Levy, Fuhrmann & Malach, 2004). Indeed, studies examining the neural processing of musico-acoustic features have tended to do so using highly controlled stimuli, the result being an incomplete understanding of the range of neural mechanisms underlying everyday music listening. A valid reason for this state of affairs is that scalp electro-encephalography (EEG) and magneto-encephalography (MEG), the key techniques used to measure electrical activity from the brains of healthy participants, provide recordings that are necessarily measured from outside of the skull. The relatively low signal to noise ratio (SNR) possible with such recordings, in turn, necessitates signal averaging of numerous trials for the effect of any stimulus features of interest to be characterized.

In contrast, intracranial electro-encephalography (iEEG) recordings, measured directly from brain tissue, provide the higher SNR that is needed to examine neuro-oscillatory correlates of music processing without the need for signal averaging. Intracranial recordings include electrocorticography data (ECoG data) recorded using electrode grids that sit on the surface of the brain (Buzsáki, Anastassiou, & Koch, 2012) and stereo-encephalography (SEEG), recorded using electrodes inserted deep within subcortical regions of the brain (e.g. see Omigie, Dellacherie, Hasboun, George, Clement, Adam, & Samson, 2015; Omigie, Dellacherie, Hasboun, Clement, Adam, & Samson, 2015; Omigie, Pearce, Lehongre, Hasboun, Navarro, Adam, Samson, 2019).

As access to iEEG data is limited by access to patients implanted for clinical and surgical reasons, intracranial recordings are, in general, rare in the context of music listening studies. However, the small number of studies that have used ECoG to examine the neuro-oscillatory correlates of naturalistic music listening suggest that, even given this method's high SNR, very few musico-acoustic features show clear tracking by neural oscillations (Potes, Gunduz, Brunner, & Schalk, 2012; Potes, Brunner, Gunduz, Knight, & Schalk, 2014; Sturm, Blankertz, Potes, Schalk, & Curio, 2014). Specifically, while such studies have shown significant correlations between levels of sound intensity and levels of auditory high gamma cortical activity (between 70 and 170 Hz in Potes et al., 2012 and Sturm et al., 2014; and between 70 and 110Hz in Potes et al., 2014), evidence for the neuro-oscillatory tracking of higher level musical features remains limited. One possibility is that expecting to see high gamma power (HGP) tracking of high level musical features is unrealistic. The key of a piece of music denotes the group of pitches, organized around a central note (known as the tonic), that form the basis of the piece's construction. A key clarity time-course (i.e. how clearly the key is reflected over time) is therefore a highly abstract feature of a piece of music, as compared to, say, the amplitude envelope (or sound intensity) of the piece. Such an abstract feature of a stimulus may, in turn, be unlikely to be tracked by HGP – a neural signature that is more commonly associated with the processing of low level sensory features of a stimulus.

However, another possibility is that neural tracking of such high level musico-acoustic features exists, albeit in specific anatomical regions and oscillatory frequency bands that have not previously been interrogated. In contrast to previous studies, which have tended to focus on high frequency power correlations in exclusively lateral cortical areas (due to the use of ECoG, where contacts sit on the surface of the brain (Sturm et al., 2014)), here we examined, using

SEEG, evidence for neuro-oscillatory tracking of sound intensity, key clarity and pulse clarity in a wide range of frequency bands and across a wide range of cortical and subcortical regions. In doing so, we sought to 1) extend previous work that was constrained in terms of anatomical and oscillatory frequency range and 2) corroborate the evidence for a role of specific limbic and frontal areas in the processing of musical structure.

Sound intensity processing

Sound intensity may be defined as the instantaneous energy in an acoustic signal. Loudness, classically defined as magnitude of an auditory sensation (Fletcher & Munson, 1933), is considered the perceptual correlate of sound intensity. Scalp EEG studies, employing event-related trial averaging to obtain event-related potentials (ERP), have provided evidence for the sensitivity of auditory cortical areas to sound-intensity (Jacobsen, Horenkamp, & Schröger, 2003; Schadow, Lenz, Thaerig, Busch, Fründ, & Herrmann, 2007). Further, neural activity in brain structures at even lower stages of the auditory pathway (e.g. the inferior colliculi and thalamus) have also been shown to linearly correlate with perceived loudness of a stimulus (Röhl & Uppenkamp, 2012).

Most recently, ECoG studies have been able to show reliable evidence of sound intensity being tracked by neural oscillatory activity during naturalistic music listening (e.g. Potes et al., 2012; Potes et al., 2014). High gamma power (HGP) is held to index the rate of firing of underlying neurons (Miller, 2010; Ray & Maunsell, 2011) and HGP (between 70 and 170 Hz) in the posterior superior temporal gyrus (STG) and precentral gyrus have been shown to track the sound intensity of a popular music track (Potes et al., 2012). The robustness of the response in these lateral frontal and temporal areas has been further demonstrated in a study in which time-courses of oscillatory activity, estimated at contacts in anatomical locations correspondent across participants, were correlated (Potes et al., 2014). In that study, HGP (70 to 110 Hz) correlations (across participants) were revealed in superior temporal and premotor cortical areas, with a further negative correlation observable between HGP and alpha power in the STG. The authors associated this pattern of neuro-oscillatory tracking with local neural firing in response to intensity (HGP) on the one hand, and, on the other, the thalamo-cortical connectivity (alpha activity) that such local firing drives (Potes et al., 2014).

Those studies provided important evidence for neuro-oscillatory tracking of music in lateral cortical areas. However, they were unable to address the extent to which changes in musico-acoustic features, including sound intensity, are also tracked by limbic areas. Sound intensity fluctuations reflects the so-called dynamics of a piece. Indeed, such fluctuations are not only indication of when musical events are sounding but are also used, within performances, to express emotional intensity (Palmer, 1997). The influence of musical dynamics on emotion perception and induction (Ilie & Thompson, 2006; Juslin 2008; Sloboda & Lehmann 2001) begs the question of whether the tracking of sound intensity should be anticipated in limbic structures involved in emotion processing. Structures like the anterior cingulate gyrus (ACG) and insula have not only been associated with auditory salience (Lockwood et al., 1999; Seeley et al., 2007) but also with modulation of the autonomic nervous system (Critchley et al., 2003). Using functional magnetic resonance imaging (fMRI), modulations in these areas have been shown in response to variations in the tempo and dynamics of musical stimuli (Chapin, Jantzen, Kelso, Steinberg, & Large, 2010). However, evidence that dynamics (changes in sound intensity) alone can result in modulations of limbic activity remains limited.

Tonality processing

Tonality, the principle of organizing a composition around a hierarchy of pitches, underlies the construction of many musical works. Simply put, any Western or non-Western music that periodically returns to a central tone (the tonic) can be said to exhibit tonality. However, the western tonal system, which is based on the division of the octave into twelve equal intervals, is structurally different to, for instance, the tonal system of most Chinese music, which is based on five core tones. Thus, in contrast to sound intensity, and on account of the learning and enculturation necessary for listeners to process music in terms of tonal hierarchies, tonality can be considered a high-level feature of musical structure.

Accordingly, processing of tonal structure has been shown to recruit a wide network of cortical brain structures that extend beyond the auditory cortices. Specifically, studies tend to emphasize the importance of dorso-medial frontal cortical regions in tracking musical key (Janata, Birk, Van Horn, Leman, Tillmann & Bharucha, 2002; Janata, 2005). Furthermore, tonal expectancy violations have been associated with the so-called early right anterior negativity or ERAN (Koelsch et al., 2001), a negative going ERP which originates from both the STG and Brodmann areas 44 and 45 (BA44/45) of the inferior frontal gyrus (IFG: Besson & Faita, 1995; Brattico, Tervaniemi, Näätänen, & Peretz, 2006; Maess, Koelsch, Gunter & Friederici, 2001; Omigie, Pearce, Williamson & Stewart, 2013; Sammler et al., 2013; Tillmann, Koelsch, Escoffier, Bigand, Lalitte, Friederici, & von Cramon, 2006).

However, while points of low key clarity in music have also been associated with increases in pre/post central gyrus, supplementary motor area (SMA) and rolandic operculum activity in neuroimaging studies using naturalistic musical stimuli (Alluri, Toiviainen, Jääskeläinen, Glerean, Sams, & Brattico, 2012), a recent electrocorticography study assessing naturalistic music listening was unable to show similarly strong evidence for these or any other frontal areas in tracking changes in tonal structure (Sturm et al., 2014). In that study, electrodes from only one out of ten patients showed active tracking of tonal structure (Sturm et al., 2014). One possibility that was suggested to account for this failure to track high level features using ECoG is that such features are processed in lower frequency bands like theta (4 -7 Hz), which have yet to be examined using ECoG (Sturm et al., 2014). Another potential reason - not considered so far- for failure to observe higher level feature tracking using ECoG, is that these features are most robustly processed in limbic areas, which have also yet to be examined with iEEG.

Violations of tonal structure have been shown to activate the autonomic nervous system (e.g. Egermann, Pearce, Wiggins, & McAdams, 2013; Steinbeis, Koelsch, & Sloboda, 2006) and therefore such events may also be expected to activate the specific brain areas that modulate this part of the peripheral nervous system. In line with this assumption, previous findings from a scalp EEG study show that frontal midline theta power, presumably from cingulate cortex, may be involved in the processing of tonality (Sammler, Grigutsch, Fritz, & Koelsch, 2007). Indeed, using SEEG, theta power in the amygdala, another key structure of the limbic system, has been shown to be greater for harmonic events that violate expectations than those that do not (Mikutta et al., 2015). Furthermore, fMRI studies of naturalistic music listening have revealed increased insula and cingulate gyrus activity at points of low key clarity (Alluri et al., 2012). Here it is important to note that the processing of musical expectation violations and the processing of modulations in key clarity are likely both strongly associated with the psychological state of uncertainty and tension arousal (Huron, 2006; Koelsch, 2012; Koelsch, 2014; Meyer, 1956; Narmour, 1990; Rohrmeier & Koelsch, 2012). Violation of musical expectations may lead to uncertainty and tension due to the listeners' anticipation of a resolution in the music after such expectation breaches (Meyer, 1956). Similarly, particularly in tonal music, which is, unlike atonal music, characterized by high key clarity (Mencke, Omigie, Wald-

Fuhrmann & Brattico, 2018), points of low key clarity (reflecting the presence of chromaticism or changes in key) may lead to tension due to the listeners' anticipation of the imminent return of a clear, albeit perhaps even a new, key.

Thus, taken together, it is possible that robust evidence of the neuro-oscillatory tracking of key clarity will be obtained by examining lower frequency oscillatory activity (Sturm et al., 2014) and this, not just in lateral but also, in medial limbic areas that are known to be involved in music-induced emotion processing (Koelsch, 2012; Koelsch, 2014; Lehne, Rohrmeier, & Koelsch, 2014). The current study sought to examine this possibility.

Rhythm processing

Finally, in addition to sound intensity and tonal clarity, the current study sought to examine the evidence for neuro-oscillatory tracking of music's temporal structural in lateral and medial cortical areas during naturalistic music listening. The processing of music's temporal structure is known to recruit both auditory and sensorimotor cortical areas involved in audition and movement, including the ventral premotor cortex (vPMC, lateral Brodman area 6), which facilitates synchronization to musical rhythms (Chen, Zatorre & Pehune, 2006) and the SMA, which is activated by the presence of a clear beat even during passive listening (Grahn & Rowe, 2012). With regard to the involvement of auditory areas (lateral temporal lobe), both neuro-imaging (Abrams, Bhatara, Ryali, Balaban, Levitin, & Menon, 2010; Fujioka, Ross & Trainor, 2015), neuropsychological, and neurophysiological studies (Creutzfeldt, & Ojemann, 1989; Di Pietro, Laganaro, Leemann, & Schneider, 2004; Fries & Swihart, 1990) have shown that STG and middle temporal gyrus (MTG) are recruited during passive exposure to implied beats, with the degree of their activation reflecting individual differences in beat perception (Grahn & McAuley, 2009).

Listening to music with metrical structure is an intrinsically predictive process and metrical structure, as is the case for tonal structure, sets up expectancies in enculturated listeners (Jones & Boltz 1989; Jones, Moynihan, MacKenzie, & Puente, 2002; Jones & Pfordresher, 1997; Large & Jones, 1999; Vuust & Witek, 2014; Zanto et al., 2006). Bilateral STG, right parietal cortex and right inferior frontal operculum (Grahn & Rowe, 2012) have been shown to be more active in the context of low pulse clarity, and are held to play the crucial role of encoding individual lengths in the absence of a clear periodic structure (Harrington, Haaland, & Hermanowicz, 1998; Schwartze, Keller, Patel & Kotz, 2011; Rao, Mayer & Harrington, 2001). However, as is the case for key clarity, points of low pulse clarity in naturalistic music have also been associated with elevated blood oxygen level dependent (BOLD) activity in limbic areas, with Alluri and colleagues (2012) reporting greater activity in cingulate and parahippocampal gyrus, at such points. Further, key oscillatory patterns in response to music's temporal structure have been revealed using MEG. For instance, the delta frequency band has been shown to entrain to the tempo of music heard (Doelling & Poeppel, 2015), while increases and decreases in beta and gamma power have been shown to track the onset of the musical beat (Fujioka, Trainor, Large, & Ross 2009; Fujioka, Ross, & Trainor, 2015).

These studies provide an interesting but still largely incomplete account of the mechanisms involved in processing naturalistic music's temporal structure, especially with regard to the nature of involvement of beta and gamma oscillations, and the neural sources of these oscillations. The current study sought to extend this work by examining the extent to which not just lateral cortical brain oscillations but also limbic brain oscillations track the uncertainty associated with changes in pulse clarity.

The current study

Taken together, studies examining changes in musical sound intensity and key and pulse clarity suggest distinct patterns of responses in frontal and temporal brain areas (Alluri et al., 2012). However, evidence for neuro-oscillatory tracking of high level features especially and the role of limbic and frontal areas, in general, remains weak (Sturm et al., 2014).

The current study sought to extend previous work examining the evidence for neuro-oscillatory tracking of musical features, by interrogating a wider range of frequency bands and a wider range of anatomical areas. In doing so, it sought to corroborate previous findings of a role of limbic and frontal areas in processing musico-acoustic features that contribute to musical expression and emotion. We recorded SEEG brain activity while patients, in the context of epilepsy treatment, listened to four musical excerpts that used rich instrumentation and expressive timing. In line with previous work, we used the MIRTOOLBOX to extract *Sound intensity, Key Clarity* and *Pulse Clarity* of the musical stimuli (Alluri et al., 2012) and then examined the extent to which these musico-acoustic time-series correlated with neuro-oscillatory power time-courses (e.g. Sturm et al., 2014) taken from both laterally and medially implanted frontal and-temporal cortical electrodes.

Critically, as changes in sound intensity, as well as in key and pulse clarity, over time are unlikely to be independent of each other, we estimated partial correlations: an association measure that allows elimination of such shared correlations between the different features of interest (Sturm et al., 2014). We predicted that tracking of sound intensity would be the most robust, in line with the low-level nature of this feature and based on evidence from previous ECoG studies. However, we expected to see evidence for neuro-oscillatory tracking of key and pulse clarity, at lower frequencies and in limbic areas, corroborating evidence from previous neuro-imaging studies measuring the brain's haemodynamic response (e.g. Alluri et al., 2012).

Materials and Methods

Participants

Data were collected from 11 patients implanted with depth electrodes for pre-surgical evaluation at the Epilepsy unit in the Pitié-Salpêtrière Hospital. However, data from 8 patients (mean age= 31 ± 9.01 years; 2 Males; see Table 1) are presented here due to highly artifacted data in two patients and no contacts in relevant regions in another. The study was approved by the ethical committee for Biomedical Research of Pitié-Salpêtrière Hospital in Paris (CPP Paris VI, INSERM C11-16). All patients gave informed written consent, and clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

Musical stimuli

The stimuli were 4 musical excerpts taken from pieces from the classical and romantic period of the Western tonal music repertoire. The excerpts presented to participants were the opening few minutes of 1) Movement 1 of C.P.E. Bach's Symphony for 2 Violins, Alto and Bass in C major, 2) Movement 2 of Haydn's String Quartet Op. 20 in E flat major, 3) Movement 3 of Mozart's String Quintet No. 4 in G Minor and, finally, 4) Mendelssohn's Songs without words: Op. 19 No. 1. The pieces varied slightly in length in order to allow them to end in musically meaningful ways and such that were all just over 3 minutes long (199, 188, 194, 192 seconds respectively). For consistency, features extracted from the musical stimuli and the iEEG recordings were cut to exactly three minutes (180 seconds) post music onset (see *Intracranial EEG data preprocessing and extraction of oscillatory activity).*

The pieces were based on the advice of musicologists who confirmed, through music theoretical analysis, the wide range the pieces presented with in terms of the three features of interest. Table 2 shows a summary of the keys, the tempo range and instrumentation of all stimuli. To avoid participants needing to adjust the volume during the experiments, due to one or the other excerpt being unpleasantly loud or too soft to hear, a custom written MATLAB script was used to adjust all four audio files, such that they had the same average sound intensity, while keeping the within-piece intensity intact. All participants successfully completed the study (i.e. they all listened to all stimuli) resulting in a total of 32 trials being acquired across participants.

Task

Participants were presented with the musical excerpts using a script programmed in Presentation (NeuroBehavioral Systems) and through external speakers connected to the presentation laptop. Participants were asked to listen to each piece in a relaxed fashion with their eyes closed. At the end of the presentation of each piece, they were required to indicate whether they liked the piece and whether they knew the pieces, using keys labelled YES or NO on the stimulus presentation laptop. All participants reported not knowing any of the pieces but reported liking all pieces. Participants were asked to press any key to continue to the next excerpt whenever they were ready.

Intracranial EEG data acquisition

Participating patients had been stereo-tactically implanted with depth electrodes to localize the epileptogenic focus of their seizures for possible resection. Invasive exploration for each patient had been planned based solely on the brain locations suspected to be at the origin of epileptic seizures. The hypotheses regarding localization of seizures resulted from analysis of data that had previously been gathered in the clinic. These included clinical examination and history, prolonged continuous surface EEG-video, structural MRI and, in some cases, both positron emission tomography and ictal single photon emission computed tomography.

Local field potential data were acquired at a sampling rate of 4000Hz (ATLAS system, Neuralynx, Bozeman, MT, USA) and bandpass filtered between 0.05 and 1000Hz. Electrodes were composed of 4–12 contacts. These contacts were 2.3 mm long, 5 to 10 mm apart, and mounted on a 1.12 mm wide flexible plastic probe (Ad-Tech Medical Instruments, Racine, WI). All electrodes were stereotactically inserted using a Leksell frame (Adam et al., 1996). Electrodes were localised using postoperative computed tomography scans that were coregistered with the 1.5 Tesla MRI scans taken preoperatively. The EpiLoc toolbox (Stereotaxy: Techniques, Images, Models) (http://pf-stim.cricm.upmc.fr) developed at the Institut du Cerveau et de la Moelle Epinière (http://pf-stim.cricm.upmc.fr) was used to automatically determine the MNI coordinates. The location of contacts was visually inspected on corresponding postoperative MRI scans using Slicer (Fedorov et al., 2012).

Intracranial EEG data coverage

Figure 1 shows the distribution of contacts in all 8 patients while Table 3 lists the brain structures that were specifically examined. In the frontal lobe, contacts were located in superior/medial frontal gyrus (SFG: 4 participants: 1, 3, 4 and 5; 43 contacts in total), middle frontal gyrus (MFG: 3 participants: 1, 3, and 5; 17 contacts in total), inferior frontal gyrus (5 participants: 3, 4, 5, 6 and 7; 20 contacts in total) and cingulate gyrus (3 participants: 1, 3 and 4; 21 contacts in total).

In the temporal lobe, contacts were located in superior temporal gyrus including temporal pole (6 participants: 2, 4, 5, 6, 7 and 8; 34 contacts in total), middle temporal gyrus (MTG: 6 participants: 2, 4, 5, 6, 7, 8; 45 contacts in total), inferior temporal gyrus (ITG: 5 participants:

2, 5, 6, 7 and 8; 19 contacts in total) and parahippocampal gyrus (PHG: 6 participants (2, 4, 5, 6, 7, 8; 43 contacts in total).

Intracranial EEG data preprocessing and extraction of oscillatory activity.

Signals from each contact were epoched from –8s to 180s relative to the onset of each excerpt. Data were analyzed in a bipolar montage (after subtraction of the signals of adjacent contacts) to minimize the influence of distant sources (Bollimunta, Chen, Schroeder & Ding, 2009). The Cartesian coordinates (x, y, z) corresponding to the new bipolar signals were calculated as the mid-distance location between the two adjacent contacts.

Time frequency decomposition was carried out using functions from the Fieldtrip toolbox (Oostenveld, Fries, Maris, Schoffelen, 2011) and custom written MATLAB scripts (MATLAB, Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) scripts. Time-courses for each of five frequency bands were estimated by bandpass filtering the signals between 1 and 3 Hz (Delta), 4 and 7Hz (Theta), 8 and 11 Hz (Alpha), 12 and 30 Hz (Beta) and 70 and 170 Hz (High Gamma). A Hilbert transform was applied to the bandpass-filtered time-courses in order to obtain the analytic signal for oscillatory activity in each frequency band. The analytical signals were, in turn, low-pass filtered at 0.5 Hz to smooth the resulting time-courses in line with previous studies (Potes et al., 2012; 2014).

Extraction of musical features

Extraction of all musico-acoustic features was carried out using MIRToolbox (Lartillot & Toivaninen, 2007). Sound intensity, which estimates the instantaneous energy in the signal by taking the square root of sum of the squares of the amplitude, and which acts as a proxy for loudness, was calculated using the *rms* function (Frame: 50ms, hop= 50%, resample to 10Hz).

Pulse clarity, a measure based on the autocorrelation function of extracted tone onsets, and reflecting the clarity of the perceived pulse in a given piece of music (Lartillot, Eerola, Toiviainen, & Fornari, 2008), was calculated using the *mirpulseclarity* function (Frame: 3s, hop= 1.653%, resample to 10Hz). Key clarity, a measure based on the degree of fit of audio-extracted pitch class profiles to Krumhansl-Schmuckler key profiles, and which reflects the degree of absence of both chromaticism and modulation (Toiviainen & Krumhansl, 2003). was calculated using the *mirkey* function (Frame: 3s, hop= 1.653%, resample to 10Hz). As with the iEEG time-courses, all musico-acoustic feature time- courses were low pass filtered at 0.5Hz in line with previous studies (Potes et al., 2012; 2014). Figure 2 shows the entire waveform and power spectra of all musical excerpts along with time-courses for each of the three extracted musico-acoustic features (Sound Intensity, Key Clarity, Pulse clarity).

Statistical Analysis

Correlational analysis was used to examine the relationship between a) the amplitude of oscillatory activity in different frequency bands and the b) time-course of the different musico-acoustic features of interest (Potes et al., 2012; 2014). Rather than estimating full correlation coefficients, we estimated partial correlation coefficients, which are obtained by correlating the residuals that are produced after the variables whose influences are to be eliminated have been regressed out. Here it is important to note that, in contrast to semi-partial correlation/regression analyses, which eliminate the influence of the interfering factor from only one variable, the current analyses eliminated influence from both variables and thus presents a conservative estimation of actual correlations between signals. Such partial spearman correlations were essential for allowing the unique relationship between each oscillatory frequency band and each feature to be determined (Sturm et al., 2014).

As the iEEG and music feature time courses are auto-correlated (i.e. subsequent samples are not independent of each other), standard tests for significance cannot be used. Thus, as proposed in Theiler et al. (1992) and as carried out previously (Sturm et al., 2014), a surrogate target function was generated for each musico-acoustic feature by i) transforming the signal from the time into the frequency domain, ii) randomly permuting the target function's phase spectrum, and then finally iii) recreating the time domain signal from the original spectral amplitudes and the permuted phases. Correlation coefficients between the oscillatory activity time-courses and 1000 surrogate target functions of each musico-acoustic feature were then computed to obtain a distribution of correlation coefficients. Finally, the likelihood of the actually observed correlation coefficient (between the original iEEG and musical feature time-course) was computed by comparing its size to the distribution of correlation coefficients obtained using the surrogate data.

Exploratory analyses in which we compared the strengths of instantaneous correlations with those seen for lagged time-courses (for the high-level features especially) did not suggest the necessity of examining and reporting cross-correlations here. Thus, in line with previous work (Sturm et al., 2014), we report instantaneous (as opposed to cross-) correlation coefficients. Like those authors, we assumed that any time lags between stimulus and brain responses were evened out, and thus not easily apparent, due to the desampling and smoothing processes required for these analyses. Here, we also opted for use of raw time-courses of the three musico-acoustic features rather than derivatives of their time-courses (which would index changes in these features as a function of time). This choice was made to facilitate the comparison of our results with the most relevant previous neuro-imaging and iEEG studies, which also used these features in their raw form (e.g. Alluri et al., 2012; Potes et al., 2014). The choice was also motivated by the fact that absolute levels of key and pulse clarity are most closely aligned with our overarching variable of interest: namely levels of music structural uncertainty, where high clarity relates to high certainty and low clarity relates to high uncertainty.

Results

Figure 3 shows, for illustration purposes, samples of 60s segments of oscillatory time-courses plotted against corresponding 60s time-segments of the three different musico-acoustic features of interest. In turn, Figures 4 and 5 display the anatomical distribution of all contacts that showed significant partial correlations in both lateral (Figure 4) and medial (Figure 5) areas.

As contacts showing significant correlations cannot be averaged across participants, their positions are projected onto a single MNI brain, and a qualitative description of their locations are provided in line with previous approaches (Alluri et al., 2012; Sturm et al., 2014). Results are shown both at a significance level corrected for multiple comparisons (false discovery rate: *fdr* $q < 0.01$) and at a significance level of $p < 0.01$ (uncorrected).

Sound intensity

As can be seen in Figures 4 and 5, delta, alpha and HGP time-courses correlated significantly with sound intensity, following *fdr* correction for family-wise error rates. In the delta range, negative correlations (represented in empty circles) with sound intensity were found in lateral temporal cortices (MTG and ITG) and anterior PHG, while positive correlations (filled circles) were found in more posterior parts of the PHG. In the alpha band, correlations were exclusively negative and were restricted to the STG while, finally, in the HGP range, positive and negative correlations were found in the STG and SFG respectively.

A qualitatively similar but more extensive coverage across areas and frequency bands was shown when considering significance at an uncorrected threshold value of $p < 0.01$. Specifically, negative correlations in the lateral temporal lobe were observed in a more broadband fashion (in theta and beta bands, in addition to delta and alpha bands), while the negative correlations observed between intensity and frontal HGP were also evident at lower frequencies (delta to alpha). Positive correlations with sound intensity were also evident in frontal cortices in delta, beta and gamma bands. Finally, with regard to correlations in medial cortical areas, and in addition to the positive and negative correlations seen in the delta band (which survived *fdr* correction), PHG contacts showed correlations with sound intensity that tended to be negative in lower frequency bands (theta and alpha) but positive in higher ones (beta and HGP bands). In contrast, ACG showed correlations that were positive in the delta band but negative in the HGP band.

Key Clarity and Pulse Clarity

No correlations with key and pulse clarity survived *fdr* correction. However, a number of contacts showed significance at a value of $p < .01$ (uncorrected). With regard to contacts in lateral cortical areas, negative correlations were observed in the temporal pole (delta and beta band), posterior STG (delta) and posterior SFG areas (delta). In contrast, contacts showing positive correlations with key clarity were observed in the IFG HG band (see Figure 4). With regard to correlations of medial areas with key clarity, exclusively negative correlations were observed (see Figure 5). While present at relatively broadband frequencies in the PHG (delta, theta, beta and HGP) they were, however, constrained to higher frequencies in the ACG (HGP only).

Finally, for pulse clarity as for key clarity, contacts were only significant at an uncorrected threshold of $p < .01$. With regard to correlations of lateral cortical areas (see Figure 4), negative correlations were observable in contacts in SFG (theta and HGP) band while, as for key clarity, positive correlations were solely in the IFG (HGP). With regard to medial areas, PHG contacts showed, similarly to key clarity, exclusively negative correlations with pulse clarity, albeit in alpha and beta bands only (see Figure 5).

Discussion

Recent opinion, not least in the field of music neuroscience research, emphasizes the importance of using ecologically valid naturalistic stimuli in laboratory studies (Alluri, et al, 2012; Burunat, Alluri, Toiviainen, Numminen, & Brattico, 2014; Burunat, Toiviainen, Alluri, Bogert, Ristaniemi, Sams, & Brattico, 2016). In keeping with this recommendation, the aim of the current study was to explore and compare, in the context of naturalistic music listening, the neuro-oscillatory correlates of the processing of sound intensity and tonal and rhythmic clarity, extending previous work by interrogating medial cortical areas and by also examining a wider range of frequency bands.

We recorded intracranial SEEG activity from patients implanted with depth electrodes in a range of areas, while they listened to real compositions from the western tonal repertoire. We then examined the extent to which oscillatory power in these areas uniquely correlated with the three musico-acoustic time-courses in a manner independent from each other. We anticipated that the lowest level musico-acoustic feature, namely sound intensity, would be more robustly tracked than higher level ones (Sturm et al., 2014). However, we anticipated that, in line with evidence from neuro-imaging studies (e.g. Alluri et al., 2012), both low and higher-level features would show evidence of being tracked by limbic and frontal areas- due to the involvement of these areas in uncertainty processing and physiological arousal.

Neuro-oscillatory tracking of sound intensity

Our results provided robust confirmation that HGP (70 and 170 Hz) and alpha power in the auditory cortices correlate (positively and negatively respectively) with the sound intensity of music that is being heard (Potes et al., 2014). HGP is often held to index the rate of firing of underlying neurons (Miller, 2010; Ray & Maunsell, 2011) and the positive correlation of STG HGP with sound intensity can be understood as reflecting this low-level sensory processing of auditory stimuli. In contrast, alpha power is known to show reductions in amplitude during sensory processing, and so the negative correlation observed here can also be interpreted as reflecting this sensitivity to low level stimulus structure (Pfurtscheller & Da Silva, 1999).

In terms of underlying mechanisms, specifically, it is believed that alpha power drops reflect interactions between the thalamus and sensory cortices that lead to a desynchronization of cortical alpha in the presence of sensory stimuli (Pfurtscheller & Da Silva, 1999; Haegens, Nácher, Luna, Romo, & Jensen, 2011; Lopes Da Silva, Van Lierop, Schrijer, & Storm Van Leeuwen, 1973). We thus suggest that the STG alpha band modulations observed here reflect thalamo-cortical interactions that are driven by local neuronal activity (HGP) in early auditory cortices. Specifically, we suggest that our findings reflect early stages of auditory processing whereby output from the thalamus travels to primary auditory cortex, before being transmitted to anterior and posterior parts of the STG that are responsible for processing higher level features. Here, it is interesting to note that we also observed, as in previous work (Potes et al, 2014), some evidence for positive tracking of sound intensity in the ventral premotor cortex.

However, unlike the robust correlations observed in STG, these correlations did not reach significance after correction for multiple comparisons.

Also, notable, despite not surviving *fdr* correction, was some support for our hypothesis that sound intensity tracking would be observable in limbic areas. Sound intensity fluctuations reflect, not just the sounding of musical events over time but also, the dynamics of a piece – a device used within performances to express emotional intensity (Palmer, 1997). Our hypothesis that sound intensity tracking would occur in limbic areas was based on previous empirical evidence of a tight relationship between sound intensity fluctuations and emotional arousal during musical listening (Ilie & Thompson, 2006; Juslin 2008; Sloboda & Lehmann 2001). iEEG indices of modulations in cingulate cortex during music listening have been shown in the context of musical pitch expectation violations (Omigie et al., 2019). In that study, they were interpreted as likely reflecting the surprise and emotional arousal (Critchley, Mathias, & Dolan, 2001) induced by such events (Omigie et al., 2019). However, a slightly more direct (albeit not exclusive) link between cingulate cortex activity and emotional arousal induced by sound intensity comes from a study using functional magnetic resonance imaging (fMRI). There, activity in this region was shown to vary in response to variations in the dynamics and tempo of musical stimuli (Chapin, Jantzen, Kelso, Steinberg, & Large, 2010).

Here, we observed both positive delta power tracking and negative HGP tracking of sound intensity in the anterior cingulate gyrus. With regard to the former, it is interesting to note that previous studies have reported increased fronto-central delta power in response to heightened subjective arousal during music listening (Lin, Duann, Chen, Jung, 2010; Mikutta, Maissen, Altorfer, Strik, & König, 2014). Frontocentral power, measured with scalp EEG in those studies, however, may simply reflect the auditory evoked responses to stimulus intensity changes originating from bilateral auditory cortices. In contrast, our finding of a positive correlation between sound intensity and delta power measured directly from the ACG specifically may be taken to reflect the heightened arousal that is caused by high amplitude musical events and passages.

In contrast, our observation that both PHG and ACG gamma activity are negatively correlated with sound intensity (i.e. greater activity in these regions was found at moments of lower intensity) is seemingly counterintuitive given that HGP is typically related to sensory processing of stimuli. Here it is very important to remember, however, that these amodal structures are unlikely to be involved in sensory processing of stimuli per se, and rather more likely to be involved in processing of abstract features of the stimuli. PHG has, for instance, been implicated in the processing of the predictability of sensory stimuli (Strange, Duggins, Penny, Dolan, & Friston, 2005), while ACG activity modulations have been reported as a function of perceptual curiosity, or more generally, information seeking with relation to a stimulus (Jepma, Verdonschot, Van Steenbergen, Rombouts, & Nieuwenhuis, 2012). Thus, based on the rich evidence for a relationship between music dynamics, anticipation and emotional arousal, it is reasonable to speculate that the higher PHG and ACG HGP activity, observed here during low sound intensity passages reflect abstract processing of the stimulus. Specifically, we suggest that it may reflect the state of uncertainty and information seeking that precedes incoming musical events.

Relative strengths of tracking of low and high level musical features using SEEG.

In contrast to the robust tracking of sound intensity exhibited in posterior STG regions, oscillatory tracking of the higher level (key and pulse clarity) musico-acoustic features did not survive correction for multiple comparisons. This was not wholly surprising given the limited success, to date, of studies seeking evidence for neuro-oscillatory tracking of high level music features (Sturm et al., 2014). Sturm and colleagues examined partial correlations of ECoG HGP (70 - 170 Hz) with five musico-acoustic features (namely Sound Intensity, Vocals on/off, Spectral centroid, Harmonic change and Pulse clarity). However, while they reported robust tracking of Vocals on/ off and considerable tracking of Sound intensity and Spectral centroid in peri-sylvian areas, correlations with both pulse and key clarity reached significance in frontal areas in only one out of 10 participants.

Both ours and that previous study (Sturm et al., 2014) employed a necessarily conservative statistical analysis strategy. We not only used partial correlations to look at unique relationships but also used surrogate data to account for autocorrelations in all music feature signals. Thus, one question that remains is whether failure to observe robust neuro-oscillatory tracking of high level musical features is due to limitations of iEEG SNR when examining non-sensory cortical areas (Potes et al., 2014), limitations of estimations of automatically extracted musical features (Lange & Frieler, 2018), or potentially even both. In any case, it is interesting to note that, despite failing to survive *fdr* correction for multiple comparisons, our results showed patterns of correlations that are very much in line with key hypotheses based on previous fMRI findings (e.g. Alluri et al., 2012). Accordingly, these are briefly described in the following section.

Neuro-oscillatory tracking of high level music features.

A key prediction of the current study was that the hierarchical structure of both tonal and rhythmic structure would entail recruitment of syntax processing areas such as the IFG. Hence, a particularly interesting observation is that of a positive correlation of vPMC/ IFG HGP activity with both key and pulse clarity. vPMC has been implicated in sequential processing and prediction of pitch, while IFG has been associated with hierarchical syntactic processing across a range of domains (Schubotz, Friederici & von Cramon., 2000; Schubotz & von Cramon, 2002). A possibility, therefore, is that the presence of a clear hierarchical structure (Lerdahl, & Jackendoff, 1983) at points of high key and pulse clarity leads to greater activity in these syntax processing areas. It is important to note that these correlations are in line with findings from Sturm and colleagues (2014) who also showed the prefrontal cortex to positively correlate with pulse clarity, albeit also in a limited fashion.

Another of our key predictions, based on previous neuroimaging work (Alluri et al., 2012) and the uncertainty entailed by low tonal and rhythmic clarity, was that the ACG and PHG would negatively track key and pulse clarity. Thus, another interesting observation was that cingulate gyrus HGP activity was negatively correlated with key clarity and that PHG modulation tended to be negatively associated with both key and pulse clarity (albeit across different frequency bands). The negative correlation of PHG theta-alpha activity with key clarity would seem to support findings of PHG activity being responsive to dissonance in music (Blood, Zatorre, Bermudez, & Evans, 1999), subjective feelings towards dissonance (Gosselin et al., 2006) and to tonality processing more generally (see Burianova, McIntosh & Grady, 2010; Groussard et al., 2010). In our study, points of low key clarity are characterized, by definition, by high chromaticity and are thus likely to have been experienced by the listeners as relatively more dissonant than points of high key clarity. However, the fact that PHG activity also negatively correlated with pulse clarity supports a more general interpretation of the role it may play in the context of naturalistic music listening. Here, we argue that PHG activity here may track the unpredictability of musical structure rather than either key or pulse structure specifically (Etkin, Egner & Kalisch, 2011). Supporting this more general interpretation of PHG activity, it is interesting to note that Alluri and colleagues (2012) showed negative BOLD activity correlations with pulse (but not key) clarity in the PHG.

Finally, while we anticipated that some amodal areas would show both key and pulse clarity tracking, we also expected some feature-specific effects. Although the failure to survive correction for multiple comparisons precludes detailed discussion of observed differences, it is worth pointing out, given the extensive work claiming a key role of beta oscillatory activity in the tracking of musical beat (Fujioka et al., 2009; Fujioka, et al., 2015), PHG beta power's negative tracking of pulse clarity that we observed here. As previous MEG studies lacked the ability to separate medial from lateral cortical sources of activity within the temporal lobe, we speculate that the negative PHG beta correlations observed here (the most robust of all correlations observed for pulse clarity) may reflect effects previously ascribed solely to lateral cortical areas. Specifically, we suggest it may reflect sustained increases in beta power that emerge as a result of the fewer recognizable beats during moments of low pulse clarity in naturalistic music (Fujioka et al., 2015).

Limitations and future outlook.

Finally, it is important to consider the ways in which future work could provide further insights than were possible in the current study. Here, it is worth pointing out that the absence of detailed behavioural data limited our ability to draw inferences regarding any differences that may exist in the psychological role of the different oscillatory frequency bands. Illustrating the relevance of collecting behavioural data in the context of similar investigations are findings of distinct neuro-oscillatory correlates of predictive coding (predicting what) and predictive timing (predicting when) (Arnal & Giraud, 2012; Chang, Bosnyak & Trainor, 2019). In the current study, we did not use an information rich task as our aim was to examine passive tracking of musico-acoustic stimulus features. However, having participants also continually rate their emotional response with regard to tension, energy and arousal (e.g. Omigie et al., 2019) or with regard to epistemic emotions like surprise or uncertainty (Omigie et al., 2012; Omigie, 2015) would provide important additional information on the subjective experience of the features in question. Such behavioural data, could for instance, be used to distinguish between potentially distinct psychological states that are reflected in correlational effects of contrasting direction (e.g. as reflected in the negative (in HGP) vs positive correlations (in delta) seen in the ACG in response to sound intensity). Together with neural data, such behavioural data would, simply put, provide a much clearer picture of how naturalistic processing of musical features is experienced.

Another way in which future studies could provide further insights would be by adopting a more sophisticated analysis strategy. One important limitation of the current study is the relative disjointedness of the findings resulted in by the currently employed analyses. In examining how each musico-acoustic feature is 'tracked' by each of the different frequency bands, our results are only able to provide a first glimpse into the potential role that different frequency bands and areas may play. Lab studies albeit using simpler stimuli, and with methods that allow quicker collection of comprehensive neural data from healthy participants, have been able to show that processing of tonality and metrical structure require the concerted effort of wide-spread brain networks. Thus, a necessary next step is to seek to determine how different areas and different frequency bands may interact over time in order to carry out an integrated processing of the different features of interest

Such future attempts to observe how different features are tracked during naturalistic music listening could adopt techniques that examine coupling between brain areas (Bastos & Schoffelen, 2016) as well as coupling within brain areas, but across different frequency bands (Tort, Komorowski, Eichenbaum, & Kopell, 2010). In previous work using iEEG, we showed that the processing of consonant versus dissonant chords is reflected in differences in functional connectivity between amygdala, orbitofrontal cortex and STG (Omigie et al., 2015). The current dataset did not provide ample data for such explorations of connectivity between areas of primary interest. However, future studies could, for instance, examine the extent to which processing of specific musical features is reflected in crosstalk between specific brain regions.

In other words, attempts could be made to determine whether connectivity between oscillatory time courses extracted from specific regions correlates with changes in particular features of interest.

Here, it is important to note that the limitations that iEEG presents in terms of incomplete coverage of the brain will continue to prove a challenge for future studies. Appropriate and useful data will be inevitably very slow to collect, requiring as it does, the chance availability of patients presenting with the required patterns of electrode implantation. However, the promise of high SNR neural activity as possible with iEEG, combined with the promise of naturalistic music as a stimulus for studying neural processing, would make all such research endeavours highly invaluable.

Conclusions

In closing, the current study sought to extend previous work examining the neuro-oscillatory correlates of the processing of sound intensity, tonality and rhythmic clarity, by interrogating subcortical as well as cortical areas and a wider range of oscillatory frequency bands. Our results confirmed that robust tracking of sound intensity is observable in the alpha and high gamma oscillatory activity of auditory cortices, likely reflecting cortico-thalamic interactions driven by neural firing. However, in providing only weak evidence for the tracking of high compared to low level musical features, our study would seem to confirm challenges inherent to combining free listening paradigms with the examination of neural correlates of high level musical features (Sturm et al., 2014). Future work will be able to provide further insights by acquiring rich behavioural data, and by developing yet more sensitive analysis strategies.

Acknowledgements.

We would like to thank all patients who very kindly took part in this study.

References

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2010). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex,* 21(7), 1507-1518.

Adam, C., Clemenceau, S., Semah, F., Hasboun, D., Samson, S., Aboujaoude, N., Samson, Y., & Baulac, M. (1996). Variability of presentation in medial temporal lobe epilepsy: a study of 30 operated cases. *Acta Neurologica. Scandinavica.* 94, 1–11.

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage*, 59(4), 3677-3689.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. Trends in cognitive sciences, 16(7), 390-398.

Bastos, A. M., & Schoffelen, J. M. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. Frontiers in systems neuroscience, 9, 175.

Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance,* 21(6), 1278.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature neuroscience,* 2(4), 382.

Bollimunta, A., Chen, Y., Schroeder, C. E., & Ding, M. (2008). Neuronal mechanisms of cortical alpha oscillations in awake-behaving macaques. *Journal of Neuroscience,* 28, 9976–9988.

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain research,* 1117(1), 162-174.

Burianova, H., McIntosh, A.R., & Grady, C.L., (2010). A common functional brain network for autobiographical, episodic, and semantic memory retrieval. *Neuroimage,* 49, 865–874.

Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex,* 57, 254-269.

Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., & Brattico, E. (2016). The reliability of continuous brain responses during naturalistic listening to music. *NeuroImage,* 124, 224-231.

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nature reviews neuroscience,* 13(6), 407.

Chapin, H., Jantzen, K., Kelso, J. S., Steinberg, F., & Large, E. (2010). Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PloS one,* 5(12), e13812.

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage,* 32(4), 1771-1781.

Creutzfeldt, O., & Ojemann, G. (1989). Neuronal activity in the human lateral temporal lobe. III. Activity changes during music. Experimental Brain Research, 77(3), 490-498.

Critchley, H. D., Mathias, C. J., Josephs, O., O' Doherty, J., Zanini, S, Dewar, BK, Cipolotti, L, Shallice, T, & Dolan R. J. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain*, 126, 2139-52.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, 29, 537-45.

Di Pietro, M., Laganaro, M., Leemann, B., & Schnider, A. (2004). Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia*, 42(7), 868-877.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences,* 112(45), E6233-E6242.

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience,* 13(3), 533-553.

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in cognitive sciences*, *15*(2), 85-93.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J-C, … Kikinis, R. (2012). 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging,* 30, 1323-41.

Fries, W., & Swihart, A. A. (1990). Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia*, 28(12), 1317-1323.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of neuroscience*, 35(45), 15187-15198.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, 1169(1), 89-92.

Gosselin, N., Samson, S., Adolphs, R., Noulhiane, M., Roy, M., Hasboun, D., ... & Peretz, I. (2006). Emotional responses to unpleasant music correlates with damage to the parahippocampal cortex. *Brain,* 129(10), 2585-2592.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage,* 47(4), 1894-1903.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience,* 29(23), 7540-7548.

Grahn, J. A., & Rowe, J. B. (2012). Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. *Cerebral cortex,* 23(4), 913-921.

Groussard, M., La Joie, R., Rauchs, G., Landeau, B., Chetelat, G., Viader, F., ... & Platel, H. (2010). When music and long-term memory interact: effects of musical expertise on functional and structural plasticity in the hippocampus. *PLoS One,* 5(10), e13225.

Haegens, S., Nácher, V., Luna, R., Romo, R., & Jensen, O. (2011). Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences of the United States of America.* 108, 19377–19382.

Harrington, D. L, Haaland, K., & Hermanowicz, N. (1998). Temporal processing in the basal ganglia. *Neuropsychology*, 12, 3–12.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R., (2004). Intersubject synchronization of cortical activity during natural vision. *Science,* 303, 1634–1640.

Huron, D. B. (2006). Sweet anticipation: Music and the psychology of expectation. MIT press.

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. Music Perception: An Interdisciplinary Journal, 23(4), 319-330.

Jacobsen, T., Horenkamp, T., & Schröger, E. (2003). Preattentive memory-based comparison of sound intensity. *Audiology and Neurotology*, 8(6), 338-346.

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, 298(5601), 2167-2170.

Janata, P., Tillmann, B., & Bharucha, J.J., (2002). Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective and Behavioural Neuroscience*. 2, 121–140.

Janata, P. (2005). Brain networks that track musical structure. *Annals of the New York Academy of Sciences,* 1060, 111–12

Jepma, M., Verdonschot, R. G., Van Steenbergen, H., Rombouts, S. A., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. Frontiers in behavioral neuroscience, 6, 5.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological review*, 96(3), 459.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological science*, 13(4), 313-319.

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale,* 51(4), 271.

Juslin, P. N., & Vastfjall, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behavioural Brain Sciences.* 31, 559–575.

Koelsch, S. (2012). Brain and Music. West Sussex, UK: John Wiley & Sons, Ltd.

Koelsch, S. (2014). Brain correlates of music-evoked emotion. Nature Reviews Neuroscience, 15(3), 170-180.

Koelsch, S., Gunter, T. C., Schroeger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: an ERP study. *NeuroReport*, 12(7), 1385-1389.

Lange, E. B., & Frieler, K. (2018). Challenges and opportunities of predicting musical emotions with perceptual and automatized features. *Music Perception: An Interdisciplinary Journal, 36(2), 217-242.*

Large, E. W. & Jones, M. R. (1999). The dynamics of attending: How we track time varying events. *Psychological Review,* 106, 119-159.

Lartillot, O., & Toiviainen., P. (2007). A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects.*

Lartillot, O., Eerola, T., Toiviainen, P., & Fornari, J. (2008, September). Multi-Feature Modeling of Pulse Clarity: Design, Validation and Optimization. In ISMIR (pp. 521-526).

Lehne, M., Rohrmeier, M., & Koelsch, S. (2014). Tension-related activity in the orbitofrontal cortex and amygdala: an fMRI study with music. *Social Cognitive and Affective Neuroscience*, 9, 1515-1523.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception: An Interdisciplinary Journal,* 1(2), 229-252.

Lin, Y. P., Duann, J. R., Chen, J. H., Jung, T. P. (2010). Electroencephalographic dynamics of musical emotion perception revealed by independent spectral components. Neuroreport 21, 410–415.

Lockwood, A. H., Salvi, R. J., Coad, M. L., Arnold, S. A., Wack, D. S., Murphy, B. W., & Burkard, R. F. (1999). The functional anatomy of the normal human auditory system: responses to 0.5 and 4.0 kHz tones at varied intensities. *Cerebral Cortex*, *9*(1), 65-76.

Lopes Da Silva, F.H., Van Lierop, T. H. M. T, Schrijer, C. F., & Storm Van Leeuwen, W. (1973). Organization of thalamic and cortical alpha rhythms: spectra and coherences. *Electroencephalography and Clinical Neurophysiology.* 35(6), 627–639.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience.* 4, 540.

Mencke, I., Omigie, D., Wald-Fuhrmann, M., & Brattico, E. (2018). Atonal Music: Can uncertainty lead to pleasure? Frontiers in neuroscience, 12, 979.

Meyer, L. B. (1956). Emotion and meaning in music. Chicago: University of Chicago Press.

Mikutta, C. A., Dürschmid, S., Bean, N., Lehne, M., Lubell, J., Altorfer, A., ...& Koelsch, S. (2015). Amygdala and orbitofrontal engagement in breach and resolution of expectancy: A case study. *Psychomusicology*. 25, 357.

Mikutta, C. A., Maissen, G., Altorfer, A., Strik, W., & König, T. (2014). Professional musicians listen differently to music. Neuroscience, 268, 102-111.

Miller, K. J. (2010). Broadband spectral change: evidence for a macroscale correlate of population firing rate? *Journal of Neuroscience.* 30(19):6477–6479.

Narmour, E. (1990). The analysis and cognition of basic melodic structures: The implication realization model. Chicago: University of Chicago Press.

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. Neuropsychologia, 50(7), 1483-1493.

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, 51(9), 1749-1762.

Omigie, D. (2015). Dopamine and epistemic curiosity in music listening. Cognitive neuroscience, 6(4), 222-224.

Omigie, D., Dellacherie, D., Hasboun, D., George, N., Clement, S., Adam, C., & Samson, S. (2015). An intracranial EEG study of the neural dynamics of musical valence processing. *Cerebral Cortex.* 25, 4038-4047.

Omigie, D., Dellacherie, D., Hasboun, D., Clement, S., Adam, C., & Samson, S. (2015). Intracranial markers of emotional valence processing and judgments in music. *Cognitive Neuroscience.* 6, 16–23.

Omigie, D., Pearce, M., Lehongre, M., Hasboun, D., Navarro, V., Adam, C., Samson, S. (2019). Intracranial Recordings and Computational Modeling of Music Reveal the Time Course of Prediction Error Signaling in Frontal and Temporal Cortices. *Journal of Cognitive Neuroscience,* 31, 855-873.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J. M. (2011) FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data.

Computational Intelligence and Neuroscience Volume 2011, Article ID 156869

Palmer, C. (1997) Music performance. *Annual Review of Psychology*, 48, 115–138.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology, 110(11), 1842-1857.

Potes, C., Gunduz, A., Brunner, P., & Schalk, G. (2012). Dynamics of electrocorticographic (ECoG) activity in human temporal and frontal cortical areas during music listening. *Neuroimage*, 61(4), 841-848.

Potes, C., Brunner, P., Gunduz, A., Knight, R. T., & Schalk, G. (2014). Spatial and temporal relationships of electrocorticographic alpha and gamma activity during auditory processing. *Neuroimage*, 97, 188-195.

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience,* 4, 317–23.

Ray, S., & Maunsell, J. H. R. (2011) Different Origins of Gamma Rhythm and High-Gamma Activity in Macaque Visual Cortex. *PLoS Biology.* 9(4), 15

Röhl, M., & Uppenkamp, S. (2012). Neural coding of sound intensity and loudness in the human auditory system. Journal of the Association for Research in Otolaryngology, 13(3), 369-379.

Rohrmeier, M., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. International Journal of Psychophysiology, 83, 164-175.

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology,* 44(2), 293-304.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H. J., ... & Friederici, A. D. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *Neuroimage.* 64, 134-146.

Schadow, J., Lenz, D., Thaerig, S., Busch, N. A., Fründ, I., & Herrmann, C. S. (2007). Stimulus intensity affects early sensory processing: sound intensity modulates auditory evoked gamma-band activity in human EEG. *International Journal of Psychophysiology,* 65(2), 152-161.

Schubotz, R. I., Friederici, A. D., & Von Cramon, D. Y. (2000). Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage,* 11(1), 1-12.

Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *Neuroimage,* 15(4), 787-796.

Schwartze, M., Keller, P. E., Patel, A. D., & Kotz, S. A. (2011). The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the detection of tempo changes. *Behavioural brain research,* 216(2), 685-691.

Seeley, W. W., V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, et al. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience.* 27, 2349–2356.

Sloboda, J. A., & Lehmann, A. C. (2001) Tracking performance correlates of changes in perceived intensity of emotion during different interpretations of a Chopin piano prelude. *Music Perception*, 19, 87–120

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of cognitive neuroscience,* 18(8), 1380-1393.

Strange, B. A., Duggins, A., Penny, W., Dolan, R. J., & Friston, K. J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable*? Neural Networks,* 18(3), 225-230.

Sturm, I., Blankertz, B., Potes, C., Schalk, G., & Curio, G. (2014). ECoG high gamma activity reveals distinct cortical representations of lyrics passages, harmonic and timbre-related changes in a rock song. *Frontiers in human neuroscience,* 8.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena,* 58(1-4), 77-94.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., & von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *Neuroimage*, 31(4), 1771-1782.

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception,* 32(6), 741-766.

Tort, A. B., Komorowski, R., Eichenbaum, H., & Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. Journal of neurophysiology, 104(2), 1195-1210.

Vuust, P., & Witek, M. A. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in psychology,* 5, 1111.

Zanto, T. P., Snyder, J.S., & Large, E.W. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology,* 2, 221–231.

Figures

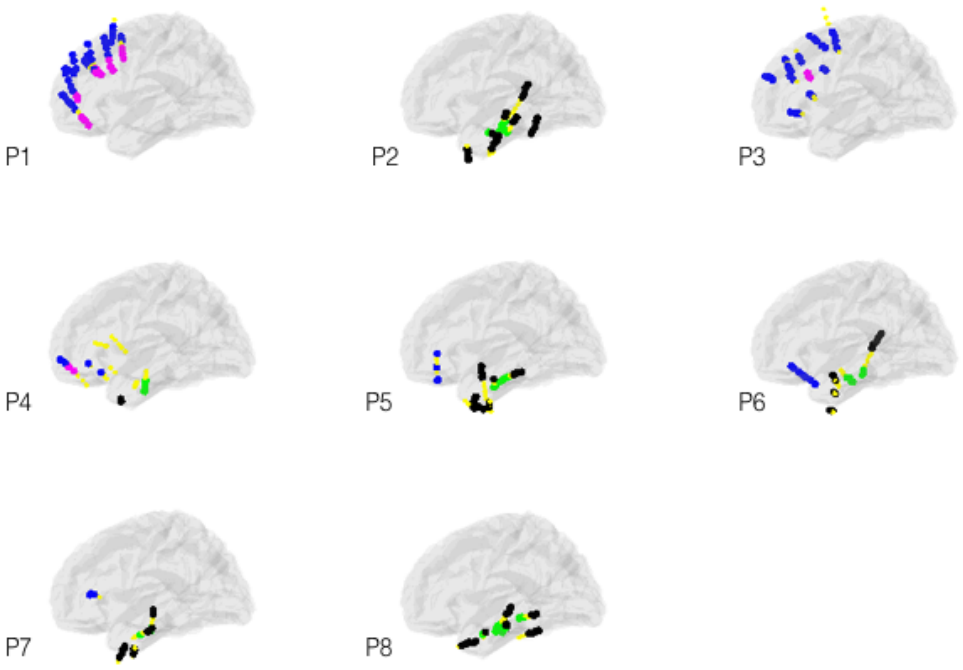

Figure 1: Distribution of recording sites. MNI brain templates showing the coverage of contacts over frontal and temporal areas in 8 patients (4 with frontal and temporal contacts, 2 with temporal-only contacts and 2 with frontal-only contacts). All contacts are projected on the left hemisphere. Data from contacts in the lateral temporal and frontal lobes that were analysed are depicted in black and blue filled circles respectively. Data from contacts in the medial temporal and frontal lobes that were analysed are depicted in green and pink filled circles respectively. Finally, data from contacts that were difficult to localize precisely or which contained artefacted signal (and were thus not analysed), are depicted in small yellow filled circles.

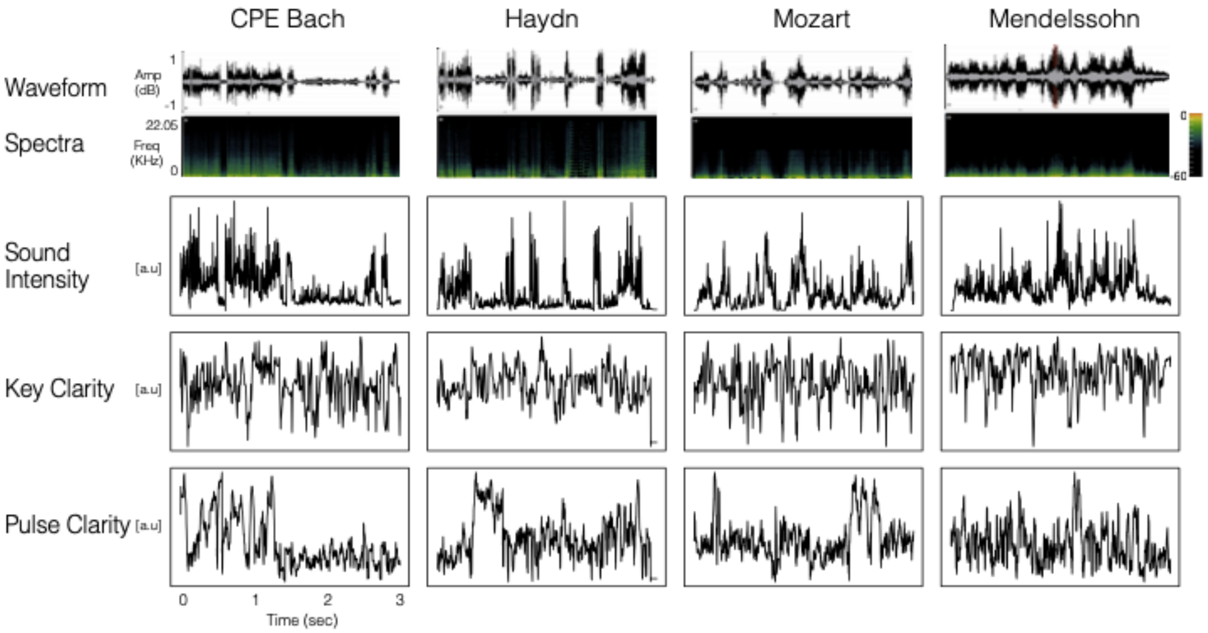

Figure 2: Musical stimuli. Plots showing the acoustic waveforms and spectrograms of the four musical excerpts, along with each of the three musico-acoustic features (Sound Intensity, Key Clarity, Pulse clarity) extracted from them.

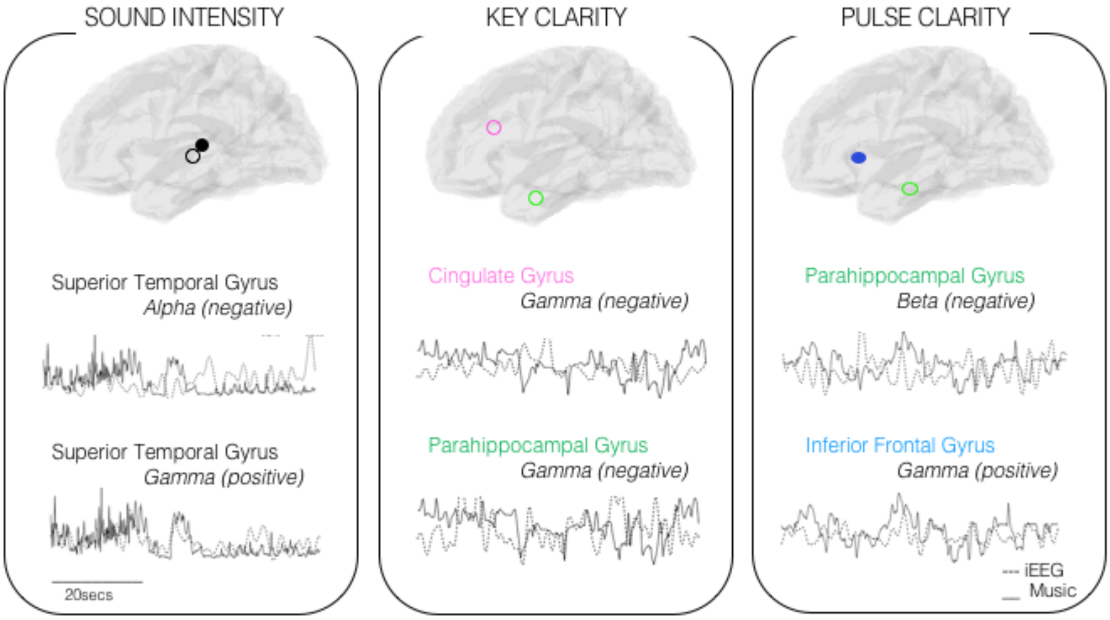

Figure 3: Sample segments of neural oscillatory time-courses for comparison with the three different features of interest. Normalized HGP and alpha oscillatory time courses extracted from contacts located in the STG are shown positively and negatively correlated with sound intensity respectively (left column). Normalized HGP oscillatory time courses extracted from contacts located in PHG and ACG are shown negatively correlated with Key clarity (middle column). Finally, normalized Beta and HGP oscillatory time courses extracted from contacts located in the PHG and IFG are shown negatively and positively correlated with Pulse clarity respectively (right column).

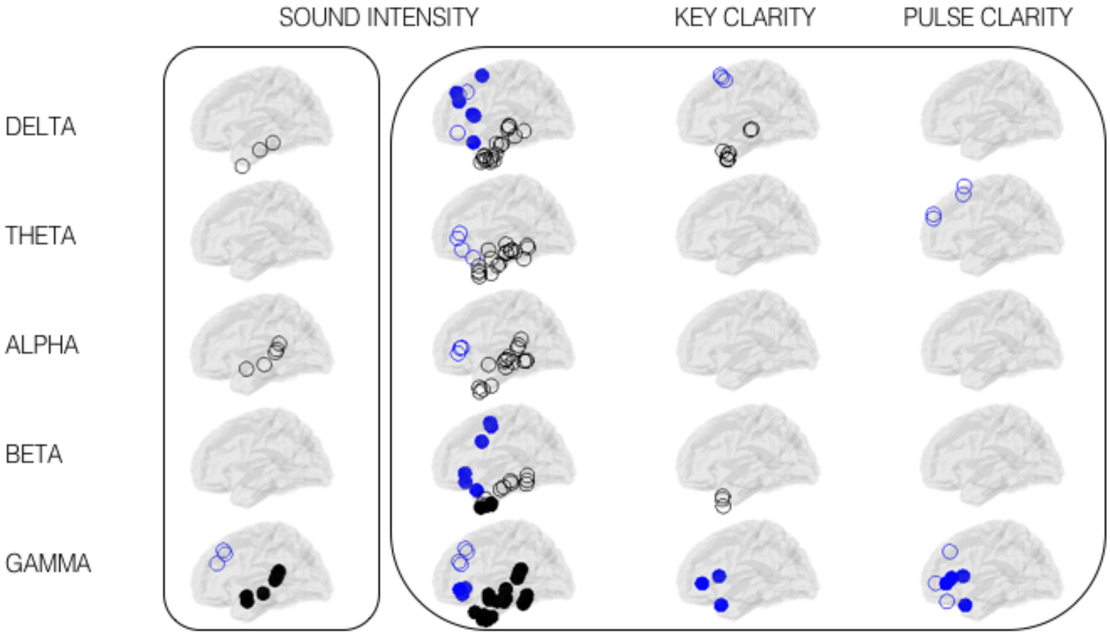

Figure 4: Neuro-oscillatory tracking in lateral cortical areas. Brain templates showing the location of contacts for which neuro-oscillatory time-courses were positively (filled circles) and negatively (empty circles) correlated with sound intensity, key clarity and pulse clarity in the lateral temporal (black) and frontal lobes (blue).

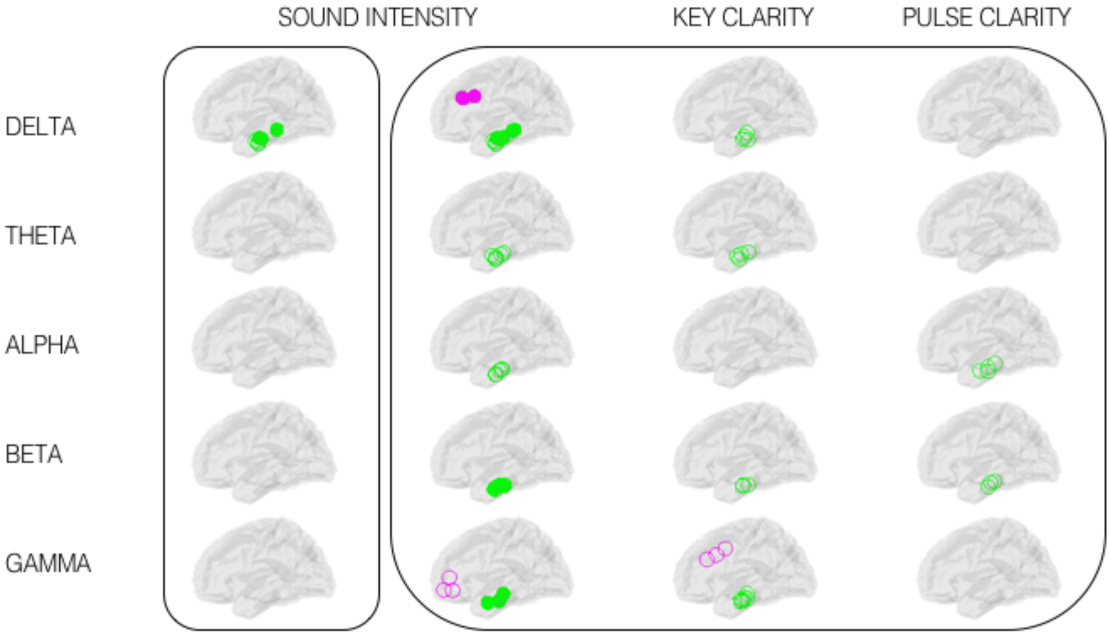

Figure 5: Neuro-oscillatory tracking in medial cortical areas. Brain templates showing the location of contacts for which neuro-oscillatory time-courses were positively (filled circles) and negatively (empty circles) correlated with sound intensity, key clarity and pulse clarity in the parahippocampal (green) and cingulate gyrus (pink)

Table 1. Demographic and background information on participating patients

Patient	Age	Sex	Handedness	Seizure frequency (per month)	Age of epilepsy onset	Epileptic focus (seizure onset regions)	Lesion	Implantation Hemisphere	Medication
1	30	F	Right	28	16	Superior frontal lobe (BA 6, 8)	Dilation of the right temporal horn	Right frontal lobe	Keppra, Lamictal
2	32	M	Right	1	24	Left temporopolar cortex	No lesion	Left temporal lobe	Keppra, Lamictal, Tegretol
3	22	F	Right	12	8	Right medial frontal lobe	No lesion	Right frontal lobe	Dépakine, Lamictal
4	34	F	Right	10	9	Parahippocampal gyrus	No lesion	Left fronto-temporal lobe	Lamictal, Zonégran
5	40	F	Right	28	23	Inferior temporal lobe	No lesion	Left fronto-temporal lobe	Vimpat, Diphantoine
6	46	F	Right	-	-	Left temporopolar cortex, middle temporal cortex.	Right hippocampal sclerosis	Right fronto-temporal and left temporal lobe	Lamictal, Keppra, Urbanyl,
7	24	F	Right	12	3	Left temporopolar cortex	No lesion	Left frontotemporal lobe	Tégrétol, Lamictal
8	20	M	Right	4	8	Left Hippocampus	MRI: Hypersignal in Right subcortical paraventricular region	Bilateral temporal lobe	Trileptal, Vimpat, Zonégran, Lamictal, Urbanyl

Table 2: Descriptive statistics of the pieces used in the experiment

	Musical Features		
Pieces	Local keys used	Tempo range	Instrumentation
CPE Bach	C major	60 to 146 bpm	Violins, Viola, Basso (Cello),
Haydn	E flat major	57 to 100 bpm	Violins, Viola, Cello
Mozart	E flat major	62 to 162 bpm	Violins, Violas, Cello
Mendelssohn	E major	70 to 180 bpm	Piano

Table 3. Locations of examined regions.

Lobe	Region	Sub region	Brodmann areas
Frontal	Lateral	Superior/Medial Frontal Gyrus	BA 6, BA8, BA9, BA10
		Middle Frontal Gyrus	BA 6, BA8, BA9
		Inferior Frontal Gyrus	BA44, BA45, BA46, BA47
	Medial	Cingulate Gyrus	BA24, BA 32, BA31
Temporal	Lateral	Superior Temporal Gyrus	BA22, BA38
		Middle Temporal Gyrus	BA 20, BA21, BA38
		Inferior Temporal Gyrus	BA20, BA21, BA37, BA36
	Medial	Parahippocampal Gyrus	BA28, BA34, BA35,

BA: Brodmann area.